High-level production of nervonic acid in the oleaginous yeast *Yarrowia lipolytica* by systematic metabolic engineering

Hang Su[#1,5], Penghui Shi[#1,4], Zhaoshuang Shen[#1,4], Huimin Meng[1], Ziyue Meng[1,5], Xingfeng Han[1], Yanna Chen[2], Weiming Fan[2], Yun Fa[1,4], Chunyu Yang[3], Fuli Li*[1,4], and Shi'an Wang*[1,4]

1. Shandong Provincial Key Laboratory of Synthetic Biology, Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China.
2. Zhejiang Zhenyuan Pharmaceutical Co., LTD, Shaoxing 312071, China
3. State Key Laboratory of Microbial Technology, Institute of Microbial Technology, Shandong University, Qingdao, 266237, China
4. Shandong Energy Institute, Qingdao 266101, China.
5. University of Chinese Academy of Sciences, Beijing 100039, China.

***Corresponding author:**

Shi'an Wang, Tel.: +86-532-80662682; Fax: +86-532-80662778.

E-mail: wangsa@qibebt.ac.cn

Fuli Li, Tel.: +86-532-80662656; Fax: +86-532-80662778.

E-mail: lifl@qibebt.ac.cn

Postal address: No.189 Songling Road, Laoshan District, Qingdao, Shandong

Province, China, Postcode 266101

[#]**These authors contributed equally: Hang Su, Penghui Shi, and Zhaoshuang Shen**

Abstract

Brain and neurological diseases are influencing more than one billion world's people. Nervonic acid (cis-15-tetracosenoic acid, C24:1 Δ15) benefits the treatment of neurological diseases and the health of brain. Currently, the sources of nervonic acid are limited to the seeds of a couple of plants. In this study, we employed the oleaginous yeast *Yarrowia lipolytica* to overproduce nervonic acid oil by systematic metabolic engineering. First, engineering the fatty acid elongation (FAE) pathway by expressing a heterologous β-ketoacyl-CoA synthase gene *CgKCS* enabled the production of nervonic acid in *Y. lipolytica.* Second, modulation of endogenous pathways by expressing a C16:0-acyl-CoA preferred fatty acid elongase gELOVL6 together with a C18:0-acyl-CoA preferred fatty acid desaturase MaOLE2 increased the content of nervonic acid in total fatty acids (TFA). Third, iterative expression of *CgKCS*, *gELOVL6* and *MaOLE2* at the genomic loci of *rDNA*, *FAD2*, *TGL4*, *GSY1* and *SNF1* dramatically improved the production of nervonic acid. Fourth, the biosynthesis of both nervonic acid and lipids were further enhanced by expression of the MaOLE2-CgKCS fusion protein and glycerol-3-phosphate acyltransferases (GPAT) and diacylglycerol acyltransferases (DGAT) from *Malania oleifera* in the endoplasmic reticulum (ER) membrane. Fifth, an ER structure regulator YlINO2 was identified in *Y. lipolytica* and the overexpression of YlINO2 led to a 39.3% increase in lipid production. Next, pilot-

scale fermentation in 50-L reactor using the strain YLNA9 exhibited a lipid titer of 96.7 g/L and a nervonic acid titer of 17.3 g/L, the highest reported titer to date for *de novo* nervonic acid production. We also found that disruption of the AMP-activated S/T protein kinase *SNF1* increased the ratio of nervonic acid (C24:1) to lignoceric acid (C24:0) by 61.6% and a ratio of 3.5:1 (nervonic acid to lignoceric acid) was achieved in the strain YLNA10. Finally, a proof-of-concept purification and separation of nervonic acid were performed and the purity of it reached 98.7%. This study suggested that oleaginous yeasts are attractive hosts for the cost-efficient production of nervonic acid and possibly other very long-chain fatty acids (VLCFAs).

Keywords: Fatty acid; nervonic acid; lipid; oleaginous yeast; metabolic engineering

1. Introduction

Brain and neurological diseases affect nearly one in six of the world's population according to the American Brain Foundation (www.americanbrainfoundation.org/). Nervonic acid (cis-15-tetracosenoic acid, C24:1 Δ15), an omega-9 very long-chain monounsaturated fatty acid (VLCMFA), is rich in the white matter and myelin sheath of human brain. More and more studies show that nervonic acid benefits neurological health (Wadhwani et al., 2011; Li et al., 2019) and has therapeutic potential for adrenoleukodystrophy and multiple sclerosis (Lewkowicz et al., 2019; Terluk et al., 2022). Moreover, nervonic acid levels in specific tissues are associated with Alzheimer's disease (Astarita et al., 2011), Parkinson's disease (Hu et al., 2021), cognition (Song et al., 2022), mood symptoms (Kageyama et al., 2018; Kageyama et al., 2021), cardiovascular death (Delgado et al., 2017) and even obesity (Keppley et al., 2020).

The source of nervonic acid has attracted increasing interests due to its tremendous potentials for nutraceutical and pharmaceutical applications. Nervonic acid accounts for more than 40% of the total fatty acids (TFA) in seeds of the rare plant *Malania oleifera* and the herbaceous plants *Tropaeolum speciosum* and *Cardamine graeca* (Chivandi et al., 2008; Tang et al., 2013; Li et al., 2019; Liu et al., 2021). The oil extracted from *Acer truncatum* seeds containing about 5% of nervonic acid is sold as a nutraceutical in China (Qiao et al., 2018). Although a couple of plants have seeds rich in nervonic acid, supply of nervonic acid by these plants is expensive due to limitations of scale-up cultivation, climate-dependence, and harvesting cost (Fan et al., 2018a).

Nervonic acid has also been identified in a few microalgae and filamentous fungi. The oleaginous microalgae *Mychonastes afer* HSO-3-1 and the filamentous fungal *Mortierella capitata* can synthesize nervonic acid accounting for 3.8% and 6.9% of TFA, respectively (Umemoto et al., 2014; Yuan et al., 2011). However, the photoautotrophic microalgae *M. afer* produces low biomass and *M. capitate* is a rare fungus lacking genetic tools, which prevent them being used as cell factories for production of nervonic acid (Xu et al., 2018; Li et al., 2020).

In recent years, the oleaginous yeast *Yarrowia lipolytica* has been used as a valuable host for the production of both lipid and nonlipid chemical products by metabolic engineering (Xu et al., 2016; Qiao et al., 2017; Holkenbrink et al., 2020; Luo et al., 2020; Muhammad et al., 2020). This yeast is generally regarded as safe (GRAS), grows fast, and has advantages in supply of acetyl-CoA and NADPH for anabolism (Abdel-Mawgoud et al., 2018; Miller and Alper, 2019). The metabolic characteristic of *Y. lipolytica* makes it suitable for production of fatty acids, such as polyunsaturated eicosapentaenoic acid (EPA) (Xue et al., 2013), gamma-linolenic acid (GLA) (Sun et al., 2017) and docosahexaenoic acid (DHA) (Gemperlein et al., 2019). Laudably, DuPont successfully synthesized EPA (C22:6, ω-3) from C18:1-acyl-CoA by elongation and desaturation in *Y. lipolytica*, and they commercialized EPA-rich oil for food and EPA-rich *Y. lipolytica* biomass for feed applications (Xue et al., 2013; Xie et al., 2015). In the context, *Y. lipolytica* is expected to be an appropriate host for production of nervonic acid.

The fatty acid profiles in triglycerides produced by *Y. lipolytica* mainly consist of

palmitic acid (C16:0), palmitoleic acid (C16:1, ω-7), and oleic acid (C18:1, ω-9) (Tai and Stephanopoulos, 2013). Several aspects should be optimized for high-production of nervonic acid by *Y. lipolytica*. First, β-ketoacyl-CoA synthases (KCS) are among the key fatty acid elongation enzymes for synthesis of VLCMFA from C18:1-acyl-CoA. The substrate preference and specificity of KCS may differ in various hosts. For example, the KCS ACJ61777.1 (GenBank No. EU871787) identified in the biennial grass *Lunaria annua* used C22:1-acyl-CoA as substrate to synthesize nervonic acid when expressed in the Ethiopia mustard *Brassica carinata* (Guo et al., 2009), while it elongated C18:1-acyl-CoA to produce mainly eicosenoic acid (C20:1, ω-9) and erucic acid (C22:1, ω-9) when expressed in the yeast *Rhodosporidium toruloides* (Fillet et al., 2017). Second, most fatty acids exist in lipid forms in oleaginous microbes and therefore the selectivity of esterifying enzymes toward acyl-CoAs affects the fatty acid profiles in lipids. Third, subcellular engineering of the endoplasmic reticulum (ER) can improve the production of VLCMFA because esterification and elongation of long-chain fatty acids occur in ER (Chen et al., 2016; Kim et al., 2019).

In this study, we employed the oleaginous yeast *Y. lipolytica* as a host to de novo synthesizing nervonic acid. The metabolic pathways for overproducing nervonic acid were designed and depicted in Fig. 1. High-level production of nervonic acid was achieved by multiplex engineering strategies. This study showed the potential of oleaginous yeasts used for cost-efficient production of nervonic acid and possibly other VLCFAs.

2. Materials and methods

2.1. Strains and plasmids

The *Y. lipolytica* strains po1g (Yeastern Biotech, Taibei, Taiwan, China) and po1g-G3 (Li et al., 2020) were used as the hosts. Yeast strains constructed in this study were listed in Table S1. Plasmids (Table S2) were constructed using Gibson assembly method and verified by DNA sequencing. Primers used were listed in Table S3. DNA sequences of heterologous genes were codon-optimized and synthesized by Beijing Genomics Institute (BGI), China.

All plasmids constructed in this study were assembled *via* Gibson assembly. The plasmid backbones and DNA fragments were amplified using the KAPA HiFi HotStart PCR kit (Kapa Biosystems, Boston, USA) and the primers described in Table S3. The assembly was performed at 50 °C for 1 h. The assembled mixture was used to transform the *Escherichia coli* Trans1-T1 component cells (TransGen Biotech, Beijing, China) for plasmid amplification. Positive clones were verified by colony PCR and DNA sequencing.

The *Y. lipolytica* native promoter TEFin (Tai and Stephanopoulos, 2013) was used in controlling the expression of the *CgKCS* gene. The organelle target signals KDEL for ER, SKL for peroxisome, and the CoxIV signal for mitochondria were fused with *CgKCS* (Avalos et al., 2013; Xu et al., 2016). To raise the copy numbers of *CgKCS*, three *CgKCS*-expressing cassettes were assembled together, resulting in a plasmid pYL-3sKCS (Table S2). The expression of *gELOVL6* and *MaOLE2* was driven by the inducible promoter yat1 under oleaginous conditions (Xue et al., 2013). To overexpress

138 *CgKCS*, *gELOVL6* and *MaOLE2*, the expression cassettes of the three genes were
139 ligated together with backbone fragments, generating the plasmids pYL-3grDNA and
140 pYL-3gD17 (Table S2).

141 Expression of genes by homologous recombination were performed at the specific
142 gene loci *FAD2* (fatty acid desaturase 2), *TGL4* (lipase 4), *rDNA*, *GSY1* (glycogen
143 synthase), *SNF1* (AMP-activated S/T protein kinase), *D17* and *PEX10* (peroxisomal
144 membrane protein) according to previous studies (Xue et al., 2013; Schwartz et al.,
145 2017; Lazar et al., 2018). The plasmids pYL-rDNA-CgKCS, pYL-GSY1-CgKCS and
146 pYL-SNF1-CgKCS were separately used to express one copy of CgKCS at the *rDNA*,
147 *GSY1* and *SNF1* locus. The plasmids pYL-2gFAD2 and pYL-2gTGL4 were adopted to
148 express two copies of *CgKCS* at the *FAD2* locus and the *TGL4* locus, respectively. The
149 fusion protein of CgKCS and MaOLE2 in the plasmid pYL-D17-gMaolE2 was
150 expressed at the *D17* locus. To express the GPATs and DGAT2s from *M. oleifera*, the
151 plasmids pYL-PEX10-Mo-35, pYL-PEX10-Mo-49, pYL-PEX10-Mo-88 and pYL-
152 PEX10-Mo-90 were constructed and the gene expression cassettes were inserted at the
153 *PEX10* locus. To express the putative regulators INO2 and INO4 of lipid synthesis, the
154 plasmids pYL-PEX10-YlINO2 and pYL-PEX10-YlINO4 were constructed and the
155 *PEX10* locus was employed to accommodate the gene expression cassettes.

156 **2.2. Culture media and conditions**

157 YPD medium, containing 20 g/L glucose, 20 g/L peptone, and 10 g/L yeast extract,
158 was used in regular cultivation of *Y. lipolytica* with rotary shaking at 250 rpm and 28
159 ºC. YNB medium used for selection of genetic transformants was made with 20 g/L

160 glucose, 1.7 g/L yeast nitrogen (without amino acids and ammonium sulfate) (Sangon
161 Biotech, Shanghai, China), 5 g/L ammonium sulfate, and 0.69 g/L CSM-Ura or CSM-
162 Leu (MP Biomedicals, Solon, OH).

163 Shake flask cultures were carried out using the fermentative medium composed of
164 150 g/L glucose, 6 g/L yeast nitrogen, and 12 g/L ammonium sulfate. A single colony
165 was inoculated into 5 ml YNB medium and cultivated at 28 ºC with shaking at 250 rpm
166 for 24 h. Afterwards, 500 μL of the seed culture was inoculated into 5 ml fresh YNB
167 medium and cultivated at 28 ºC with shaking at 250 rpm for 24 h. Then, precultures
168 were inoculated into 30 ml of fermentative medium in 250 ml shake flasks to an optical
169 density (OD_{600}) of 0.08 and incubated at 250 rpm and 28 ºC for 144 h.

170 Bioreactor fermentation was carried out in a 50-L bioreactor (HND Bio-
171 engineering Equipment, Jiangsu, China). The fermentative medium contained 150 g/L
172 glucose, 6 g/L yeast extract, 12 g/L ammonium sulfate, 1.5 g/L $MgSO_4 \cdot 7H_2O$, 6 g/L
173 KH_2PO_4, 3 g/L $Na_2HPO_4 \cdot 12H_2O$, 3 mg/L $CaCl_2 \bullet 2H_2O$, 1 mL/L trace metals stock (1000
174 ×), and 1 mL/L vitamins stock (1000 ×). One liter of trace metals stock contained 1 mg
175 boric acid, 0.2 mg KI, 0.67 mg $FeCl_3 \bullet 6H_2O$, 0.125 mg $CuSO_4 \bullet 5H_2O$, 0.89 mg
176 $MnSO_4 \bullet H_2O$, 0.48 mg $Na_2MoO_4 \bullet 2H_2O$, 1.42 mg $ZnSO_4 \bullet 7H_2O$, 20 mg $FeSO_4 \bullet 7H_2O$,
177 and 0.73 mg $CoCl_2 \bullet 6H_2O$. One liter of vitamins stock contained 0.05 mg biotin, 0.8 mg
178 calcium pantothenate, 0.004 mg folic acid, 4 mg inositol, 0.8 mg niacin, 0.4 mg *p*-
179 aminobenzoic acid, 0.8 mg pyridoxine HCl, 0.4 mg riboflavin, 1.5 mg thiamine HCl.
180 The vitamins stock was sterilized by 0.22 μm hydrophilic filters. The strains used in
181 fermentation were pre-cultivated in the medium containing 20 g/L glucose, 1.7 g/L

182 yeast nitrogen (without amino acids and ammonium sulfate), 5 g/L ammonium sulfate,
183 and 0.69 g/L CSM-Ura. Next, 2.3 L of the precultures were inoculated into the
184 fermentative medium and cultivated at 28 ºC. Oxygen was supplied in the form of
185 filtered air via sparging rate of 20–80 L/min of air using agitation in 50–500 rpm range
186 to maintain a dissolved oxygen level at 40% in 24 h and below 5% after 24 h. The pH
187 of the cultures was constantly controlled at 5.5 using 12 M NaOH. During the process
188 of fermentation, glucose concentration was detected by an SBA-40D Biosensor
189 (Shandong Academy of Sciences, China) every 12 h.

190 **2.3. Extraction and quantification of lipid and protein**

191 Total lipids were extracted using a previous procedure with minor modification
192 (Xiao et al., 2015). Briefly, yeast cultures were collected by centrifugation and 0.3 g
193 wet cells of each sample was used for lipid extraction. The wet cells were resuspended
194 in 3 ml of 4 mol/L hydrochloric acid in glass tubes and shaken slightly for 1.5 h.
195 Subsequently, the cells were boiled for 8 min and then were frozen at –20 °C for half
196 an hour. Next, 6 ml of chloroform: methanol (1:1) was added to the tubes and
197 centrifuged at 2,000 g for 10 min. The supernatant was transferred into a clean tube and
198 3 ml of 0.15% NaCl was added and then centrifuged. Finally, the supernatant was
199 collected and dried using the blowing concentrator with nitrogen gas.

200 To detect the lipid content and composition by gas chromatography (GC), the
201 extracted lipid was esterified by methanol to fatty acid methyl esters (FAMEs). The
202 lipid was first incubated with 3.9 mL of methanol: sulfuric acid (98:2) at 85 °C for 3 h.
203 Next, 1.5 mL of saturated NaCl was added and FAMEs were extracted through the

204 addition of 1.5 mL hexane. The FAMEs were analyzed by an Agilent 7890B-GC
205 equipped with a flame ionization detector and a capillary column HP-5 (30 m, 0.32
206 mm). 1 μL of sample was injected at 250 °C using helium as the carrier gas at a flow
207 rate of 1 mL/min. The GC oven temperature was held at 140 °C for 1 min, and then
208 ramped to 180 °C for 10 min at a speed of 10 °C/min, 210 °C for 4 min at a speed of
209 5 °C/min, and 250 °C for 4 min at a speed of 5 °C/min. FAMEs were identified and
210 quantified using commercial FAME standards purchased from Sigma-Aldrich
211 (Shanghai, China). The content of proteins was determined by the Kjeldahl method
212 using the Kjeldahl instrument K9840 (Hanon Shandong Scientific Instruments Co., Ltd.,
213 Jinan, China).

214 **2.4. Structural analysis and point mutation of AtADS2**

215 To evaluate the substrate preference of AtADS2, the 3D structural model of
216 AtADS2 was constructed by SWISS-MODEL server using mouse SCD (PDB ID: 4ymk)
217 as the template (Schwede et al., 2003). Pictures of AtADS2 structures were generated
218 with the program UCSF Chimera (Pettersen et al., 2004). The docking of acyl-CoA into
219 the binding pocket of AtADS2 was performed with the AutoDock Vina software (Trott
220 et al., 2010). The 3D structures of acyl-CoA and the minimized energy structures were
221 constructed by BIOVIA Discovery Studio 4.5. Next, AutoDock Tools 1.5.4 was used to
222 assign hydrogens, Gasteiger charges and rotatable bonds to acyl-CoA. A docking grid
223 dimension was set to 64 Å × 70 Å × 64 Å and default values were used for other
224 parameters. Based on molecular docking and characteristics of the binding pocket of
225 AtADS2, some polar amino acids at the periphery of the binding pocket were replaced

with non-polar amino acids (Table S5). The *AtADS2* mutant genes were expressed in the *Y. lipolytica* strain YLVL6 to estimate the substrate preference according to the fatty acid profiles.

2.5. Determination of the distribution of nervonic acid in TAG

The distribution of nervonic acid in triacylglycerol (TAG) was detected according to a previous publication (Fan et al., 2018b). Briefly, 10 mg of TAG and 3 ml of methanol was mixed with 10 mg of 1,3-specific lipase from *Thermomyces lanuginose* (Lipozyme TL IM; Novozymes, Bagsvaerd, Denmark) and shaken at 30 °C for 8 h. The hydrolytic products were separated by thin layer chromatography (TLC) on Partisil K6 Silica gel 60 plates (250 μm thickness, 20 × 20 cm; Merck) with developing solvent, hexane: diethyl ether: acetic acid (70:30:1, v/v). Lipids were visualized under UV light by brief exposure to iodine vapor. Then, the free fatty acids and 2-monoacylglycerols (2-MAGs) was converted to the corresponding methyl ester by incubating with 2% H_2SO_4-methanol solution at 85°C for 3 h. The compositions of the fatty acid methyl esters were analyzed by GC and GC-MS.

2.6. Illumina genome sequencing and bioinformatic analyses

The genome sequences of representative edited clones were sequenced by Illumina technique. Standard genome sequencing and standard bioinformatic analyses were provided by Oebiotech (Shanghai, China). The filtered reads were mapped to the reference genome *Y. lipolytica* CLIB89 (W29) (GenBank: GCA_001761485.1) using the BWA 0.7.16a software. The aligned sequence reads were visualized by Integrative Genomics Viewer (IGV).

248 **2.7. Identification of DGAT2 and GPAT genes from *M. oleifera***

249 Putative acyltransferase DGAT and GPAT have been annotated in the genome of
250 *M. oleifera* and transcriptomic analysis showed that two *DGAT2* and two *GPAT* genes
251 highly expressed in seeds at fast oil accumulation stage (Yang et al., 2018). However,
252 it is not clear whether these acyltransferases prefer esterifying C24:1-acyl-CoAs to
253 generate nervonic acid lipids. In this study, the *DGAT2* and *GPAT* genes in *M. oleifera*
254 were identified by comparative sequence analysis. The amino acid sequence of DGAT2
255 (AEE78802.1) and GPAT (AAG23437.1) was used as a BLAST query against the *M.*
256 *oleifera* genome database to retrieve gene homologs, which were used as candidates for
257 identifying C24:1-acyl-CoAs preferred acyltransferases.

258 **2.8. Identification of regulators of lipid synthesis in *Y. lipolytica***

259 To identify regulators for lipid synthesis in *Y. lipolytica*, the regulators INO2
260 (NP_010408.1) and INO4 (NP_014533.1) in the phospholipid biosynthesis of
261 *Saccharomyces cerevisiae* were used as BLAST queries against the *Y. lipolytica* W29
262 genome (GCA_009372015.1) to retrieve the gene homologues. Homology modeling of
263 the retrieved YlINO2 and YlINO4 from *Y. lipolytica* was performed using SWISS-
264 MODEL server (Schwede et al., 2003). The bHLH ScINO2-ScINO4 transcription
265 activation complex bound the promoter region (PDB ID: 7XQ5) and the structures of
266 TFIIIB-TBP/Brf2/DNA and SANT domain of Bdp1 (PDB ID: 5N9G) in Protein
267 Databank (PDB) were used as templates for homology modeling of YlINO4 and
268 YlINO2, respectively. The predicted protein structures were visualized using PyMOL
269 Molecular Graphics System (Version 1.7.0.0). Figure 2F, 6A, 7A, 7B and 8A were

270 created by Figdraw (www.figdraw.com). Structure-based multiple sequence alignment
271 of the bHLH transcription factors was carried out on the T-COOFFEE online service
272 (Notredame et al., 2000).

273 **2.9. Separation and purification of nervonic acid**

274 Crude lipids extracted from cell sediment of the YLNA10 culture from 50 L-
275 reactor were used to separate nervonic acid. The extracts were hydrolyzed in the
276 solution of sodium hydroxide (pH 9.0) and ethanol at 80-85ºC with stirring for 3 h.
277 When no stratification was observed with the naked eye, the saponification solution
278 was acidified to pH 2.0 by sulfuric acid and heated to volatilize ethanol at 75ºC. Next,
279 free fatty acids in the saponification solution were extracted by using *n*-hexane for three
280 times, washed and dehydrated, and then evaporated. The extracts of free fatty acids
281 were dissolved in *n*-hexane for purification analysis.

282 Silica gel column chromatography was used in preliminary purification of
283 nervonic acid. Mobile phases consisting of *n*-hexane and ammonia (0.1%, 0.5%, or
284 10%) or acetic acid (0.5%, 1%, or 2%) were first evaluated by TLC on silica gel plates.
285 Next, the extracts of free fatty acids were isolated by silica gel column (26×3.0 cm i.d.)
286 using hexane containing 1% acetic acid as the elution buffer. The initial 75 ml eluent
287 was discarded, and following 230 ml eluent was collected and concentrated by rotary
288 evaporator. The residues were dissolved in 1 ml of methanol for further purification by
289 preparative chromatography. The effusion mainly containing nervonic acid was
290 determined and collected according to the HPLC spectrogram of nervonic acid standard.
291 Next, the collection was dried under a gentle nitrogen flow, dissolved in methanol, and

filtered through 0.22-μm nylon syringe filters for further separation by UHPLC (Ultra high performance liquid chromatography) (Thermo U3000) equipped with a UV detector through a C_{18} column Poroshell 120 EC-C18 (4.6 × 150 mm, 4 μm) (Agilent Technologies, Inc.). Mobile phase consisted of acetonitrile/methanol/0.4% acetic acid/ tetrahydrofuran (80:12:5:3, v/v/v/v). Twenty microliters of samples were automatically injected. A flow rate of 0.6 mL/min of the mobile phase was used.

The reagents methanol, acetonitrile, *n*-hexane, and tetrahydrofuran used in this study are all HPLC grade and purchased from Merck Co. Acetic acid, NaOH, ammonia, sulfuric acid, and ethanol are purchased from Sinopharm Chemical Ceagent Co., Ltd. (Shanghai, China). Nervonic acid, methyl cis-15-tetracosenoate, and ethyl 15-tetracosenoate (>99%) were purchased from Shanghai McLean Biochemical Technology Co., Ltd., China. Silica gel was provided by Qingdao Ocean Chemical Plant, China.

2.10. Transmission electron microscopy and fluorescence microscope

Yeast stock cells were inoculated into 5 ml YNB medium and cultivated at 28 ºC with shaking at 250 rpm for 24 h. Afterwards, 300 μL of the cultures were grown for 24 h in 30 mL of YNB medium at 28 °C and 250 rpm. Collect the cells by centrifugation and fix them by 2.5% glutaraldehyde at room temperature for 2 h. Then, gently wash the fixed cells by 0.1 M phosphate buffer solution (PBS) at pH 7.4 for three times. Further fix the cells by 1% osmic acid solution for 1 h and then wash them with PBS for three times. Dehydration of the fixed cells by acetone at the concentrations of 30%, 50%, 70%, 80%, 90%, 95% and 100%. Next, the cells were infiltrated by acetone and

314 resin at a ratio of 7:3 (v/v) for 5 h, 3:7 (v/v) for 5 h, and then embedded by pure resin
315 for 12 h. The specimen was stained by 2% uranyl acetate for 15 min and lead citrate for
316 5 min, and observed using a transmission electron microscope (FEI Tecnai G2 Spirit,
317 OR, USA). To observe the intracellular lipid droplets, the yeast cells were stained with
318 Nile red and observed by a fluorescence microscope (OLYMPUS BX35, Tokyo).

319 **2.11 Prediction and visualization of transmembrane segments in CgKCS and**
320 **MaOLE2**

321 The transmembrane segments (TMS) of CgKCS and MaOLE2 embedded in ER
322 were predicted using matching apparent free energy differences (ΔG prediction server
323 v1.0) via the Sec61 translocon (Hessa et al., 2007). The anticipated TMS with ΔG≤ 1.5
324 kcal/mol were considered to TMS (Ojemalm et al., 2013). For membrane protein
325 topology visualization, the PROTTER online servers was used (Omasits et al., 2014).
326 Structure homology modeling of MaOLE2 was carried out using the SWISS-MODEL
327 online servers (Schwede et al., 2003). The three-dimensional structure of a mouse
328 stearoyl-CoA desaturase (PBD ID: 4YMK) in PDB was used as the template for
329 homology modeling of MaoLE2. Three-dimensional structures modeling of CgKCS
330 was carried out using I-TASSER unified platform (Yang et al., 2015). The predicted
331 protein structures were visualized using the program UCSF Chimera (Pettersen et al.,
332 2004) and PyMOL Molecular Graphics System (Version 1.7.0.0).

333 **3. Results**

334 **3.1. Nervonic acid synthesis by fatty acid elongation**

335 The β-ketoacyl-CoA synthase KCS, one of the fatty acid elongase (FAE)
336 components, is considered the rate-limiting enzyme for VLCFAs synthesis (Huai et al.,
337 2015). We expressed the codon-optimized *CgKCS* gene from *C. graeca* in the *Y.*
338 *lipolytica* strain po1g and achieved production of nervonic acid accounting for 1.4% of
339 total fatty acids (TFA) in the strain YL-CgKCS (Fig. 2A). The strain po1g cannot
340 synthesize eicosenoic acid (C20:1) and erucic acid (C22:1), so it is considered that
341 C18:1-acyl-CoA was used as the substrate of CgKCS in YL-CgKCS (Fig. 2B). The
342 fatty acid elongases AtFAE1 in *Arabidopsis thaliana* and BtFAE1 in *Brassica*
343 *tournefortii* prefer extending oleic acid to eicosenoic acid and erucic acid, respectively
344 (Singh et al., 2014; Sun et al., 2013; Trenkamp et al., 2004). Protein topology prediction
345 showed that AtFAE1, BtFAE1 and CgKCS have extremely similar transmembrane
346 structures (Fig. S1A-S1E). To test whether C20:1-acyl-CoA and C22:1-acyl-CoA can
347 be used as substrates for KCS to produce nervonic acid, the genes *AtFAE1* and *BtFAE1*
348 were expressed in *Y. lipolytica*. As expected, expression of *AtFAE1* and *BtFAE1* resulted
349 in the production of eicosenoic acid (4.4% of the TFA) in YL-AtFAE1 and erucic acid
350 (5.4% of the TFA) in YL-BtFAE1 (Fig. 2C and 2D). The expression of *BtFAE1* also led
351 to the production of a small amount of nervonic acid (0.47% of TFA; Fig. 2D).
352 Multisequence alignment showed that the catalytic residues in CgKCS, AtFAE1 and
353 BtFAE1 are extremely conservative (Fig. S1F), while structural modeling indicated that
354 the size of substrate-binding regions were consistent with the preference of them to
355 different acyl-CoAs (Fig. S2). Nevertheless, the strain YL-3FAEs co-expressing of
356 *AtFAE1*, *BtFAE1* and *CgKCS* did not improve the production of nervonic acid (Fig. 2E),

357 indicating that CgKCS may not be able to effectively use C20:1-acyl-CoA and C22:1-
358 acyl-CoA as substrates in *Y. lipolytica*.

359 The biosynthesis of fatty acids is functionally partitioned in multiple cellular
360 compartments, including the cytosol, mitochondria, ER, peroxisome, and lipid droplets
361 (Kerner and Hoppel, 2000; Xu et al., 2016; Yang et al., 2019). We next investigated
362 whether CgKCS expression in specific subcellular organelles could help direct acyl-
363 CoA toward the formation of nervonic acid (Fig. 2F). Mitochondria was included in the
364 subcellular localization tests because of mitochondrial FAS found in yeast and the
365 interaction between mitochondria associated member and ER (Hiltunen et al., 2009;
366 Malina et al., 2018). Expression of *CgKCS* or *CgKCS* fused with ER, peroxisomal or
367 mitochondria signals led to similar production of nervonic acid (1.1%-1.3% of the TFA).
368 Co-expression of *CgKCS* fused with ER and peroxisome signals resulted in an increase
369 of nervonic acid to 2.9% of the TFA (Fig. 2G). Because expression of CgKCS without
370 additional target signals allowed production of nervonic acid and CgKCS is a
371 transmembrane protein (Fig. S1A), we speculated that CgKCS may contain inherent
372 signals for organelles localization in *Y. lipolytica*. To assess this speculation, we fused
373 *CgKCS* to a superfolder GFP gene (*sfGFP*) (Pedelacq et al., 2006). CgKCS-sfGFP
374 exhibited both punctate and accumulated localization in the early stage of cultivation
375 and targeted in lipid droplets in the lipid accumulation stage (Fig. 2H). This observation
376 indicates that CgKCS can target multiple cellular organelles in *Y. lipolytica*, including
377 ER and lipid droplets. Given the effectiveness of both original CgKCS and CgKCS
378 fusing with additional target signals, we employed combinatorial patterns to express

these genes to further improve nervonic acid production.

3.2. Evaluation of nervonic acid synthesis by desaturation

Theoretically, nervonic acid could be synthesized from both C18:1-acyl-CoA and C18:0-acyl-CoA by expression of KCS and desaturases (Fig. 1). However, few fatty acid desaturases with VLCFA activities have been identified. To the best of our knowledge, the acyl-coenzyme A desaturase-like protein AtADS2 in *A. thaliana* is the only reported enzyme able to catalyze C24:0-acyl-CoA to synthesize nervonic acid (Smith et al., 2013). We expressed the *AtADS2* gene in the strain YL-3FAEs due to the TFA of this strain containing 7% lignoceric acid (C24:0) (Fig. 2E). Determination of the fatty acids in the *AtADS2*-overexpresed strains showed that nervonic acid was nearly undetected while oleic acid improved from 26% to 54% in TFA (Fig. S3A and S3B), suggesting that AtADS2 desaturated C18:0-acyl-CoA by Δ9-desaturase activity. This finding is in contrast to expressing *AtADS2* in *S. cerevisiae*, in which AtADS2 presented Δ15-desaturase activity leading to nervonic acid production from C24:0-acyl-CoA (Smith et al., 2013).

Structural analysis of AtADS2 by SWISS-MODEL server showed that the three-dimensional structure of AtADS2 is highly similar to that of mouse stearoyl-CoA desaturase (mSCD, PDB ID:4ymk) (Bai et al., 2005). Like typical SCDs, cross-sections of the AtADS2 surface have a tunnel-like substrate binding pocket, which determines the chain length of acyl-CoA substrates (Fig. S3C). Compared with C18:0-acyl-CoA, C24:0-acyl-CoA has a longer hydrophobic hydrocarbon chain that cannot effectively combine with the binding pocket. We explored improving the substrate preference of

401 AtADS2 toward C24:0-acyl-CoA by modifying amino acids around the pocket to
402 enhance its hydrophobicity (Table S4). Unexpectedly, the content of nervonic acid
403 decreased in all mutants, while the content of oleic acid increased from 36% to above
404 47% in TFA along with a slight increase of eicosenoic acid (Fig. S3D). Since AtADS2
405 and the mutants tested here prefer desaturating C18:0-acyl-CoA to produce oleic acid
406 in *Y. lipolytica*, the strategy of desaturation of C24:0-acyl-CoA for production of
407 nervonic acid was given up in this study.

408 **3.3. Modulation of endogenous pathways improved nervonic acid production**

409 Palmitoleic acid is a ω-7 fatty acid that cannot be used as the substrate for
410 producing the ω-9 nervonic acid (Fig. 1). With the overexpression of *CgKCS*,
411 palmitoleic acid increased significantly, from 13.5% to 27.3% of TFA (Fig. 2A and 2B).
412 We assumed that expression of C16/C18 FAEs that prefer to convert C16:0-acyl-CoA
413 to C18:0-acyl-CoA together with Δ9 desaturases with high substrate preference for
414 C18:0-acyl-CoA should decrease the synthesis of palmitoleic acid and increase the
415 C18:1-acyl-CoA precursor for nervonic acid production (Fig. 1).

416 We next evaluated the substrate preference of heterologous FAEs for C16:0-acyl-
417 CoA and fatty acid desaturases for C18:0-acyl-CoA *in vivo* (Table S5). The goat
418 ELOVL6 (gELOVL6) (Shi et al., 2017), rat ELO2 (rELO2) (Yazawa et al., 2011), and
419 CpLCE1 from *Cryptosporidium parvum* (Frltzler et al., 2007), were chosen according
420 to previous reports about the high elongation preference of them toward C16:0-acyl-
421 CoA. The tested fatty acid desaturases included OLE2 from *Mortierella alpine*
422 (MaOLE2) (Wan et al., 2013), FAT6 from *Caenorhabditis elegans* (CeFAT6) (Watts

423 and Browse, 2000), and D9DMB from *Cunninghamella echinulata* (Wongwathanarat
424 et al., 1999). All the genes were optimized according to the codon usage bias of *Y.*
425 *lipolytica*.

426 Expression of *CeFAT6* and *rELO2* in *Y. lipolytica* po1g increased the content of
427 linoleic acid from 21.6% in TFA to 36.4% and 41.7%, respectively. In contrast,
428 expression of *MaOLE2* or *gELOVL6* in po1g led to an effective increase of oleic acid
429 from 36.3% in TFA to more than 50% along with a decrease of palmitoleic acid from
430 12.1% in TFA to below 8.4% (Fig. 3A and 3B). Because CgKCS prefers using C18:0-
431 acyl-CoA as the substrate for nervonic acid synthesis, *MaOLE2* and *gELOVL6* were
432 employed in constructing nervonic acid overproducing strains. Expression of *gELOVL6*
433 in the *CgKCS*-overexpressing strain YL-3CgK decreased the contents of palmitoleic
434 acid and palmitic acid in TFA by 7.5% and 57.2% in the strain YL-3CgKE, respectively.
435 Meanwhile, nervonic acid increased from 3.3% to 13.2% of the TFA in YL-3CgKE (Fig.
436 3C). Furthermore, expressing an additional copy of each of *gELOVL6* and *MaOLE2* in
437 YL-3CgKE reduced palmitoleic acid by 40.5% and increased nervonic acid to 15.9%
438 of TFA in YL-3CgKEM, (Fig. 3C). A maximum nervonic acid titer of 0.34 g/L was
439 achieved in flask cultures. These results confirmed the effectiveness of expression of
440 *gELOVL6* and *MaOLE2* to decrease palmitoleic acid and improve the precursor for
441 nervonic acid production in *Y. lipolytica*. The distribution of nervonic acid in TAG was
442 then detected by hydrolysis using a 1,3-specific lipase as well as TLC and GC-MS
443 analysis. The results indicated that nervonic acid produced by the engineered *Y.*
444 *lipolytica* localized at *sn*-1and *sn*-3 positions in TAG (Fig. 3D).

3.4. Construction of nervonic acid overproduction strains by homologous recombination in lipid-overproduction strains

Although the nervonic acid content in TFA reached 15.9% in the strain YL-3CgKEM, the lipid titer was only 2.1 g/L in flask cultures. To improve the titer of both nervonic acid and lipids, we iteratively overexpressed *CgKCS*, *gELOVL6* and *MaOLE2* in the lipid-overproduction strain po1g-G3 (Li et al., 2020) by both random integration and homologous recombination (Fig. 4A). Six copies of *CgKCS* were expressed by random genome insertion to generate the strain YLVL6, which produced 0.69 g/L of nervonic acid (6.5% of TFA) in flask (Fig. 4B and 4C). However, further expression of additional copies of *CgKCS* by random genetic integration without clearly improving nervonic acid production in YLVL9 (Fig. S4A and S4B). Genome resequencing revealed that the additional *CgKCS* expression cassettes eliminated the expression of previous *CgKCS* by recombination with their TEF1 promoters (Fig. S4C). Thus, we next adopted homologous recombination to overexpress heterologous genes at given loci. Simultaneous expression of *CgKCS*, *gELOVL6* and *MaOLE2* at the ribosomal DNA loci of YLVL6 generated the strain YLVL7, in which the content and titer of nervonic acid reached 10.0% of the TFA and 1.46 g/L in flask cultures, respectively (Fig. 4B and 4C). Additional expression of *CgKCS*, *gELOVL6* and *MaOLE2* at the *D17* locus (Schwartz et al., 2017) increased nervonic acid to 11.8% of the TFA and 1.71 g/L in the strain YLVL8 (Fig. 4B and 4C). Next, double *CgKCS* expression cassettes were co-expressed at the fatty acid desaturase-2 (FAD2) gene locus in YLVL8, producing the strain YLVL10. Disruption of *FAD2* led to a decrease of lipid production, from 14.5 g/L

467 in YLNA8 to 11.6 g/L in YLVL10, with a slight increase of the nervonic acid titer, from
468 1.71 g/L to 1.76 g/L as nervonic acid content increased to 15.2% from 11.8% of the
469 TFA (Fig. 4B to 4D). The strain YLVL10 produced VLCFAs at 3.74 g/L as high as 25-
470 folds of that in the initial strain po1g-G3 in flasks. The content of VLCFAs in YLVL10
471 was 32.3% of the TFA. The ratio of nervonic acid and lignoceric acid to the total
472 VLCFAs reached 47.1% and 38.2%, respectively.

473 Meanwhile, we constructed a series of YLNA strains (YLNA1 to 7), in which all
474 overexpressed genes were integrated into the genome of po1g-G3 by homologous
475 recombination (Fig. 5A). In the strain YLNA7, seven copies of *CgKCS* were expressed
476 at the genomic loci of *rDNA*, *FAD2*, *TGL4*, *GSY1* and *SNF1*. A nervonic acid titer of
477 2.60 g/L were achieved in YLNA7 (Fig. 5A to C), which was apparently higher than
478 the 1.76 g/L achieved in YLVL10. Interestingly, the ratio of nervonic acid to lignoceric
479 acid (C24:0) increased from 1.38 in YLNA6 to 2.23 in YLNA7 by disruption of the
480 AMP-activated S/T protein kinase *SNF1* (Fig. 5C and 5D), which facilitates the
481 purification of nervonic acid. The ratio of nervonic acid to lignoceric acid in YLNA7
482 was also significantly higher than that in YLVL10 (2.23 vs 1.23).

483 Since the fatty acid desaturase MaOLE2 prefers to desaturate C18:0-acyl-CoA to
484 generate C18:1-acyl-CoA as the substrate of CgKCS for nervonic acid biosynthesis, it
485 is hypothesized that the proximity of MaOLE2 and CgKCS in ER benefits the
486 production of nervonic acid (Fig. 6A). Protein transmembrane prediction showed that
487 MaOLE2 has five transmembrane helices (TM1–TM5) arranged in a trumpet-like shape
488 (Fig. 6B and 6D). The dimetal (Zn^{2+}) catalytic activity sites locate in two helical turns

489 toward the cytoplasm (Fig. 6B). CgKCS has two transmembrane segments (TM1 and
490 TM2) at the N-terminal (Fig. 6C and 6E). When the transmembrane regions were
491 removed, the truncated CgKCS maintained only a weak longation ability (Fig. S5).
492 Based on the structure information, a flexibility glycine linker (GGGGGGGGGG) was
493 used to fuse the C-terminal of MaOLE2 with the N-terminal of CgKCS. The fusion
494 protein was expressed at the *D17* locus in YLNA7, producing the strain YLNA8. The
495 nervonic acid content in YLNA8 (17.3%) was slightly higher than YLNA7 (16.5%)
496 (Fig. 5C and 5D). The ratio of nervonic acid to lignoceric acid in YLNA8 also increased
497 from 2.23 to 2.27.

498 **3.5. Expansion of ER enhancing lipid and nervonic acid production**

499 The INO2/INO4 transcription factor complex has been demonstrated to activate
500 phospholipid biosynthesis and ER structural alteration in *S. cerevisiae* (Fig. 7A and 7B)
501 (Schwank et al., 1997). Over-expression of *INO2* and *INO4* improved the production
502 of terpenoids, 3-hydroxypropionic acid, ethanol and lycopene by expanding the ER size
503 or enhancing cellular stress response in *S. cerevisiae* (Chen et al., 2016, 2017; Kim et
504 al., 2019; Fang et al., 2020). However, the homologues of INO2/INO4 have not been
505 identified in *Y. lipolytica*.

506 To recognize the ER regulatory factors in *Y. lipolytica*, a putative inositol-3-
507 phosphate synthase YlINO1 (NC 006068.1) was first identified by two-sequence
508 alignment. The YlINO1 was confirmed by prediction of the interaction between
509 ScINO2/INO4 and the YlINO1 promoter by using the software JASPAR2020 (Fig.
510 S6A), indicating *Y. lipolytica* having INO2/INO4 complex. YlINO1 and ScINO1 (NC

511 001142.9) showed a sequence identity of 67.2% (Fig. S6B). Subsequently, putative ER
512 regulators YlINO2 (GenBank No. AOW01275.1) and YlINO4 (GenBank No.
513 AOW06291.1) were identified by homology analysis in the genome of *Y. lipolytica*. In
514 the GenBank database, YlINO2 was annotated as a hypothetical protein with no specific
515 function, whereas YlINO4 was annotated as a ScINO4-like protein. YlINO4 shared a
516 sequence identity of 37.3% with ScINO4 and they had similar three-dimensional
517 structure (Fig. S7A-S7C). The sequence similarity between templet (Bdp1) and YlINO2
518 is only 15.2% (Fig. 7C). Structure-based multi-sequence alignment showed that the
519 putative YlINO2/INO4 and the Sc INO2/INO4 shared the same essential amino acids
520 for DNA recognition (Fig. 7D, Fig. S7D and S7E), indicating that YlINO2/YlINO4 is
521 ER transcription factors.

522 Next, the putative YlINO2 and YlINO4 were separately expressed at the *PEX10*
523 locus in YLNA8. Overexpression of YlINO4 did not result in clear changes in the
524 production of lipid and nervonic acid (Fig. 7E). Overexpression of YlINO2 resulted in
525 a 39.3% increase in lipid production from 15.2 g/L in YLNA8 to 21.1 g/L in YLNA9,
526 as well as a significant increase in biomass from 28.0 g/L to 36.6 g/L (Fig. 7E). The rate
527 of sugar consumption speeded up quickly from 48 h that resulted in rapid increases in
528 the biomass and lipid titer in YLNA9 expressing YlINO2 (Fig. 7F). A lipid content of
529 59.5% in the DCW and a nervonic acid content of 16.5% in TFA were achieved in
530 YLNA9 (Fig. 7G and 7H). The titer of nervonic acid increased 18.2% from 2.96 g/L to
531 3.5 g/L in YLNA9 (Fig. 7E). The nervonic acid was confirmed by GC-MS (Fig. S8).
532 The ER membrane expansion was observed in the YlINO2 overexpressed strains

YLNA9 by transmission electron microscopy (Fig. S9), further indicated that YlINO2 is an ER structure regulator.

3.6. Engineering the esterification of C24:1-acyl-CoAs

The biosynthesis of TAG from fatty acyl-CoA and glycerol-3-phosphate requires glycerol-3-phosphate acyltransferases (GPAT), lysophospholipid acyltransferases (LPAT), phosphatidic acid phosphatase (PAP), and diacylglycerol acyltransferases (DGAT) (Fig. 8A). Most fatty acids exist in the lipid forms of TAG. Once C16-acyl-CoAs and C18-acyl-CoAs are esterified to form lipids, they cannot be elongated to C24:1-acyl-CoAs for the synthesis of nervonic acid. So, esterified enzymes with substrate preference for C24:1-acyl-CoAs rather than C16-acyl-CoAs or C18-acyl-CoAs are expected to improve the production of nervonic acid. In view of the seeds of *M. oleifera* rich in nervonic acid (Chivandi et al., 2008), *DGAT* and *GPAT* genes in this plant were systematically evaluated to identify C24:1-acyl-CoAs preferred ones.

Eight putative GPATs and ten putative DGAT2s sequences were extracted from the *M. oleifera* genome by homologous sequence analysis (Fig. 8B and 8C). Further functional prediction using UniPort in NCBI indicated that two GPATs (GenBank No. oleifera 006088 and 006090) and two DGAT2s (GenBank No. oleifera 010035 and 015949) possibly prefer to esterify C24:1-acyl-CoAs. Both GPAT sequences 006088 and 006090 showed a sequence identity of above 55% compared with the GPAT AAG23437.1 in *A. thaliana*. The sequence identity of DGAT2 010035 and 015949 to the DGAT2 AEE78802.1 in *A. thaliana* was 25.4% and 37.8%, respectively. Separated expression of the four putative acyltransferases at the *PEX10* locus in YLNA8 increased

555 the lipid titer by 22.1% (strain MoGPAT-88), 18.8% (strain MoGPAT-90), 15.6% (strain
556 MoDGAT2-35) and 17.2% (strain MoDGAT2-49), respectively (Fig. 8D). A highest
557 nervonic acid titer of 3.36 g/L was achieved in MoGPAT-88 (entitled YLNA10), which
558 is 28.1% higher than that in the strain YLNA8 (Fig. 8D). The content of nervonic acid
559 in TFA improved from 17.3% to 18.0% in YLNA10 (Fig. 8D-8F). The ratio of nervonic
560 acid to lignoceric acid in YLNA10 increased to 3.5 from 2.27 in YLNA8.

561 **3.7. Production of nervonic acid by two-stage fed-batch bioreactor**

562 The fermentative medium and conditions were optimized in flasks and a 3-L reactor.
563 Central composite design (CCD) and response surface methodology were used to
564 optimize the concentrations of glucose and nitrogen sources (yeast extract and
565 ammonium sulfate) in flasks using strain YLVL3 (Table S6 and S7). In two rounds of
566 optimization experiments, the medium containing 150 g/L glucose, 6 g/L yeast extract,
567 and 12 g/L ammonium sulfate exhibited optimal lipid production (Fig. S10), with lipid
568 and nervonic acid titer that were 2.9-fold and 6.7-fold the levels obtained in the control
569 media, respectively (Fig. S10). The strain YLVL6 was also cultivated by fed-batch
570 fermentation in the 3-L bioreactor and the feeding medium composition was evaluated
571 for maximal lipid production. The results showed that feeding glucose and ammonium
572 sulfate with a C/N molar ratio of 100:1 during fermentation resulted in an 84.6%
573 increase in lipid titer (41.4 g/L) compared to that achieved by feeding solely glucose
574 (Fig. S11A). The lipid titer was further improved to 43.7 g/L by controlling the
575 dissolved oxygen level at 20% up to 24 h and below 5% after 24 h, and the biomass
576 reached 159.2 g/L at 120 h (Fig. S11B). Under the optimized conditions, a nervonic

577 acid titer of 4.2 g/L was achieved by 120 h for strain YLVL6 in the 3-L reactor.

578 On the basis of the optimized fermentation medium and conditions, pilot-scale
579 fermentation in 50-L reactor was separately performed using the strains YLVL10 and
580 YLNA9 (Fig. 9A-9C). Lipid and nervonic acid synthesized constantly by YLVL10 in
581 192 h during the fermentation process (Fig. 9C). A VLCFAs titer of 24.0 g/L was
582 obtained in YLVL10 at 192 h, which consisted of behenic acid (C22:0) 1.6 g/L, erucic
583 acid (C22:1) 2.2 g/L, lignoceric acid (C24:0) 10.0 g/L and nervonic acid (C24:1) 10.2
584 g/L (Fig. 9B). Lipid and nervonic acid synthesized constantly by YLNA9 in 216 h (Fig.
585 9C). The fastest lipid production in YLNA9 happened between 48 h and 72 h which
586 was earlier than that of 72 h to 96 h in YLVL10, and the highest yield during this period
587 in both YLNA9 and YLVL10 reached 0.267 g FAME/g glucose (98% of the maximum
588 theoretical yield). A largest lipid titer of 96.7 g/L and a lipid content of 52.1% to DCW
589 were achieved in YLNA9 (Fig. 9C, Fig. S12). The titer of nervonic acid and VLCFAs
590 in YLNA9 reached 17.3 g/L and 28.2 g/L, respectively (Fig. 9C), as far as we known,
591 both of which were more than two-folds of the highest levels in previous reports (Fillet
592 et al., 2017). The productivity of nervonic acid is 0.135 $g \cdot L^{-1} \cdot h^{-1}$.

593 **3.8. Separation and purification of nervonic acid**

594 Both nervonic acid and lignoceric acid have 24 carbons and the positions of them
595 in GC spectrums are close (Fig. 5B). Separation of nervonic acid from lignoceric acid
596 and other fatty acid components was explored. The lipids produced by the strain YLNA-
597 10 in 50-L reactor were extracted and saponified to release free fatty acids. The
598 standards of methyl cis-15-tetracosenoate and ethyl 15-tetracosenoate were used to

estimate saponification conditions. The recoveries of methyl cis-15-tetracosenoate and ethyl 15-tetracosenoate reached 93.38 ± 6.20 % and 101.71 ± 10.53%, respectively, indicating the feasibility of the processes used in this study.

After saponification, silica gel column chromatography was used to separate nervonic acid in view of the weak polarity of it. Mobile phases consisting of *n*-hexane and ammonia or acetic acid were first evaluated on TLC. The results showed that using the solutions containing *n*-hexane and acetic acid or 0.5% ammonia successfully presented nervonic acid (Fig. S13). Particularly, the mobile phase consisting of *n*-hexane and 1% acetic acid exhibited best separation efficiency and product purity and was used in silica gel column chromatography. Furthermore, discarding 75 ml of initial eluent and collecting 230 ml of subsequent eluent resulted in maximum recovery of nervonic acid. Next, the products purified by silica gel column chromatography were further separated by a C_{18} column using UHPLC and the eluent containing nervonic acid was collected according to the elution time of nervonic acid standard. Nervonic acid was verified by HPLC-MS/MS and the purity of the nervonic acid solution reached 98.7% (m/m) determined by HPLC (Fig. 10). Thus, these processes successfully separated nervonic acid from lignoceric acid and other fatty acid components.

4. Discussion

The number of lipid structures collected in the LIPID MAPS® Structure Database (LMSD) exceeds 47, 000 so far. Many lipids have similar structures except the chain length of fatty acids because of the non-specificity of elongation and esterification

620 reactions toward fatty acid units. Therefore, evaluation of fatty acid elongases and
621 esterifying enzymes is crucial for producing tailored fatty acid species (Xin et al., 2019;
622 Wang et al., 2021). In this work, we identified C16:0-acyl-CoA preferred elongases and
623 C18:0-acyl-CoA preferred desaturases that combined to effectively promote the
624 formation of C18:1-acyl-CoA, a common precursor for production of long-chain
625 unsaturated fatty acids (LCUFAs) and very long-chain unsaturated fatty acids
626 (VLCUFAs). To synthesize VLCFAs, KCS is among the key enzymes used for
627 lengthening fatty acid chains (Huai et al., 2015). Here, we found that CgKCS prefers to
628 directly convert C18:1-acyl-CoA to C24:1-acyl-CoA by three rounds of two-carbon
629 addition in the yeast *Y. lipolytica*. Meanwhile, increase of C20:1-acyl-CoA and C22:1-
630 acyl-CoA by expression of *AtFAE1* and *BtFAE1* did not help CgKCS to produce more
631 C24:1-acyl-CoA and nervonic acid. In contrast, expression of CgKCS in *R. toruloides*
632 led to approximately equal amount of erucic acid (C22:1) and nervonic acid (Fillet et
633 al., 2017). In addition, some KCSs prefers using C22:1-acyl-CoA as substrate to
634 synthesize nervonic acid in plants (Guo et al., 2009). Accordingly, we concluded that
635 the products of KCSs depend on both enzyme specificity and the hosts, and therefore,
636 both aspects should be seriously considered when VLCFAs are produced by metabolic
637 engineering.

638 In oleaginous prokaryotes and eukaryotes, most acyl-CoAs are incorporated onto
639 glycerol-3-phosphate (G3P) to generate TAGs as a form of energy storage. TAG
640 biosynthesis is an evolutionarily conserved process, while acyltransferases GPAT and
641 DGAT involved in the synthesis of TAGs do not possess rigid substrate specificity

642 toward acyl-CoAs (Liu et al., 2012; Korbes et al., 2016). Acyltransferases with similar
643 amino acid sequences can exhibit completely different acyl-donor specificities (Jeppson
644 et al., 2020). In the nervonic acid producing plant *M. oleifera*, *GPAT* and *DGAT* genes
645 have been identified, which exhibited apparently differential transcript levels in seeds
646 (Yang et al., 2018). In this study, we evaluated the high expressed genes in *M. oleifera*
647 seeds and demonstrated that expression of the GPAT Maole_006088.T1 encoding gene
648 in *Y. lipolytica* effectively improved the titer and content of nervonic acid. This result
649 indicated that regulating esterification specificity is also a strategy to overproduce
650 tailored fatty acid species, while this strategy has not received enough attention.

651 Lipids play a vital function in the physiology of cells and lipid metabolism
652 correlates with various organelles in eukaryotes (Kerner and Hoppel, 2000; Xu et al.,
653 2016; Yang et al., 2019). The organelle ER is particularly critical in the elongation and
654 desaturation of fatty acids and lipid formation. Biosynthesis of polyunsaturated fatty
655 acids (PUFAs), VLCFAs, and total TAGs require enzymes being targeted on ER, while
656 expression of too many heterologous enzymes on ER may increase the load of it. Here,
657 we identified a hypothetical protein YlINO2 (AOW01275.1) in *Y. lipolytica*, which is
658 possibly a counterpart of *S. cerevisiae* INO2. Overexpression of *YlINO2* expanded the
659 ER area and apparently improved the titer of both TAG and nervonic acid in *Y. lipolytica*,
660 highlighting the significance of ER engineering for production of lipids and tailored
661 fatty acids.

662 Some VLCFAs and VLCFA derived products have been utilized as lubricants,
663 detergents, cosmetics and pharmaceuticals, such as erucic acid, docosanol, and wax

esters (Gao et al., 2020; Taylor et al., 2010; Yu et al., 2017). Microbial production of VLCFAs from renewable feedstocks presents a promising alternative route to sourcing these materials from plant oils or the petrochemical industry. In this study, we engineered the oleaginous yeast *Y. lipolytica* to produce the VLCUFA nervonic acid and achieved the highest reported titer to date, 17.3 g/L in a 50-L bioreactor, indicating a bright prospect for commercial use. In contrast to unsaturated fatty acids and long-chain saturated fatty acids, very long-chain saturated fatty acids (VLSFAs) have received limited attention. Interestingly, recent studies reported that arachidic acid (20: 0), behenic acid (22: 0), and lignoceric acid (24: 0) need to be distinguished from other saturated fatty acids for their potential health benefits in reducing risks of diabetes, heart disease mortality and aging (Fretts et al., 2020; Jensen et al., 2020; Lemaitre and King, 2022). Here, we engineered *Y. lipolytica* to produce the VLCFAs at 28.2 g/L in a 50-L bioreactor, indicating this oleaginous yeast is a superior host for production of valuable VLCFAs.

Data availability

The authors declare that the data supporting the findings of this study are available within the Article and its Supplementary Information files.

Author contributions

S.W., F.L. and W.F. conceived the project. S.W. and F.L. designed the experiments. H.M., H.F., H.S., P.S. and J.L. constructed the engineered strains. H.S. and J.X.

performed bioreactor fermentation. S.W., H.S. and H.M. analyzed the data. S.W. and H.S. wrote the manuscript. S.W. and F.L. reviewed and edited the manuscript. All authors contributed to the manuscript, and read and approved the final version.

Declaration of competing interest

The authors declare competing financial interests: S.W., F.L., and W.F. have filed patent applications on this work through Qingdao Institute of Bioenergy and Bioprocess Technology, CAS and Zhejiang Zhenyuan Pharmaceutical Co., LTD, China. S.W., F.L., and W.F. have commercial interests in Zhejiang Zhenyuan Pharmaceutical Co., LTD.

Acknowledgements

This work was funded by National Key R&D Program of China (2022YFC2106200 and 2022YFE0207100), Shandong Energy Institute (SEI I202136), and Zhejiang Zhenyuan Pharmaceutical Co., LTD, China (Y86101190B).

Reference

Abdel-Mawgoud, A.M., Markham, K.A., Palmer, C.M., Liu, N., Stephanopoulos, G., Alper, H.S., 2018. Metabolic engineering in the host *Yarrowia lipolytica*. Metab. Eng. 50, 192-208.

Astarita, G., Jung, K.M., Vasilevko, V., DiPatrizio, N.V., Martin, S.K., Cribbs, D.H., et al., 2011. Elevated stearoyl-CoA desaturase in brains of patients with Alzheimer's disease. PloS. One. 6(10), e24777.

Avalos, J.L., Fink, G.R., Stephanopoulos, G., 2013. Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. Nat. Biotechnol. 31, 335-341.

Bai, Y., McCoy, J. G., Levin, E. J., Sobrado, P., Rajashankar, K. R., Fox, B. G., et al., 2015. X-ray structure of a mammalian stearoyl-CoA desaturase. Nature. 524, 252-256.

Chen, Y., Xiao, W. H., Wang, Y., Liu, H., Li, X., Yuan, Y. J., 2016. Lycopene overproduction in *Saccharomyces cerevisiae* through combining pathway engineering with host engineering. Microb. Cell Fact. 15, 113.

Chivandi, E., Davidson, B.C., Erlwanger, K.H., 2008. A comparison of the lipid and fatty acid profiles from the kernels of the fruit (nuts) of *Ximenia caffra* and *Ricinodendron rautanenii* from Zimbabwe. Ind. Crop. Prod. 27, 29-32.

Delgado, G.E., Kramer, B.K., Lorkowski, S., Marz, W., von Schacky, C., Kleber, M.E., 2017. Individual omega-9 monounsaturated fatty acids and mortality-the Ludwigshafen risk and cardiovascular health study. J. Clin. Lipidol. 11, 126-135.

Fan, Y., Meng, H.M., Hu, G.R., Li, F.L., 2018a. Biosynthesis of nervonic acid and perspectives for its production by microalgae and other microorganisms. Appl. Microbiol. Biot. 102, 3027-3035.

Fan, Y., Yuan, C., Jin, Y., Hu, G.R., Li, F.L., 2018b. Characterization of 3-ketoacyl-coA synthase in a nervonic acid producing oleaginous microalgae *Mychonastes afer*. Algal Res. 31, 225-231.

Fang, T.S., Yan, H.B., Li, G.Z., Chen, W.P., Liu, J., Jiang, L.H., 2020. Chromatin

remodeling complexes are involvesd in the regulation of ethanol production during static fermentation in budding yeast. Genomics. 112, 1674-1679.

Fillet, S., Ronchel, C., Callejo, C., Fajardo, M.J., Moralejo, H., Adrio, J.L., 2017. Engineering *Rhodosporidium toruloides* for the production of very long-chain monounsaturated fatty acid-rich oils. Appl. Microbiol. Biot. 101, 7271-7280.

Fretts, A.M., Jensen, P.N., Hoofnagle, A., McKnight, B., Howard, B.V., Umans, J., et al., 2020. Plasma ceramide species are associated with diabetes risk in participants of the strong heart study. J. Nutr. 150, 1214-1222.

Gao, Q., Yang, J.L., Zhao, X.R., Liu, S.C., Liu, Z.J., Wei, L.J., et al., 2020. *Yarrowia lipolytica* as a metabolic engineering platform for the production of very-long-chain wax esters. J. Agr. Food. Chem. 68, 10730-10740.

Gemperlein, K., Dietrich, D., Kohlstedt, M., Zipf, G., Bernauer, H.S., Wittmann, C., et al., 2019. Polyunsaturated fatty acid production by *Yarrowia lipolytica* employing designed myxobacterial PUFA synthases. Nat. Commun. 10, 1-12.

Guo, Y. M., Mietkiewska, E., Francis, T., Katavic, V., Brost, J. M., Giblin, M., et al., 2009. Increase in nervonic acid content in transformed yeast and transgenic plants by introduction of a *Lunaria annua* L. 3-ketoacyl-CoA synthase (KCS) gene. Plant Mol. Biol. 69, 565-575.

Hessa, T., Meindl-Beinker, N. M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., et al., 2007. Molecular code for transmembrane-helix recognition by the sec61 translocon. Nature, 450(7172), 1026-1030.

Hiltunen, J. K., Schonauer, M. S., Autio, K. J., Mittelmeier, T. M., Kastaniotis, A. J.,

Dieckmann, C. L., 2009. Mitochondrial fatty acid synthesis type II: more than just fatty acids. J. Biol. Chem. 284, 9011-9015.

Holkenbrink, C., Ding, B. J., Wang, H. L., Dam, M. I., Petkevicius, K., Kildegaard, K., et al., 2020. Production of moth sex pheromones for pest control by yeast fermentation. Metab. Eng. 62, 312-321.

Hu, D. D., Cui, Y. J., Zhang, J., 2021. Nervonic acid ameliorates motor disorder in mice with Parkinson's disease. Neurochem. J. 15, 317-324.

Huai, D.X., Zhang, Y.Y., Zhang, C.Y., Cahoon, E.B., Zhou, Y.M., 2015. Combinatorial effects of fatty acid elongase enzymes on nervonic acid production in *Camelina sativa*. PloS. One. 10(6), e0131755.

Jeppson, S., Mattisson, H., Demski, K., Lager, I., 2020. A predicted transmembrane region in plant diacylglycerol acyltransferase 2 regulates specificity toward very-long-chain acyl-CoAs. J. Biol. Chem. 295, 15398-15406.

Jensen, P.N., Fretts, A.M., Hoofnagle, A.N., Sitlani, C.M., McKnight, B., King, I.B., et al., 2020. Plasma ceramides and sphingomyelins in relation to atrial fibrillation risk: the cardiovascular health study. J. Am. Heart. Assoc. 9, e012853.

Kageyama, Y., Deguchi, Y., Hattori, K., Yoshida, S., Goto, Y., Inoue, K., et al., 2021. Nervonic acid level in cerebrospinal fluid is a candidate biomarker for depressive and manic symptoms: A pilot study. Brain Behav. 11(4): e02075.

Kageyama, Y., Kasahara, T., Nakamura, T., Hattori, K., Deguchi, Y., Tani, M., et al., 2018. Plasma nervonic acid is a potential biomarker for major depressive disorder: a pilot study. Int. J. Neuropsychoph. 21, 207-215.

769 Keppley, L.J.W., Walker, S.J., Gademsey, A.N., Smith, J.P., Keller, S.R., Kester, M., et
770 al., 2020. Nervonic acid limits weight gain in a mouse model of diet-induced
771 obesity. Faseb. J. 34, 15314-15326.
772 Kerner, J., Hoppel, C., 2000. Fatty acid import into mitochondria. BBA-Mol. Cell. Biol.
773 L. 1486, 1-17.
774 Kenar, J.A., Moser, B.R., List, G.R., 2017. Naturally occurring fatty acids: source,
775 chemistry, and uses. Fatty Acids. 23-82.
776 Kim, J.E., Jang, I.S., Son, S.H., Ko, Y.J., Cho, B.K., Kim, S.C., et al., 2019. Tailoring
777 the *Saccharomyces cerevisiae* endoplasmic reticulum for functional assembly of
778 terpene synthesis pathway. Metab. Eng. 56, 50-59.
779 Korbes, A.P., Kulcheski, F.R., Margis, R., Margis-Pinheiro, M., Turchetto-Zolet, A.C.,
780 2016. Molecular evolution of the lysophosphatidic acid acyltransferase (LPAAT)
781 gene family. Mol. Phylogen. Evol. 96, 55-69.
782 Lazar Z., Liu N., Stephanopoulos G., 2018. Holistic Approaches in Lipid Production
783 by *Yarrowia lipolytica*. Trends Biotechnol. 36(11):1157-1170.
784 Lemaitre, R.N., King, I.B., 2022. Very long-chain saturated fatty acids and diabetes and
785 cardiovascular disease. Curr. Opin. Lipidol. 33, 76-82.
786 Lewkowicz, N., Piatek, P., Namiecinska, M., Domowicz, M., Bonikowski, R., Szemraj,
787 J., et al., 2019. Naturally occurring nervonic acid ester improves myelin synthesis
788 by human oligodendrocytes. Cells. 8(8), 786.
789 Li, J.X., Xu, J., Ruan, J.C., Meng, H.M., Su, H., Han, X.F., et al., 2020. Disrupting a
790 phospholipase A2 gene increasing lipid accumulation in the oleaginous yeast

Yarrowia lipolytica. J. Appl. Microbiol. 130, 100-108.

Li, Q., Chen, J., Yu, X.Z., Gao, J.M., 2019. A mini review of nervonic acid: source, production, and biological functions. Food. Chem. 301, 125286.

Liu, F., Wang, P. D., Xiong, X. J., Zeng, X. H., Zhang, X. B., Wu, G., 2021. A review of nervonic acid production in plants: prospects for the genetic engineering of high nervonic acid cultivars plants. Front. Plant Sci. 12, 626625.

Liu, Q., Siloto, R.M., Lehner, R., Stone, S.J., Weselake, R.J., 2012. Acyl-CoA:diacylglycerol acyltransferase: molecular biology, biochemistry and biotechnology. Prog. Lipid Res. 51, 350-377.

Li, X.P., Liu, J.P., Chen, G.Y., Zhang, J.G., Wang, C.B., Liu, B., 2019. Extraction and purification of eicosapentaenoic acid and docosahexaenoic acid from microalgae: A critical review. Algal Res. 43, 101619.

Luo, Z.S., Liu, N., Lazar, Z., Chatzivasileiou, A., Ward, V., Chen, J., et al., 2020. Enhancing isoprenoid synthesis in *Yarrowia lipolytica* by expressing the isopentenol utilization pathway and modulating intracellular hydrophobicity. Metab. Eng. 61, 344-351.

Lv, E.M., Ding, S.X., Lu, J., Yi, W.M., Ding, J.C., 2020. Separation and purification of fatty acids by membrane technology: a critical review. Int. J. Chem. React. Eng. 18(9): 20190224.

Malina, C., Larsson, C., Nielsen, J., 2018. Yeast mitochondria: an overview of mitochondrial biology and the potential of mitochondrial systems biology. FEMS Yeast Res. 18(5), foy040.

Mansour, M.P., 2005. Reversed-phase high-performance liquid chromatography purification of methyl esters of C_{16}-C_{28} polyunsaturated fatty acids in microalgae, including octacosaoctaenoic acid [28:8(n-3)]. J. Chromatogr. A. 1097, 54-58.

Miller, K. K., Alper, H. S., 2019. *Yarrowia lipolytica*: more than an oleaginous workhorse. Appl. Microbiol. Biot. 103, 9251-9262.

Muhammad, A., Feng, X.D., Rasool, A., Sun, W.T., Li, C., 2020. Production of plant natural products through engineered *Yarrowia lipolytica*. Biotechnol. Adv. 43, 10755.

Notredame C., Higgins DG., Heringa J., 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. 302(1):205-217

Ojemalm, K., Botelho, S. C., Stüdle, C., Von Heijne, G., 2013. Quantitative analysis of SecYEG-mediated insertion of transmembrane α-helices into the bacterial inner membrane. J. Mol. Biol. 425, 2813-2822.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., Waldo, G.S., 2006. Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. 24, 79-88.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al., 2004. UCSF chimera—A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605-1612.

Omasits, U., Ahrens, C. H., Müller, S., Wollscheid, B., 2014. Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics (Oxford, England). 30(6), 884–886.

Qiao, K.J., Wasylenko, T.M., Zhou, K., Xu, P., Stephanopoulos, G., 2017. Lipid production in *Yarrowia lipolytica* is maximized by engineering cytosolic redox metabolism. Nat. Biotechnol. 35, 173-177.

Qiao, Q., Xue, W., Feng, Z., 2018. Variability of seed oil content, fatty acid composition, and nervonic acid content in *Acer truncatum*, native to 14 regions of China. Grasas. Aceites. 69(4), 274.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., et al., 2011. Integrative genomics viewer. Nat. Biotechnol. 29, 24-26.

Schwank, S., Hoffmann, B., Schuller, H.J., 1997. Influence of gene dosage and autoregulation of the regulatory genes INO2 and INO4 on inositol/choline-repressible gene transcription in the yeast *Saccharomyces cerevisiae*. Curr. Genet. 31, 462-468.

Schwede, T., Kopp, J., Guex, N., Peitsch, M. C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381-3385.

Singh, B.K., Bala, M., Rai, P.K., 2014. Fatty acid composition and seed meal characteristics of *Brassica* and allied genera. Natl. Acad. Sci. Lett. 37, 219-226.

Smith, M.A., Dauk, M., Ramadan, H., Yang, H., Seamons, L.E., Haslam, R.P., et al., 2013. Involvement of *Arabidopsis* ACYL-COENZYME A DESATURASE-LIKE2 (At2g31360) in the biosynthesis of the very-long-chain monounsaturated fatty acid components of membrane lipids. Plant. Physiol. 161(1), 81-96.

Song WT., Zhang K., Xue T., Han JR., Peng FD., Ding CG., et al., 2022. Cognitive improvement effect of nervonic acid and essential fatty acids on rats ingesting

857 *Acer truncatum* Bunge seed oil revealed by lipidomics approach. Food Funct.
858 13(5):2475-2490.
859 Sun, M. L., Madzak, C., Liu, H. H., Song, P., Ren, L. J., Huang, H., et al., 2017.
860 Engineering *Yarrowia lipolytica* for efficient gamma-linolenic acid production.
861 Biochem. Eng. J. 117, 172-180.
862 Sun, X.Q., Pang, H., Li, M.M., Peng, B., Guo, H.S., Yan, Q.Q., et al., 2013.
863 Evolutionary pattern of the *FAE1* gene in *Brassicaceae* and its correlation with the
864 erucic acid trait. PloS. One. 8(12), e83535.
865 Schwartz, C., Shabbir-Hussain, M., Frogue, K., Blenner, M., Wheeldon, I., 2017.
866 Standardized markerless gene integration for pathway engineering in *Yarrowia*
867 *lipolytica*. ACS Synth. Biol. 6, 402-409.
868 Tai, M., Stephanopoulos, G., 2013. Engineering the push and pull of lipid biosynthesis
869 in oleaginous yeast *Yarrowia lipolytica* for biofuel production. Metab. Eng. 15, 1-
870 9.
871 Tang, T.F., Liu, X.M., Ling, M., Lai, F., Zhang, L., Zhou, Y.H., et al., 2013. Constituents
872 of the essential oil and fatty acid from *Malania oleifera*. Ind. Crop. Prod. 43, 1-5.
873 Taylor, D.C., Falk, K.C., Palmer, C.D., Hammerlindl, J., Babic, V., Mietkiewska, E., et
874 al., 2010. *Brassica carinata* - a new molecular farming platform for delivering bio-
875 industrial oil feedstocks: case studies of genetic modifications to improve very
876 long-chain fatty acid and oil content in seeds. Biofuel. Bioprod. Bior. 4, 538-561.
877 Terluk, M.R., Tieu, J., Sahasrabudhe, S.A., Moser, A., Watkins, P.A., Raymond, G.V.,
878 et al., 2022. Nervonic acid attenuates accumulation of very long-chain fatty acids

and is a potential therapy for adrenoleukodystrophy. Neurotherapeutics. 19(3):1007-1017.

Trenkamp, S., Martin, W., Tietjen, K., 2004. Specific and differential inhibition of very-long-chain fatty acid elongases from *Arabidopsis thaliana* by different herbicides. P. Natl. Acad. Sci. USA. 101, 11903-11908.

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455-461 .

Umemoto, H., Sawada, K., Kurata, A., Hamaguchi, S., Tsukahara, S., Ishiguro, T., Kishimoto, N., 2014. Fermentative production of nervonic acid by *Mortierella capitata* RD000969. J. Oleo. Sci. 63, 671-679.

Wadhwani, N.S., Dhobale, M.V., Mehendale, S.S., Pisal, H.R., Joshi, S.R., 2011. Reduced levels of placental long chain polyunsaturated fatty acids in preterm deliveries. Prostaglandins. Leukot. Essent. Fatty. Acids. 85, 149-153.

Wang, Q. T., Feng, Y. B., Lu, Y. D., Xin, Y., Shen, C., Wei, L., et al., 2021. Manipulating fatty-acid profile at unit chain-length resolution in the model industrial oleaginous microalgae *Nannochloropsis*. Metab. Eng. 66, 157-166.

Xiao, A.F., Jiang, X.L., Ni, H., Yang, Q.M., Cai, H.N., 2015. Study on the relationship between intracellular metabolites and astaxanthin accumulation during *Phaffia rhodozyma* fermentation. Electron. J. Biotech. 18, 148-153.

Xie, D. M., Jackson, E. N., Zhu, Q., 2015. Sustainable source of omega-3 eicosapentaenoic acid from metabolically engineered *Yarrowia lipolytica*: from

fundamental research to commercial production. Appl. Microbiol. Biot. 99, 1599-1610.

Xin, Y., Shen, C., She, Y. T., Chen, H., Wang, C., Wei, L., et al., 2019. Biosynthesis of Triacylglycerol molecules with a tailored PUFA profile in industrial microalgae. Mol. Plant. 12, 474-488.

Xu, F., Fan, Y., Miao, F.H., Hu, G.R., Sun, J., Yang, G.F., et al., 2018. Naphthylacetic acid and tea polyphenol application promote biomass and lipid production of nervonic acid-producing microalgae. Front. Plant. Sci. 9, 506.

Xu, P., Qiao, K.J., Ahn, W.S., Stephanopoulos, G., 2016. Engineering *Yarrowia lipolytica* as a platform for synthesis of drop-in transportation fuels and oleochemicals. P. Natl. Acad. Sci. USA. 113, 10848-10853.

Xue, Z.X., Sharpe, P.L., Hong, S.P., Yadav, N.S., Xie, D.M., Short, D.R., et al., 2013. Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. Nat. Biotechnol. 31, 734-740.

Yang, K.X., Qiao, Y.G., Li, F., Xu, Y., Yan, Y.J., Madzak, C., et al., 2019. Subcellular engineering of lipase dependent pathways directed towards lipid related organelles for highly effectively compartmentalized biosynthesis of triacylglycerol derived products in *Yarrowia lipolytica*. Metab. Eng. 55, 231-238.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang, Y., 2015. The I-TASSER Suite: Protein structure and function prediction. Nat. Methods, 12, 7-8.

Yang, T., Yu, Q., Xu, W., Li, D.Z., Chen, F., Liu, A., 2018. Transcriptome analysis reveals crucial genes involved in the biosynthesis of nervonic acid in woody

Malania oleifera oilseeds. BMC Plant Biol. 18, 247.

Yu, T., Zhou, Y.J.J., Wenning, L., Liu, Q.L., Krivoruchko, A., Siewers, V ., et al., 2017. Metabolic engineering of *Saccharomyces cerevisiae* for production of very long chain fatty acid-derived chemicals. Nat. Commun. 8, 1-11.

Yuan, C., Liu, J.H., Fan, Y., Ren, X.H., Hu, G.R., Li, F.L., 2011. *Mychonastes afer* HSO-3-1 as a potential new source of biodiesel. Biotechnol. Biofuels. 4(1), 47.

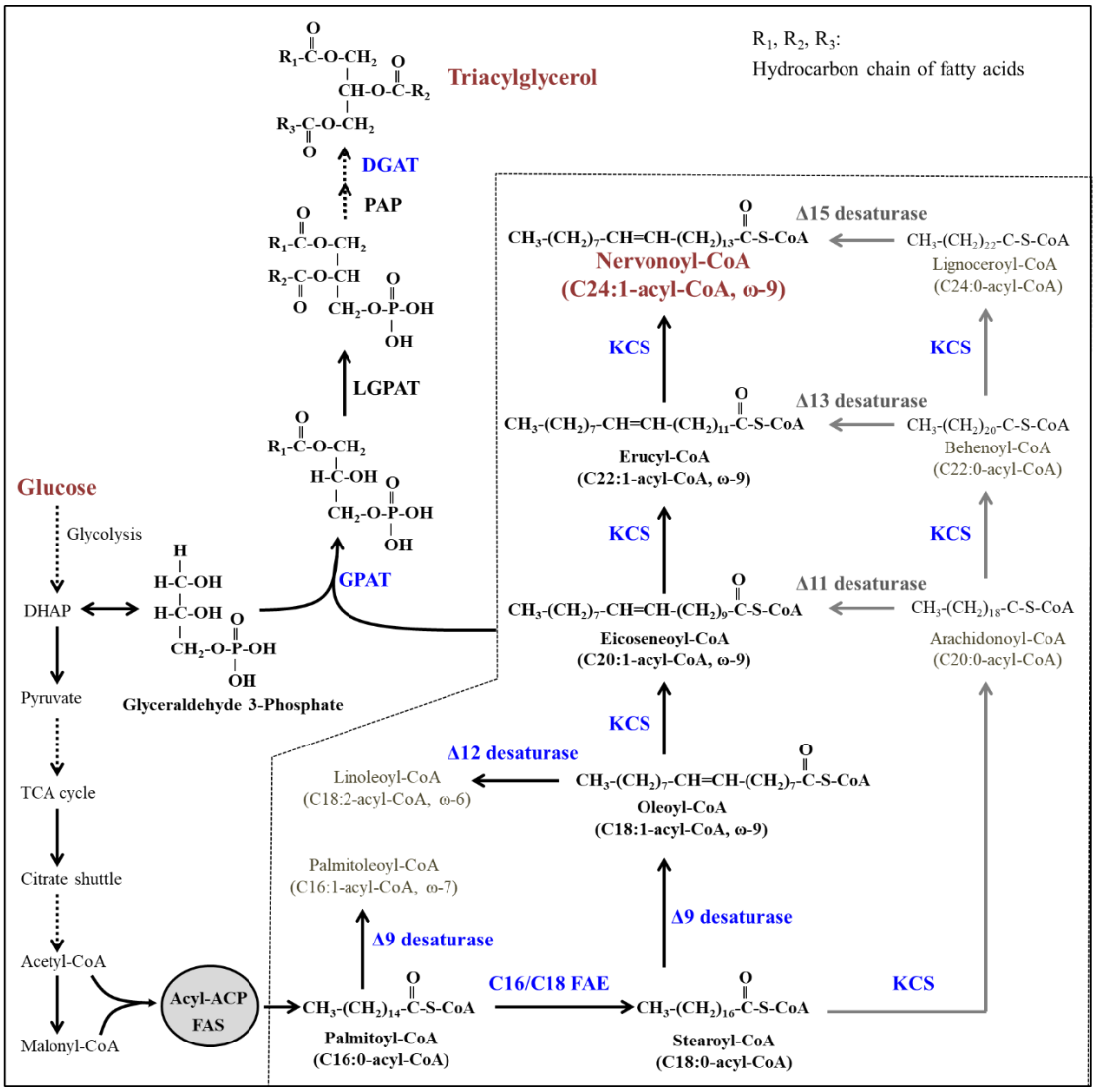
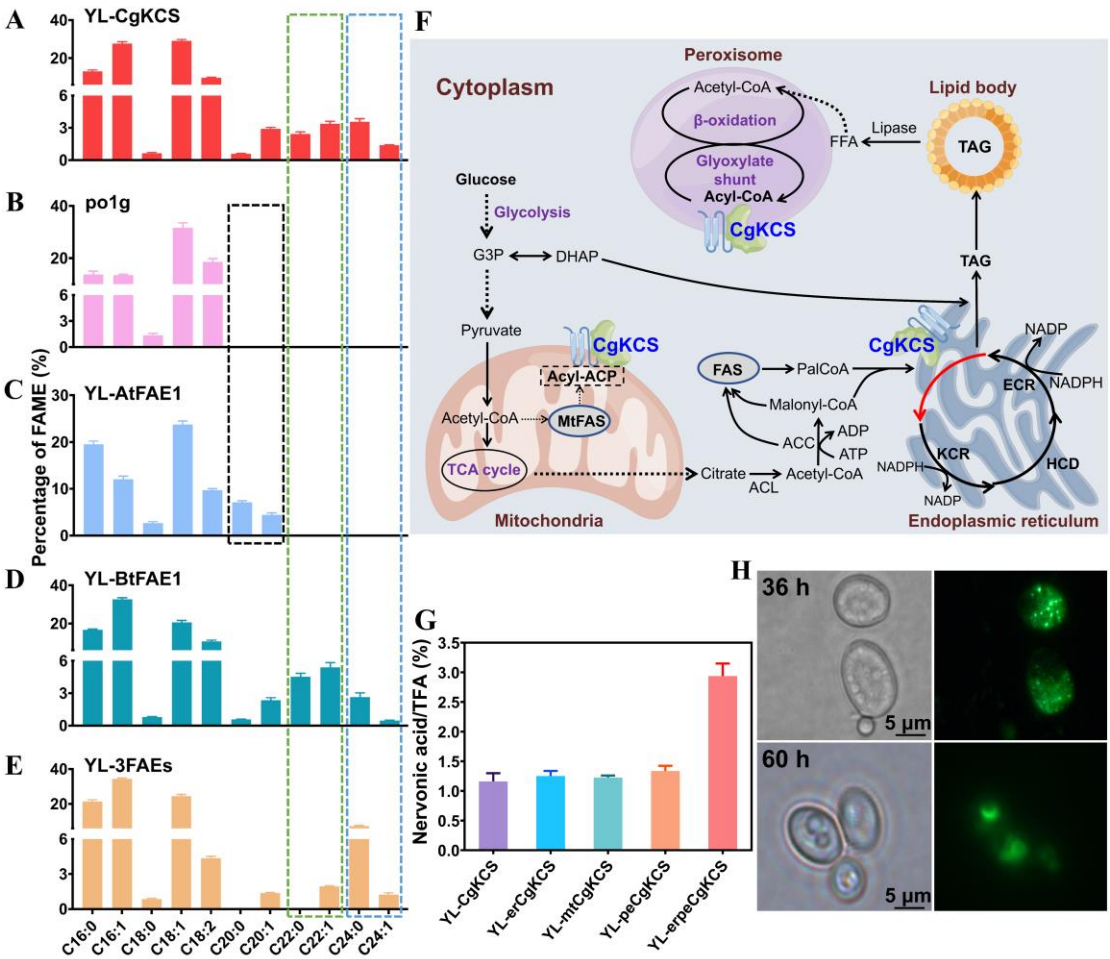


Fig. 1. Biosynthetic pathways for nervonic acid production in engineered *Y. lipolytica*. The solid line-box and the dotted line-box denote native pathways and heterologous pathways, respectively. The grey arrows represent competing pathways of the biosynthesis of nervonic acid. The dotted arrows show the hypothetical desaturation steps of fatty acids. FAE, fatty acid elongases.



935

936 **Fig. 2.** Evaluation of fatty acid elongases for production of nervonic acid. (A)
937 Expression of *CgKCS* in *Y. lipolytica* po1g. (B) The control strain po1g. (C) Expression
938 of *AtFAE1* in po1g. (D) Expression of *BtFAE1* in po1g. (E) Co-expression of *AtFAE1*,
939 *BtFAE1* and *CgKCS* in po1g. (F) Diagram of fatty acid metabolism related to multiple
940 organelles and subcellular localization of *CgKCS*. ACL, ATP-citrate lyase; ACC,
941 acetyl-CoA carboxylase; CgKCS, β-ketoacyl-CoA synthase of *C. graeca*; FAS, fatty
942 acid synthase; FFA, free fatty acids; PalCoA, palmitoyl-CoA. (G) The relative amount
943 of nervonic acid content in the TFA produced by the strains expressing CgKCS and
944 CgKCS fused with organelle targeting signals. Cultivations were carried out for 144 h
945 at 250 rpm and 28 ºC in flasks. Data are mean ± s.d. from three replicates. (H)

946 Microscopic observation of the strains expressing the fused protein CgKCS-sfGFP for
947 evaluating the location of CgKCS in *Y. lipolytica*. CgKCS-sfGFP exhibited both
948 punctate and accumulated localization at the early stage of cultivation (36 h) and
949 gathered in lipid droplets at lipid accumulation stage (60 h).

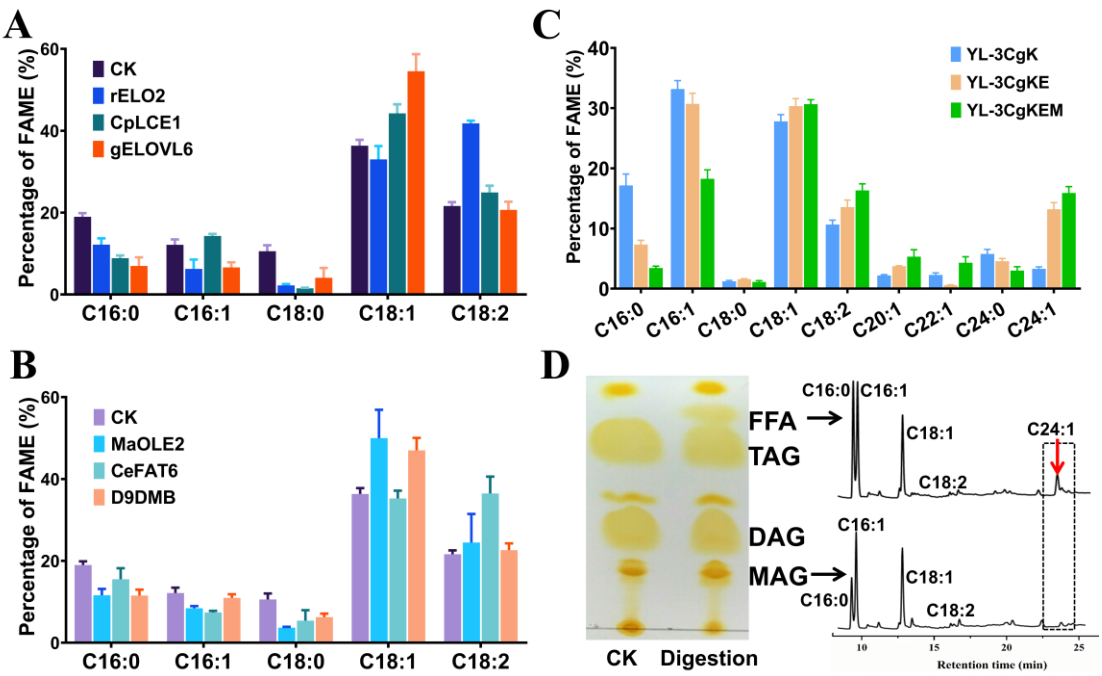


Fig. 3. Expression of C16:0-acyl-CoA-specific FAEs and C18:0-acyl-CoA-specific fatty acid desaturases to improve nervonic acid production. (A) Expression of genes encoding C16:0-acyl-CoA-specific FAEs. (B) Expression of genes encoding C18:0-acyl-CoA-specific fatty acid desaturases. (C) Co-expression of *gELOVL6*, *MaOLE2*, and *CgKCS*. (D) Determination of the distribution of nervonic acid in TAG by hydrolysis using a 1,3-specific lipase as well as TLC and GC-MS analysis. Cultivations were carried out for 144 h at 250 rpm and 28 ºC in flasks. Data are mean ± s.d. from three replicates. Statistical analysis was performed by two-way ANOVA and Tukey's multiple comparison test.

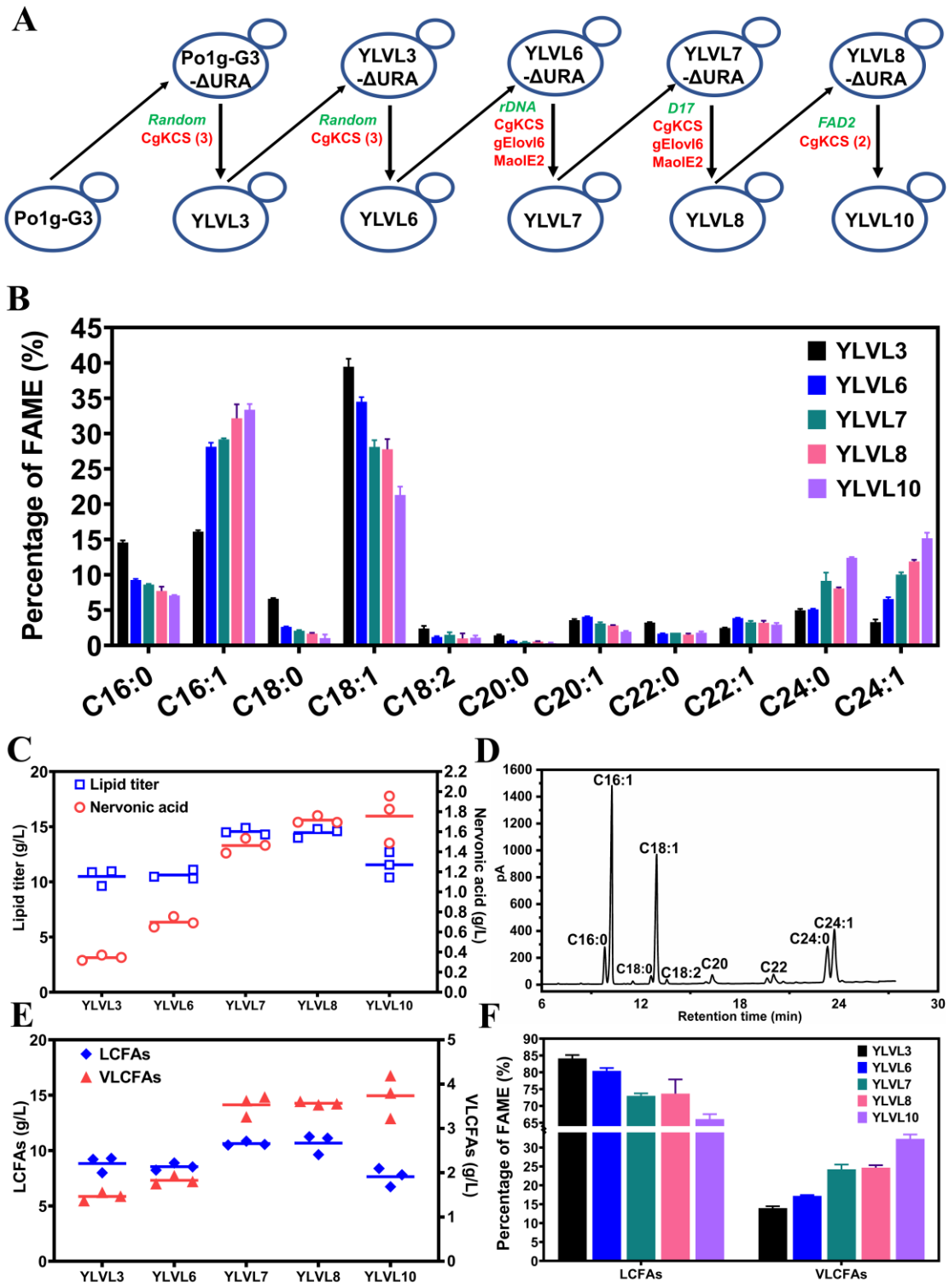
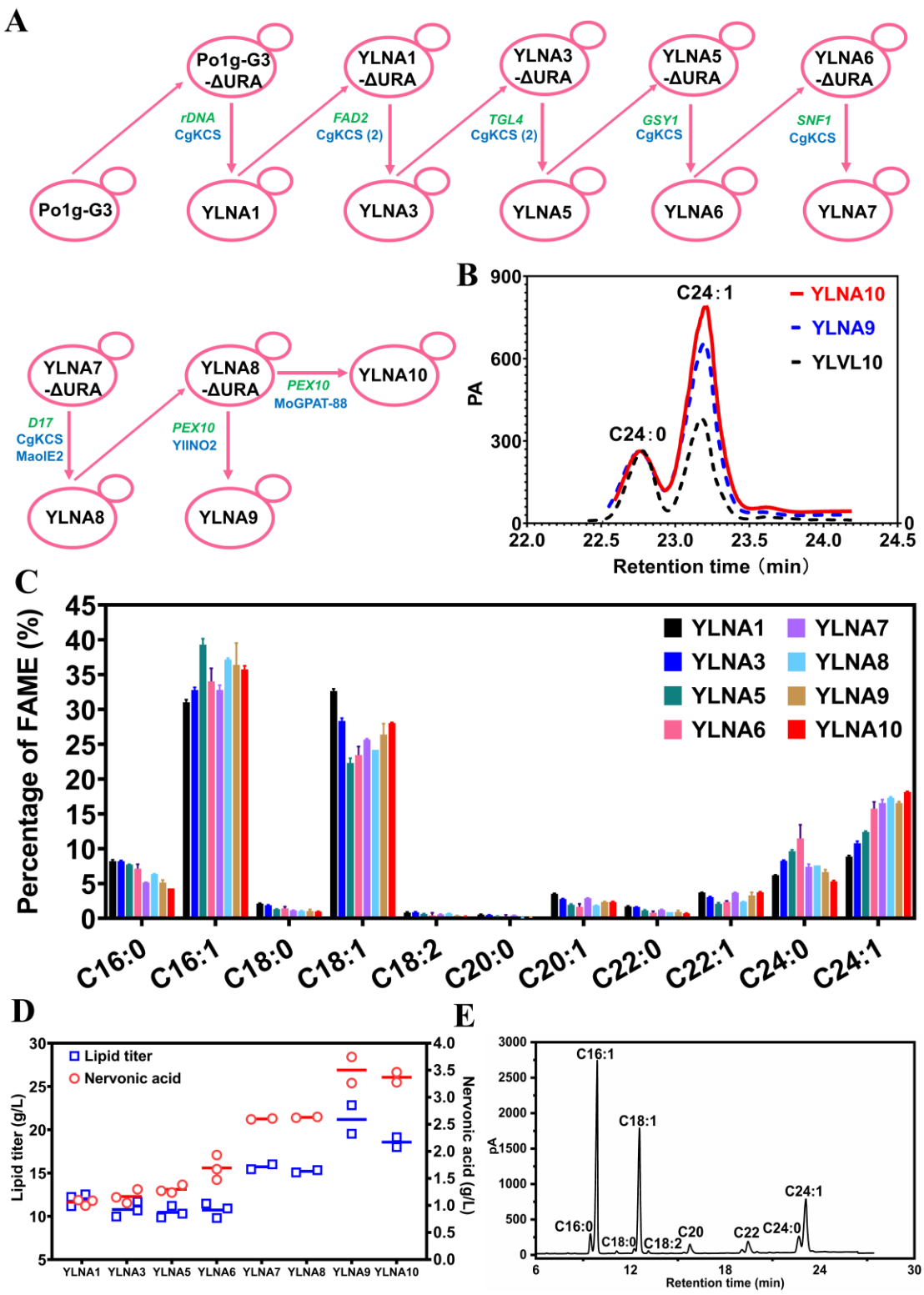


Fig. 4. Construction of VLCFAs-producing strains expressing *CgKCS*, *gELOVL6* and *MaOLE2* by random insertion and homologous recombination. (A) Schematic diagram showing the construction of YLVL strains. The overexpressed genes are indicated by red symbol. The red numbers in parentheses represented the gene copies more than one

on plasmids. The green italics symbol denotes genes integrated into the genomes by random integration or homologous recombination. The results of lipid and VLCFAs biosynthesis in the strains YLVL3, 6, 7, 8 and 10, including the fatty acid profiles (B), the titer of lipid and nervonic acid (C), the fatty acid composition in YLVL10 detected by gas chromatography (D), the titer of LCFAs and VLCFAs (E), and the fraction of LCFAs and VLCFAs (F). The yeasts were cultured for 144 h at 250 rpm and 28 ºC in flasks. Data are mean ± s.d. from three replicates.



974

975 **Fig. 5.** The construction of nervonic acid overproducing strains by homologous

976 recombination. (A) Schematic diagram showing the construction of YLNA strains by

977 expressing the genes *CgKCS*, *MaOLE2*, YlINO2 and MoGPAT-88 at multiple genomic

loci. The overexpressed genes are indicated by blue symbol. The blue numbers in parentheses represented the gene copies more than one on plasmids. The green italics symbol denotes genes integrated into the genomes by homologous recombination. (B) Comparison of the production of nervonic acid and lignoceric acid in YLNA9, YLNA10 and YLVL10. The results of lipid and VLCFAs biosynthesis in the strains YLNA1, 3, and 5 to 10, including the fatty acid profiles (C), the titer of lipid and nervonic acid (D), the fatty acid composition in YLNA10 detected by gas chromatography (E). The yeasts were cultured for 144 h at 250 rpm and 28 ºC in flasks. Data are mean ± s.d. from three replicates.

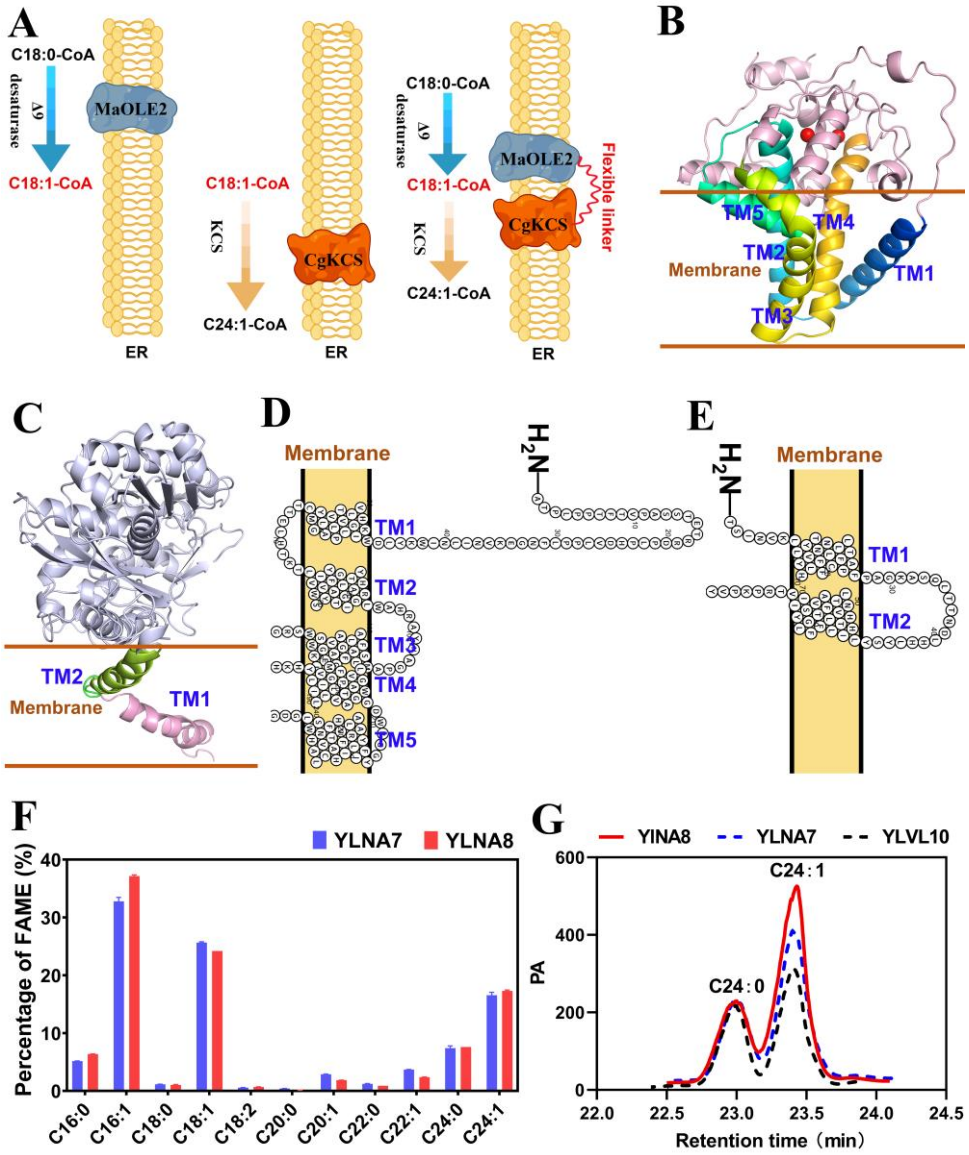


Fig. 6. Overexpression of MaOLE2-CgKCS in ER. (A) Schematic diagram showing the physical locations of MaOLE2 and CgKCS in ER. (B) Three-dimensional architecture of MaOLE2. The red spheres represent the catalytic site of zinc ions. (C) Three-dimensional architecture of CgKCS. (D) Five transmembrane domains in MaOLE2 are organized in a two-dimensional topology. (E) Two transmembrane domains in CgKCS are organized in a two-dimensional topology. (F) Comparison of the fatty acid composition between YLNA7 and YLNA8. (G) Comparison of the production of nervonic acid and lignoceric acid in YLNA7, YLNA8 and YLVL10.

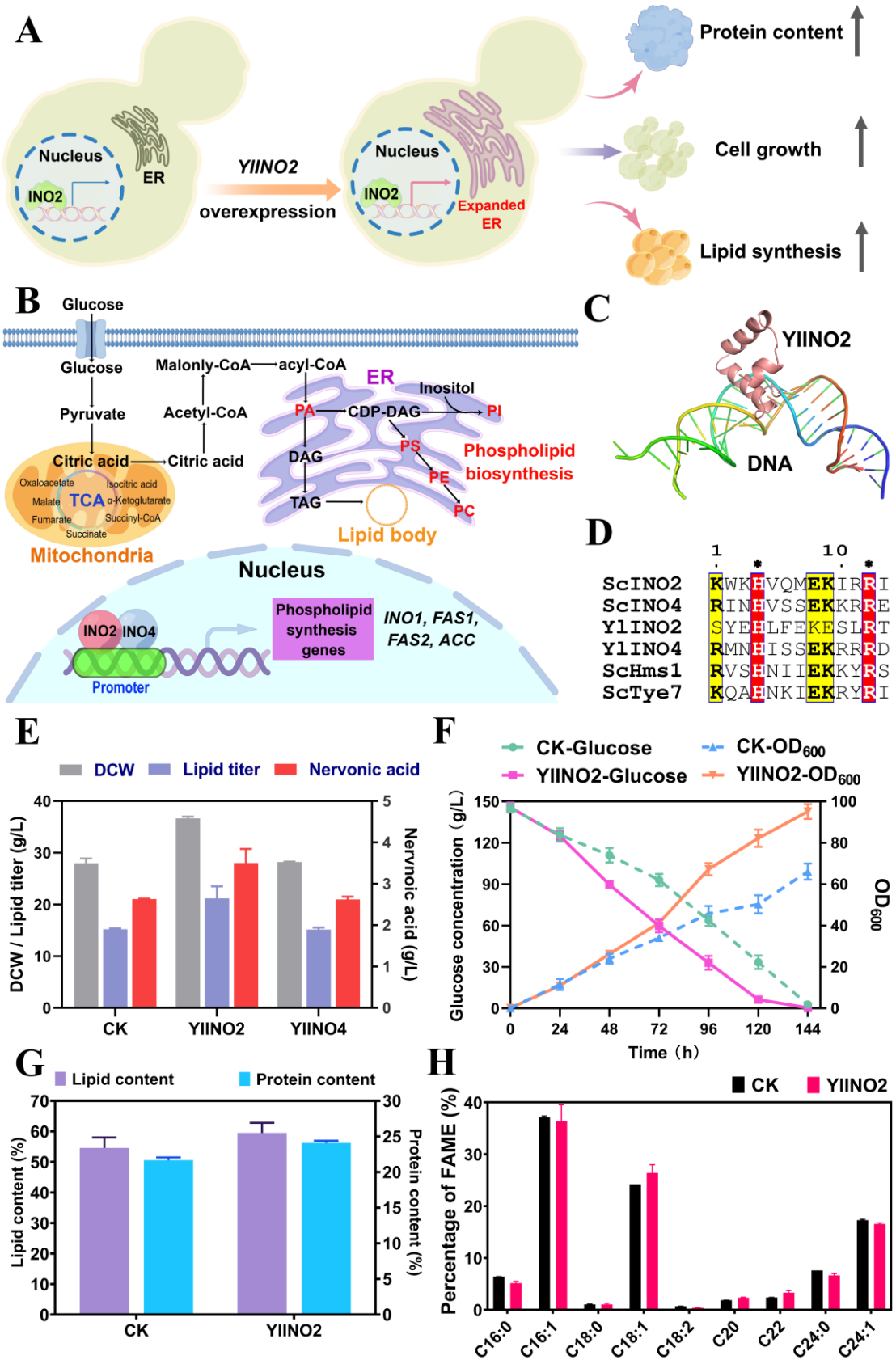


Fig. 7. Identification of an ER regulator YlINO2 promoting the production of lipid. (A) Schematic diagram showing the functions of INO2. (B) The regulatory network of the

INO2/INO4 complex. PA, lipidphosphatidic acid; PI, phosphatidylinositol; PS, phosphatidylserine; PE, phosphatidylethanolamine; PC, phosphatidylcholine. (C) Three-dimensional architecture showing the combination of YlINO2 with DNA. (D) Structure-based multiple sequence alignment of bHLH transcription factors. Sc, *S. cerevisiae*; Yl, *Y. lipolytica*. The star symbol above the red boxes indicates the conserved residues. (E) Comparison of DCW, and the titer of lipid and nervonic acid in the CK strain YLNA8 and the strains YlINO2 (YLNA9) and YlINO4. (F) The cell growth and glucose consuming in the CK strain YLNA8 and YlINO2 (YLNA9). Data are mean ± s.d. from three replicates. (G) The content of lipid and protein in CK (YLNA8) and YlINO2 (YLNA9). (H) Comparison of fatty acid composition in CK (YLNA8) and YlINO2 (YLNA9).

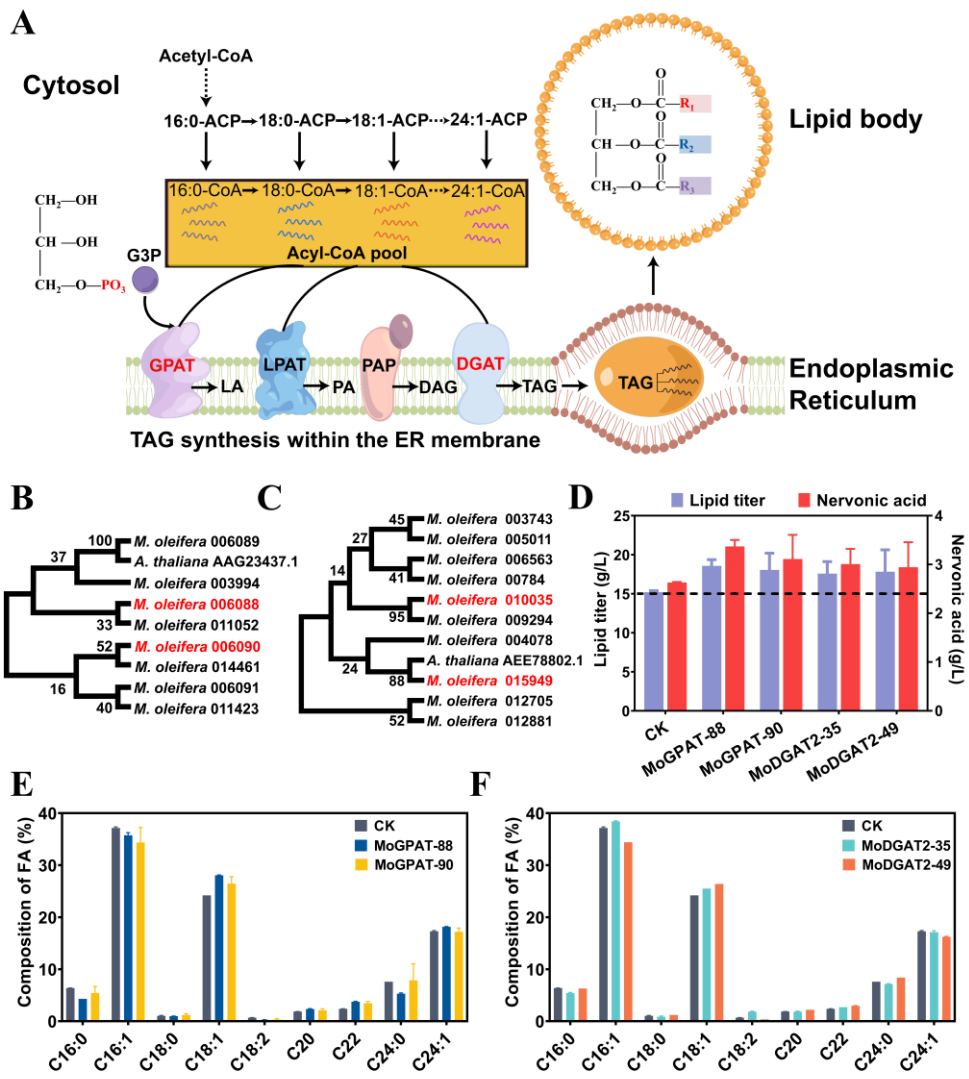


Fig. 8. Nervonic acid and lipids improved by overexpressing *GPAT* and *DGAT2* from *M. oleifera*. (A) The biosynthesis of TAG in ER catalyzed by GPAT, LPAT, PAP and DGAT. G3P, glycerol 3-phosphate; GPAT, glycerol-3-phosphate acyltransferase; LPAT, lysophosphlipid acyltransferase; PAP, phosphatidic acid phosphatase; DGAT, diacylglycerol acyltransferase. Phylogenetic trees of the acyltransferases GPAT (B) and DGAT2 (C) from *M. oleifera* constructed by the Neighbor-Joining method using the software MEGA11. The titer of lipid and nervonic acid in the acyltransferases overexpressed strains (D). The fatty acid composition in GPAT (E) and DGAT2 (F) overexpressed strains.

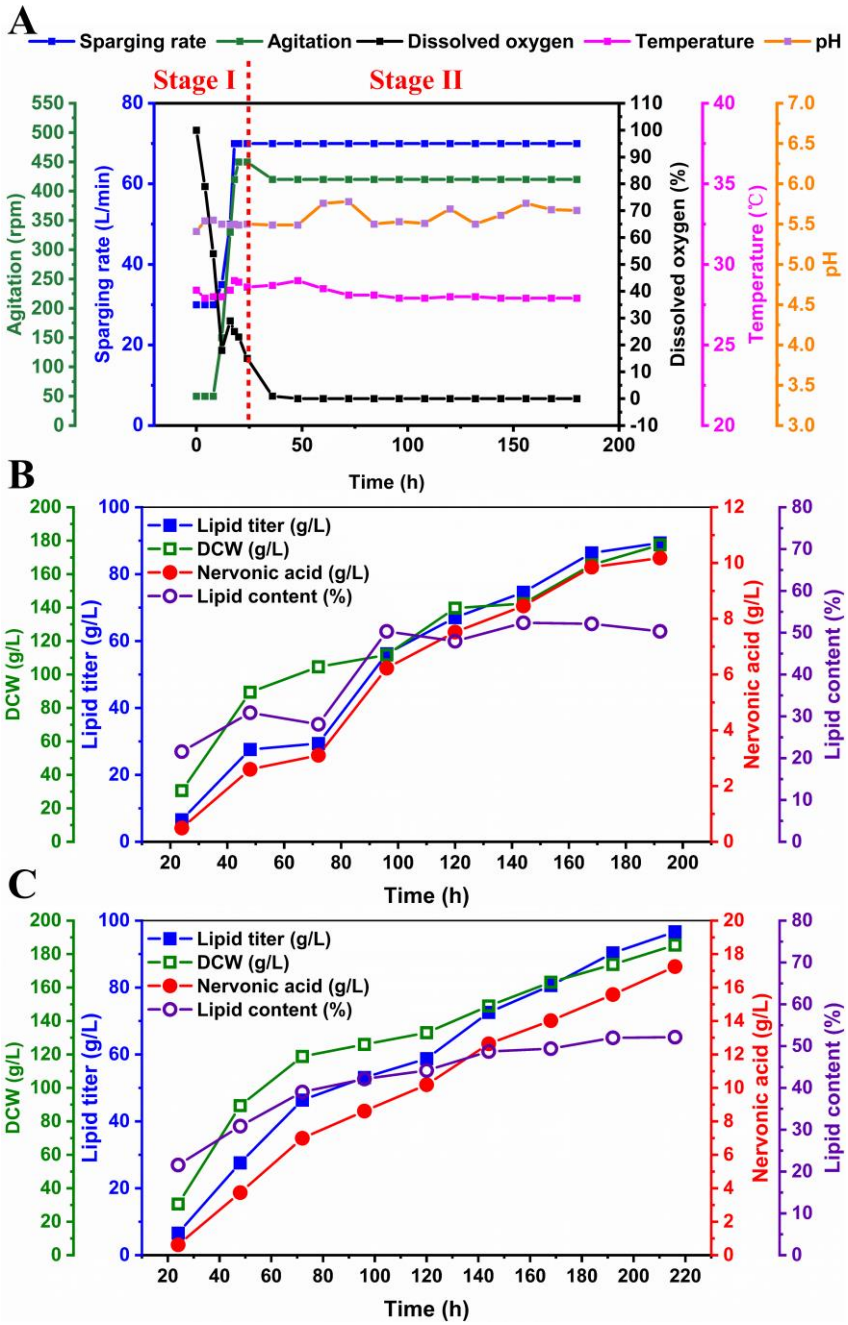


Fig. 9. Fed-batch fermentation in a 50-L reactor. (A) Conditions for controlling the two-stage fermentation process. (B) Fed-batch fermentation of YLVL10 in a 50-L bioreactor. (C) Fed-batch fermentation of YLNA9 in a 50-L bioreactor. Fermentation was conducted at 28 °C, pH 5.5. Substrate feeding to 150 g/L in every 12 h. DCW, dry cell weight.

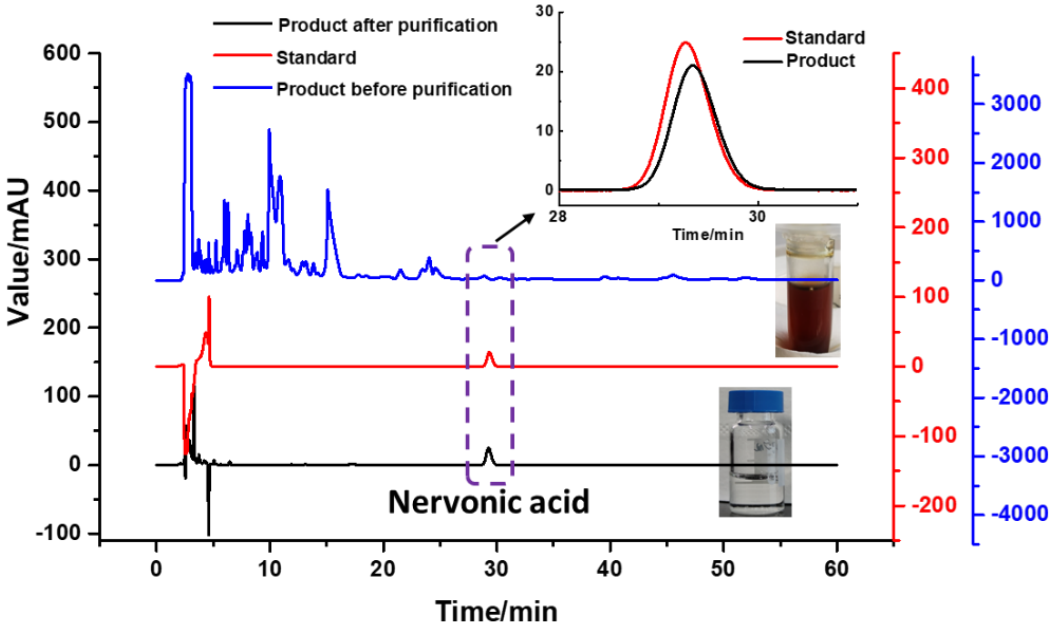


Fig. 10. Comparison of nervonic acid products before and after purification determined by HPLC. The brown liquid showed the initial lipids. The transparent liquid presented the purified nervonic acid solution.