## SUPPLEMENTAL FILE

## 1. Methods

### 1.1 ColabFold Modeling

ColabFold offers a free and user-friendly platform for protein structure prediction. By integrating the rapid homology search capabilities of MMseqs2 with AlphaFold2 or RosettaFold, ColabFold significantly accelerates the modeling process. Its optimized search algorithms and model efficiency allow for predicting up to 1,000 protein structures daily on a server equipped with a single GPU, achieving a 40-60-fold speed improvement (GPU) [1].

To predict structural models of the CYP155 protein, a GPU runtime was selected to enhance inference speed and minimize overall runtime. The single amino acid sequence of CYP155 was entered as the target, and predictions were carried out using Google ColabFold v1.5.5 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb) within the Google Collaboratory environment. Template information was not utilized, and the "Num relax" parameter was set to 1. For the modeling process, ColabFold structures were generated using the multiple sequence alignment (MSA) mode with MMseqs2 (parameters: UniRef+Environmental), msa_mode (parameters: mmseqs2_uniref_env), and pair mode (parameters: unpaired paired), along with other default settings. Five structural models for CYP155 were generated, and the model with the highest confidence score was selected for further analysis. The optimal CYP155 model had a predicted local distance difference test (pLDDT) value of 92.5 and a post-translational modification (pTM) value of 0.89. Amino acid sequence alignment and visualization were performed using Geneious

software.

### 1.2 Molecular Docking

Molecular docking was performed using AutoDockTools 1.5.6 to dock lupeol and heme into CYP155. Among the 10 docking results, the complex with lupeol's C28 position closest to the heme center was chosen for further analysis. The mutant structures of E120Q and R155Q were generated with PyMOL 2.5.5. Prior to molecular dynamics simulations, the energy of the E120Q and R155Q mutant complex structures was minimized.

### 1.3 Molecular Dynamics Simulation

Molecular dynamics simulations were carried out using Amber 24 (San Francisco, CA, USA). The system's force field parameters were calculated with the ff19SB force field [2] and HEME force field parameters were constructed using Amber's standard force field (http://amber.manchester.ac.uk/). Solvation was carried out using the TIP3P water model, and counterions were added to neutralize the system.

After energy minimization, the system was heated from 0 K to 301.15 K (28°C) over 500 ps. System restraints were applied in the NVT ensemble, followed by pre-equilibration at 301.15 K (28°C). The final production simulations were conducted in the NPT ensemble under periodic boundary conditions for 200 ns, with all covalent bonds involving hydrogen atoms constrained using the SHAKE algorithm.

### 1.4 LC-MS Analysis Conditions

The analysis was carried out using a Thermo Scientific Q Exactive mass spectrometer paired with an Ultimate 3000 UPLC system. For mass spectrometry, the spray voltage

was set at 3200 V, and the capillary temperature was maintained at 300°C. The sheath and auxiliary gas flows were set to 40 Arb and 8 Arb, respectively. The maximum spray current was 100 μA, and the probe heater temperature was set to 300°C. Electrospray ionization (ESI-MS) was used as the ionization source.

For liquid chromatography, an Accucore$^{TM}$ aQ column (100 × 2.1 mm, 2.6 μm) was employed, with the column temperature set at 30°C. The detection wavelength was set to 202 nm, and the flow rate was 0.2 mL/min. The mobile phase consisted of solvent A (methanol) and solvent B (0.1% formic acid in water), with a gradient composition of A: B = 82:18.

**1.5 Quantification of acetyl-CoA and NADPH/NADP+ Ratio.**

Recombinant *Y. lipolytica* strains are cultured and pre-cultured. The pre-culture is diluted into 50 mL of fresh medium to an initial $OD_{600}$ of 0.05 and grown until $OD_{600}$ reaches 0.4. Cells are then harvested by centrifugation at 12,000 rpm for 3 minutes. Metabolic activity is quenched by adding 10 mL of pre-chilled (−80°C) methanol, followed by a second centrifugation to remove the supernatant. The cell pellet is extracted with 2 mL of boiling ethanol for 15 minutes and vortexed with glass beads for 20 minutes to disrupt the cells and release metabolites. The mixture is then centrifuged to separate the supernatant, which is vacuum-dried and resuspended in 200 μL of double-distilled water ($ddH_2O$). Acetyl-CoA concentration is determined using the Acetyl-Coenzyme A Assay Kit MAK039 (Sigma-Aldrich, USA). The reported acetyl-CoA levels correspond to a 30-fold concentrated sample (10 mL culture at $OD_{600}$ of 0.4 concentrated to 200 μL).

The NADPH/$NADP^+$ ratio was measured using the NADPH/$NADP^+$ Detection Kit (WST-8 method) (Beyotime Biotechnology). This assay was employed to assess the intracellular redox metabolic status of the strains. NADPH was extracted from the samples, and glucose-6-phosphate dehydrogenase catalyzed the reduction of $NADP^+$ to NADPH. The concentration of $NADP^+$ was then determined by measuring the absorbance at 450 nm using a microplate reader, following the instructions provided in the kit.

**Table. S1 Comparative analysis of BA biosynthesis in different studies.**

| Host strain | Modification strategy | Titer in shake flask culture (mg/L) | Reference |
|---|---|---|---|
| *S. cerevisiae* | CrAS and CrAO enzymes, derived from Rosa moschata leaves, were identified and co-expressed with AtLUP1 from *Arabidopsis thaliana.* | 0.1 | [3] |
| | The BA and fatty acid pathways were optimized by using different promoter combinations for the genes HMGR, ERG9, CrAO, LUS, and HFA1. The flux through the competitive ergosterol pathway was reduced by down-regulating ERG7 using an inhibitory promoter. | 10 | [4] |
| | Supply of NADPH and oxygen through co-expression of mutant 2,3-butanediol dehydrogenase (*mBDH1*) and *Vhb*. | 12 | [5] |
| | A gene for lupeol C-28 oxidase (BPLO) was isolated from birch. The productivity of two yeast strains, WAT11 and CEN.PK, was compared. The loss of Gal80p function on the galactose-inducible promoter impacted the expression of synthetic BA genes. | 1 | [6] |
| | Co-expression of *AtLUP1* from *Arabidopsis thaliana* and CYP716A11 from *Catharanthus roseus*, along with *AtATR2* from *Arabidopsis thaliana*, overexpression of the native *ERG1*, and a truncated *HMGR* (*tHMGR*); optimization of extraction and fermentation processes. | 28 | [7] |

| | | | |
|---|---|---|---|
| | RoCYP01 (CYP716A155) and RoCPR1 were identified from *Rosmarinus officinalis* and co-expressed with *AtLUP* from *Arabidopsis thaliana*. Enzymes involved in converting acetyl-CoA to IPP were overexpressed, along with squalene synthase (AtSQS2), squalene-epoxidase (AtSQE2) from *A. thaliana*, and FPP synthase (SmFPS) from *Salvia miltiorrhiza*. | 193.5 | [3] |
| | Optimize the copy number of key enzymes *ERG1* and *AtLUP1*; increase the supply of precursor acetyl-CoA and the cofactor NADPH by knocking out *MLS1* and overexpressing *GND1*, *TAL1*, and *TKL1*; compartmentalize the BA biosynthesis pathway in peroxisomes. | 210.88 | [8] |
| | Dual engineering of peroxisomes and lipid droplets was performed, optimizing combinations of *CYP* and *CPR*. The endogenous *ERG7* was replaced with a heterologous version to down-regulate ergosterol, and endogenous reductase genes in *S. cerevisiae* were knocked out to balance reducing power. Multi-copy integration of the rate-limiting enzyme BPLO (CYP716A180) and *ATR1* was also carried out. | 77.53 | [9] |
| *Y. lipolytica* | Screening was conducted for effective combinations of cytochrome P450 monooxygenases (CYPs) and NADPH-cytochrome P450 reductases (CPRs). Overexpression of *ERG1*, *ERG9*, and *HMGR* in the MVA module was implemented, alongside strategies to increase acetyl-CoA levels through overexpression of acetyl-CoA synthetase or enhancing the β-oxidation pathway. Additionally, redox cofactor supply was boosted by introducing NADPH or NADH-producing enzymes, Rtme and EMT. | 51.87 | [10] |
| | Co-expressing *AtLUP1, CYP716A12, and AtCPR1*; overexpression of *ERG10, HMGR, ERG20, ERG9,* and *ERG1*. | 26.53 | [11] |
| | Modular pathway optimization: introducing the NOG pathway and the IUP pathway to enhance acetyl-CoA and IPP supply; *LUS* endoplasmic reticulum localization, | 271.3 | This study |

enzyme engineering to improve CYP716A155 enzyme activity, and organelle engineering along with the subcellular dynamics of MCSs to enable efficient compound exchange between organelles; Enhance cytosolic NADPH supply (overexpress *GPD1* from *Clostridium acetobutylicum* and *MCE2* from *Mucor circinelloides*); and balancing carbon flux.

**Table. S2 Free energy calculation of WT, E120Qand R155Q.**

| Associative free energy calculation (kcal/mol) | | | |
|---|---|---|---|
| | WT | E120Q | R150Q |
| ΔVDWAALS | -44.2830±5.48 | -55.9847±1.93 | -51.5416±2.64 |
| ΔEEL | -0.0035±1.45 | 0.0179±1.27 | -0.0330±1.64 |
| ΔEPB | −12.7594±4.43 | 17.6468±2.66 | 16.9068±2.13 |
| ΔENPOLAR | -4.5362±0.17 | -4.5156±0.09 | -5.1730±0.07 |
| ΔGGAS | -44.2795±5.76 | -55.9668±2.29 | -51.5745±3.45 |
| ΔGSOLV | 8.2232±4.40 | 13.1312±2.64 | -11.7337±2.10 |
| ΔTOTAL | -36.0563±3.06 | -42.8356±3.03 | -39.8408±2.11 |

ΔVDWAALS: Van der Waals energy.
ΔEEL: electrostatic energy.
ΔEPB: energy calculated by Poisson Boltzmann.
ΔENPOLAR: non-polar energy.
ΔGGAS: molecular mechanical term energy (meteorological energy) =ΔVDWAALS + ΔEEL.
ΔGSOLV: solvation energy = ΔEPB + ΔENPOLAR.
ΔTOTAL: total energy = ΔGGAS + ΔGSOLV.

**Table. S3 Strains and plasmids used in the study**

| Strain name | Description of strains and plasmids | Source |
|---|---|---|
| Po1h | *MatA, ura3-302, xpr2-322, axp1-2, Ura⁻, ΔAEP, ΔAXP, Suc⁺* | [12] |
| Y L 7166-01 | Po1h derivative, integration site AXP, harboring $P_{hp4d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ | This study |
| Y L 7166-001 | Po1h derivative, integration site AXP, harboring $P_{hp4d}$-*CYP716A180*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ | This study |
| Y L 7166-02 | YL7166-01 derivative, integration site AXP, harboring $P_{hp4d}$-*CYP716A155*-T2A-*AtATR1*-2A-*RcLUS*-$T_{xpr2}$ and $P_{hp4d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-002 | YL7166-02 derivative, integration site AXP, harboring $P_{hp4d}$-*CYP716A180*-T2A-*AtATR1*-2A-*RcLUS*-$T_{xpr2}$ and *tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-03 | YL7166-01 derivative, integration site AXP, harboring $P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-04 | Po1h derivative, integration site AXP, harboring $P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-2A-*RcLUS*-$T_{xpr2}$ and $P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-05 | YL7166-04 derivative, integration site AXP, harboring 2copy $P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy $P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-06 | YL7166-05 derivative, integration site AXP, harboring 3copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$*vtHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ | This study |
| Y L 7166-08 | YL7166-04 derivative, integration site AXP, harboring $P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-2A-*RcLUS*-$T_{xpr2}$ and $P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ and $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ | This study |
| Y L 7166-09 | YL7166-04 derivative, integration site AXP, harboring $P_{hp4d}$-*LmPK*-$T_{xpr2}$ and $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ | This study |

| | | |
|---|---|---|
| Y L 7166-10 | YL7166-05 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*LmPK*-$T_{xpr2}$ and $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ | This study |
| Y L 7166-11 | YL7166-05 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ and $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ | This study |
| Y L 7166-12 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG10*-$T_{xpr2}$ | This study |
| Y L 7166-13 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG13*-$T_{xpr2}$ | This study |
| Y L 7166-14 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG12*-$T_{xpr2}$ | This study |
| Y L 7166-15 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG8*-$T_{xpr2}$ | This study |
| Y L 7166-16 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG19*-$T_{xpr2}$ | This study |
| Y L 7166-17 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IDI*-$T_{xpr2}$ | This study |

| | | |
|---|---|---|
| Y L 7166-18 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*ERG20*-$T_{xpr2}$ | This study |
| Y L 7166-19 | YL7166-011 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ | This study |
| Y L 7166-20 | YL7166-019 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*-$T_{xpr2}$ | This study |
| Y L 7166-22 | YL7166-019 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ | This study |
| Y L -01 | Po1h derivative, integration site AXP, harboring $P_{hp4d}$-*RcLUS*-$T_{xpr2}$ and $P_{hp4d}$- *CYP (CYP716A155)*-$T_{xpr2}$ and $P_{hp4d}$-*CPR* (*AtATR1*)-$T_{xpr2}$ | This study |
| Y L 7166-39 | Po1h derivative, integration site AXP, harboring $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$ and $P_{hp4d}$- *CYP(tCYP716A155)* -$T_{xpr2}$ and $P_{hp4d}$-*CPR*(*AtATR1)* -$T_{xpr2}$ | This study |
| Y L 7166-40 | Po1h derivative, integration site AXP, harboring $P_{hp4d}$-*LUS-SKL*-$T_{xpr2}$ and $P_{hp4d}$- *tCYP (tCYP716A155)-SKL*-$T_{xpr2}$ and $P_{hp4d}$-*tCPR*(t*AtATR1)-SKL*-$T_{xpr2}$ | This study |
| Y L 7166-47 | YL7166-022 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$ | This study |

| | | |
|---|---|---|
| Y L 7166-48 | YL7166-022 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+ $P_{hp4d}$-*ERG20*-$T_{xpr2}$ | This study |
| Y L 7166-58 | YL7166-048 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ | This study |
| Y L 7166-60 | YL7166-058 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ | This study |
| Y L 7166-61 | YL7166-060 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2+}$ $P_{hp4d}$-*ylYEF1*-$T_{xpr2}$ | This study |
| Y L 7166-62 | YL7166-060 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE2*-$T_{xpr2}$ | This study |

| | | |
|---|---|---|
| Y L 7166-71 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE2*-$T_{xpr2}$ and integration site D9, $P_{hp4d}$-*PXA1*-$T_{xpr2}$+$P_{hp4d}$-*FAT1*-$T_{xpr2}$+$P_{hp4d}$-*ANT1*-$T_{xpr2}$ | This study |
| Y L 7166-72 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE2*-$T_{xpr2}$ and integration site D9,$P_{hp4d}$-*PXA1*-$T_{xpr2}$+$P_{hp4d}$-*FAT1*-$T_{xpr2}$+$P_{hp4d}$-*PXA2*- $T_{xpr2}$ | This study |
| Y L 7166-73 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE*$_2$-$T_{xpr2}$ and integration site D9,$P_{hp4d}$-*PXA1*-$T_{xpr2}$+$P_{hp4d}$-*FAT1*-$T_{xpr2}$ | This study |
| Y L 7166-74 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+$P_{hp4d}$-*MCE2*-$T_{xpr2}$ and integration site *F16*,$P_{hp4d}$-*POT1*-$T_{xpr2}$+$P_{hp4d}$-*PAT1*- | This study |

| | | |
|---|---|---|
| | $T_{xpr2}$ | |
| Y L 7166-75 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE*$_2$-$T_{xpr2}$ and integration site *F16*,$P_{hp4d}$-*POT1*-$T_{xpr2}$+$P_{hp4d}$-*PAT1*-$T_{xpr2}$+$P_{hp4d}$-*POX2*-$T_{xpr2}$ | This study |
| Y L 7166-76 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE*$_2$-$T_{xpr2}$ and integration site *F16*,$P_{hp4d}$-*POT1*-$T_{xpr2}$+$P_{hp4d}$-*PAT1*-$T_{xpr2}$+$P_{hp4d}$-*MFE1*-$T_{xpr2}$ | This study |
| Y L 7166-77 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE*$_2$-$T_{xpr2}$ and integration site *F16*,$P_{hp4d}$-*POT1*-$T_{xpr2}$+$P_{hp4d}$-*PAT1*-$T_{xpr2}$ and integration site *D9*,$P_{hp4d}$-*MFE1*-$T_{xpr2}$+$P_{hp4d}$-*PXA1*-$T_{xpr2}$+$P_{hp4d}$-*FAT1*-$T_{xpr2}$ and integration site *E14,*$P_{hp4d}$-*TGL4*-$T_{xpr2}$+$P_{hp4d}$-*TGL3*-$T_{xpr2}$ | This study |

| | | |
|---|---|---|
| Y L 7166-78 | YL7166-062 derivative, integration site AXP, harboring 2copy$P_{hp16d}$-*CYP716A155*-T2A-*AtATR1*-T2A-*RcLUS*-$T_{xpr2}$ and 2copy$P_{hp16d}$-*tHMGR*-T2A-*ERG1*-T2A-*ERG9*-$T_{xpr2}$ and $P_{hp4d}$-*BbPK*-$T_{xpr2}$ *and* $P_{hp4d}$-*CkPTA*-$T_{xpr2}$ and $P_{hp4d}$-*IPK*-T2A-*CK*-T2A-*E.coli IDI*-$T_{xpr2}$ and $P_{hp4d}$-*CYP716A155*$^{E120Q}$-$T_{xpr2}$ and $P_{hp4d}$-*LUS-KDEL*-$T_{xpr2}$+$P_{hp4d}$-*ERG20*-$T_{xpr2}$ and $P_{hp4d}$-*Ino2*-$T_{xpr2}$ and $P_{hp4d}$-*Pex30*-T2A-*Rnt1*-T2A-*Ypo1*-$T_{xpr2}$ and $P_{hp4d}$-*GPD1*-$T_{xpr2}$+ $P_{hp4d}$-*MCE*$_2$-$T_{xpr2}$ and integration site *F16*,$P_{hp4d}$-*POT1*-$T_{xpr2}$+$P_{hp4d}$-*PAT1*-$T_{xpr2}$ and integration site *D9*,$P_{hp4d}$-*MFE1*-$T_{xpr2}$+$P_{hp4d}$-*PXA1*-$T_{xpr2}$+$P_{hp4d}$-*FAT1*-$T_{xpr2}$ and integration site *E14,*$P_{hp4d}$-*TGL4*-$T_{xpr2}$+$P_{hp4d}$-*TGL3*-$T_{xpr2}$, and integration site *D9*,$P_{hp4d}$-*DGA1*-$T_{xpr2}$ | This study |
| Y L 7166-79 | Based on YL7166-78, $P_{ERG7}$::$P_{CTR3}$ | This study |
| Y L 7166-80 | Based on YL7166-79,△*TF* | This study |
| Y L 7166-81 | Based on YL7166-80, $P_{PFK}$::$P_{PKI}$ | This study |
| Y L 7166-82 | Based on YL7166-81, $P_{HXK1}$::$P_{TDH1}$ | This study |
| Y L 7166-83 | Based on YL7166-82,△*Mhy1* | This study |
| Y L 7166-84 | Based on YL7166-83, integration site *XPR2*, $P_{hp4d}$-*Vhb*-$T_{xpr2}$ | This study |
| **Plasmid** | | |
| pINA1312 | Multi-copy integration vector with hp4d promoter, *XPR2* terminator, selection marker *URA3*, *KmR* | [13] |
| pUAxp7166 | Mono-copy integration vector with hp4d promoter, *XPR2* terminator, Carrying selection marker *URA3, AmpR*<br>Integration site AXP, Cre/lox-based for repeated, targeted, and markerless gene integration | Lab stock |
| pURCyl-sg1HR | Specific plasmid information has not yet been published. | Lab stock |

Note: '+' indicates that the plasmid was constructed in vitro using restriction enzyme digestion and ligation with compatible cohesive ends.

**Table. S4 Genes used in this study.**

| Gene | Origin | Database and accession number | Optimization |
|---|---|---|---|
| *RcLUS* | *Ricinus communis* | DQ268869 | Yes |
| *AtATR1* | *Arabidopsis thaliana* | At4g24520 | Yes |
| *CYP716A155* | *Rosmarinus officinalis* | MK592859 | Yes |
| *CYP716A180* | *Betula platyphylla* | AHL46848 | Yes |
| *YlHMGR1* | *Yarrowia lipolytica* | YALI_E04807 | No |
| *ERG1* | *Y. lipolytica* | YALI_E15730 | No |
| *ERG9* | *Y. lipolytica* | YALI_A10076 | No |
| *LmPK* | *Leuconostoc mesenteroides* | AY804190.1 | Yes |
| *BbPK* | *Bifidobacterium bifidum* | LFII01000014.1 | Yes |
| *PTA* | *Bacillus subtilis* | UATI01000005.1 | Yes |
| *ERG10* | *Y. lipolytica* | YALI_B08536 | No |
| *ERG13* | *Y. lipolytica* | YALI_F30481 | No |
| *ERG12* | *Y. lipolytica* | YALI_B16038 | No |
| *ERG8* | *Y. lipolytica* | YALI_E06193 | No |
| *ERG19* | *Y. lipolytica* | YALI_F05632 | No |
| *IDI* | *Y. lipolytica* | YALI_F04015 | No |
| *ERG20* | *Y. lipolytica* | YALI_E05753 | No |
| *AtIPK* | *A. thaliana* | NM_102426.6 | Yes |
| *ScCK* | *Saccharomyces cerevisiae* | NM_001182020.1 | Yes |
| *IDI* | *Escherichia coli* | NC_002695.2 | Yes |
| Ypo1 | *S. cerevisiae* | NM_001184125.1 | Yes |
| Rtn1 | *S. cerevisiae* | NM_001180541.3 | Yes |
| *POT1* | *Y. lipolytica* | YALI_E18568 | No |
| *MFE1* | *Y. lipolytica* | YALI_E15378 | No |
| *PXA1* | *Y. lipolytica* | YALI_A06655 | No |
| *POX2* | *Y. lipolytica* | YALI_F14495 | No |
| *FAA1* | *Y. lipolytica* | YALI_D17864 | No |
| *ANT1* | *Y. lipolytica* | YALI_E03058 | No |
| *DGA1* | *Y. lipolytica* | YALI_E32769 | No |
| *OLE1* | *Y. lipolytica* | YALI_C07638 | No |
| *PAT1* | *Y. lipolytica* | YALI_E11099 | No |
| *TGL4* | *Y. lipolytica* | YALI_F13550 | No |
| *TGL3* | *Y. lipolytica* | YALI_D21511 | No |
| *ylYEF* | *Y. lipolytica* | CR382131.1 | No |
| *MCE2* | *Mucor circinelloides* | DQ975377.1 | Yes |
| *GPD1* | *Clostridium acetobutylicum* | LZYY01000009.1 | Yes |
| *ERG7* | *Y. lipolytica* | YALI_F06787 | No |

| | | | |
|---|---|---|---|
| *TF* | *Y. lipolytica* | YALI_B06928 | No |
| *HXK1* | *Y. lipolytica* | YALI_B29133 | No |
| *PFK* | *Y. lipolytica* | YALI_D20222 | No |
| *Mhy1* | *Y. lipolytica* | YALI_B28150 | No |

**Table. S5 The sequences of sgRNA used in this study.**

| gene name | sgRNA | Source |
|---|---|---|
| *HXK1* | TCGATCTTTATGGAACACCA | This study |
| *PFK1* | CATGCGGTTTCAAATGCACG | This study |
| *TF* | TCTGCTTCTCCAGTACGCCG | This study |
| *ERG7* | TATTTGATGCGATATGAGCG | This study |
| *Mhy*1 | TGTTCGACGGAGAGAGCGGC | This study |

**Table. S6 Summary of BA titer or content improvements via various metabolic engineering strategies**

| Engineering strategy | Fold improvement (Titer or Content) |
|---|---|
| Enhance the flux through the MVA pathway and 2,3-oxidosqualene toward betulinic acid | Titer: Approximately 46-fold improvement |
| Enhance acetyl-CoA precursor availability | Titer: Approximately 3-fold improvement |
| Engineering cytosolic redox metabolism (Increase NADPH availability) | Titer: Approximately a 13% increase |
| Protein engineering of CYP716A155 | Titer: Approximately a 9% increase |
| Subcellular engineering | Titer: Approximately a 13% increase |
| Engineering lipid metabolism | Titer: Approximately a 5% increase |
| Downregulation of the sterol-competing pathway | Content: Approximately a 7% increase |
| Fine-tuning of the glycolytic pathway | Titer: Approximately a 20% increase |

**Note:** The unit of titer is milligrams per liter (mg/L), and the unit of content is milligrams per gram (mg/g).

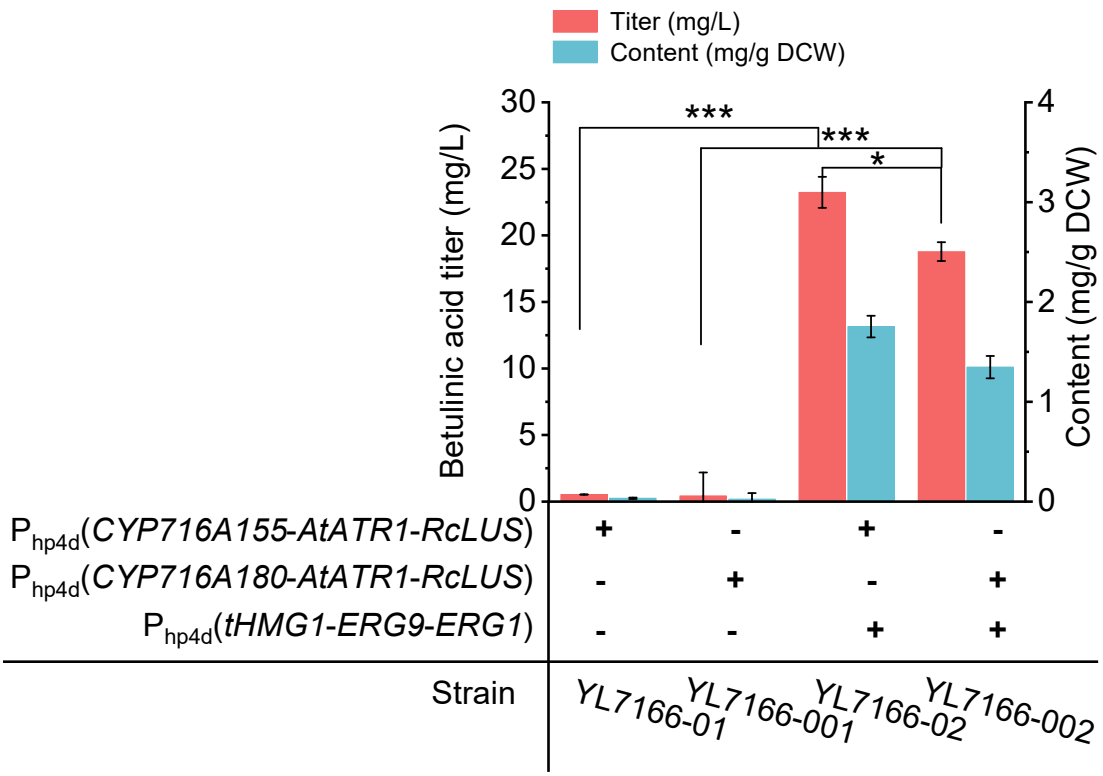



**Fig. S1** BA titer and content in strains screened for cytochrome P450 oxidases (CYP) from different sources.

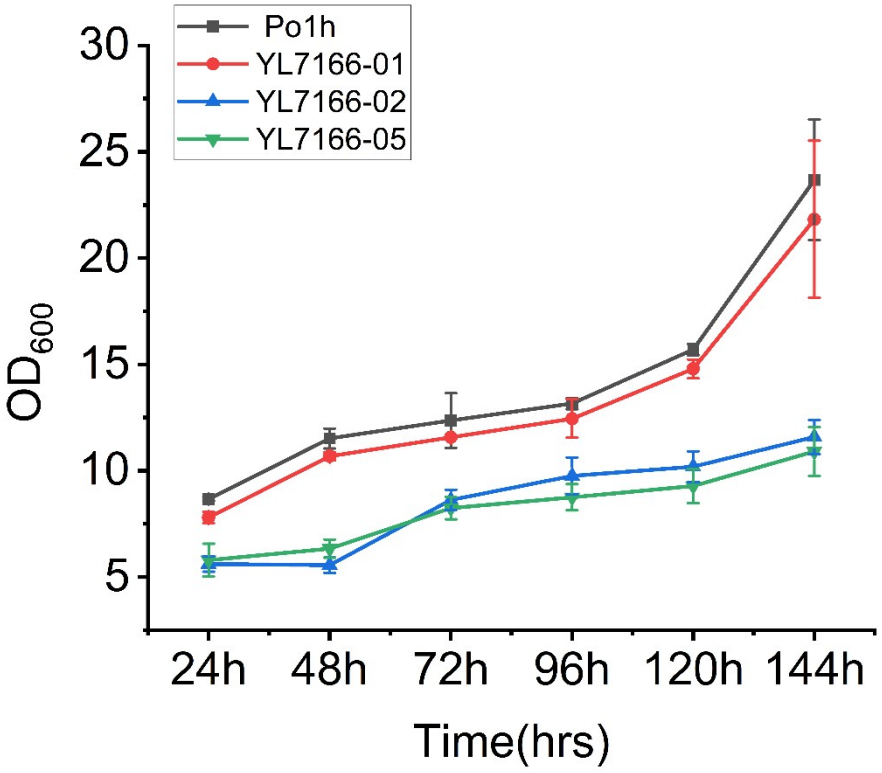



**Fig. S2** Strain Growth Curves. $OD_{600}$ values of Po1h, YL7166-01, YL7166-02, and YL7166-05 measured at 24h, 48h, 72h, 96h, 120h, and 144 h.

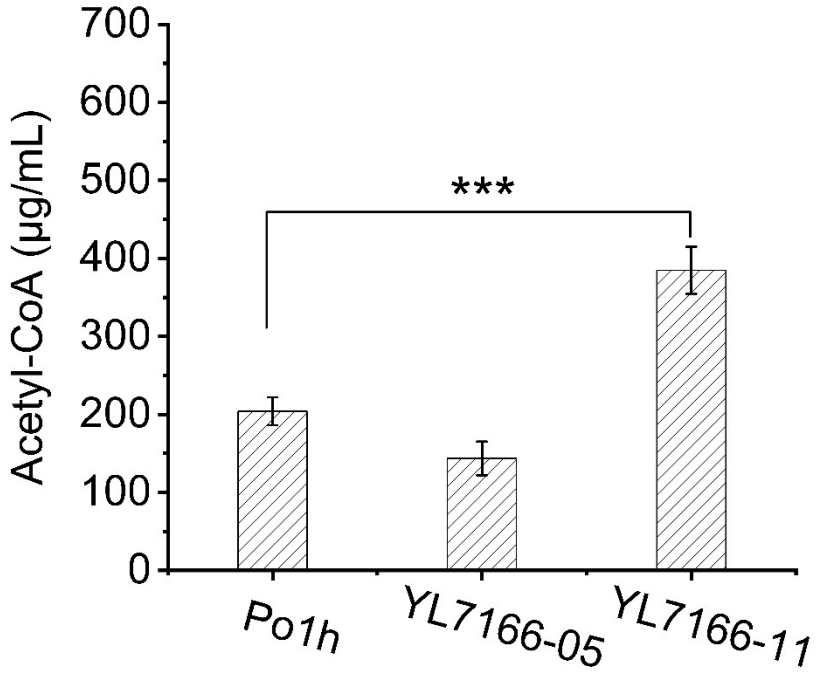



**Fig. S3** Boosting intracellular acetyl-CoA using NOG pathway strategies. Bars represent the mean ±standard deviation (SD), n=3.

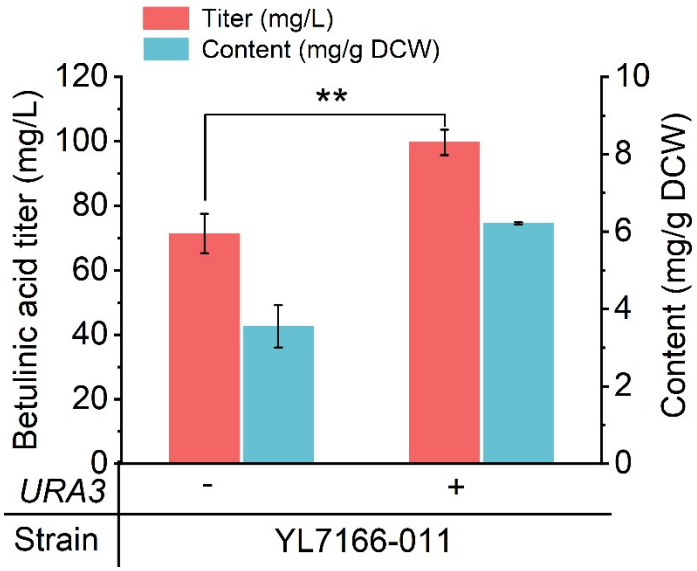



**Fig. S4** Effect of URA3 rescue via the plasmid pINA1312 on BA titer and content in engineered strains. Bars represent the mean ±standard deviation (SD), n=3.

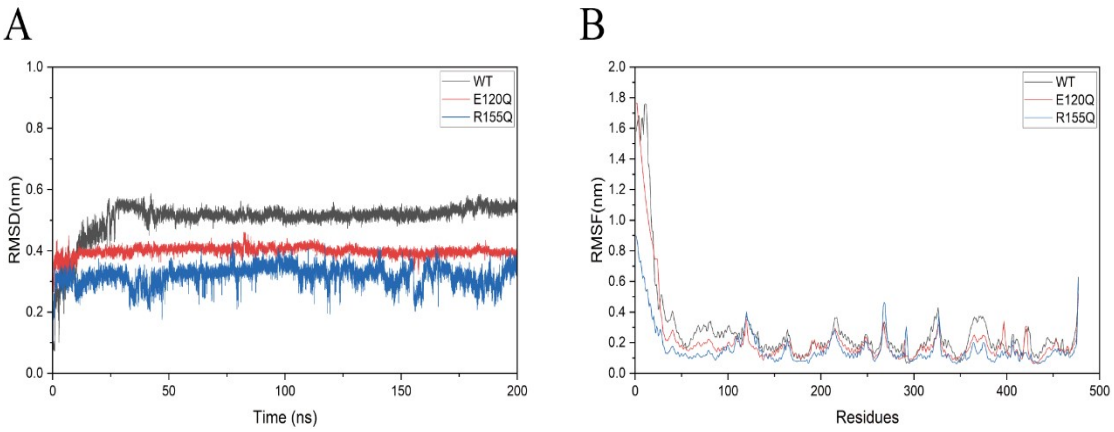



**Fig. S5** (A) Root mean square deviations (RMSDs) of ligands in mutant and wild-type complexes. (B) Root mean square fluctuations (RMSFs) of mutant and wild-type complexes.

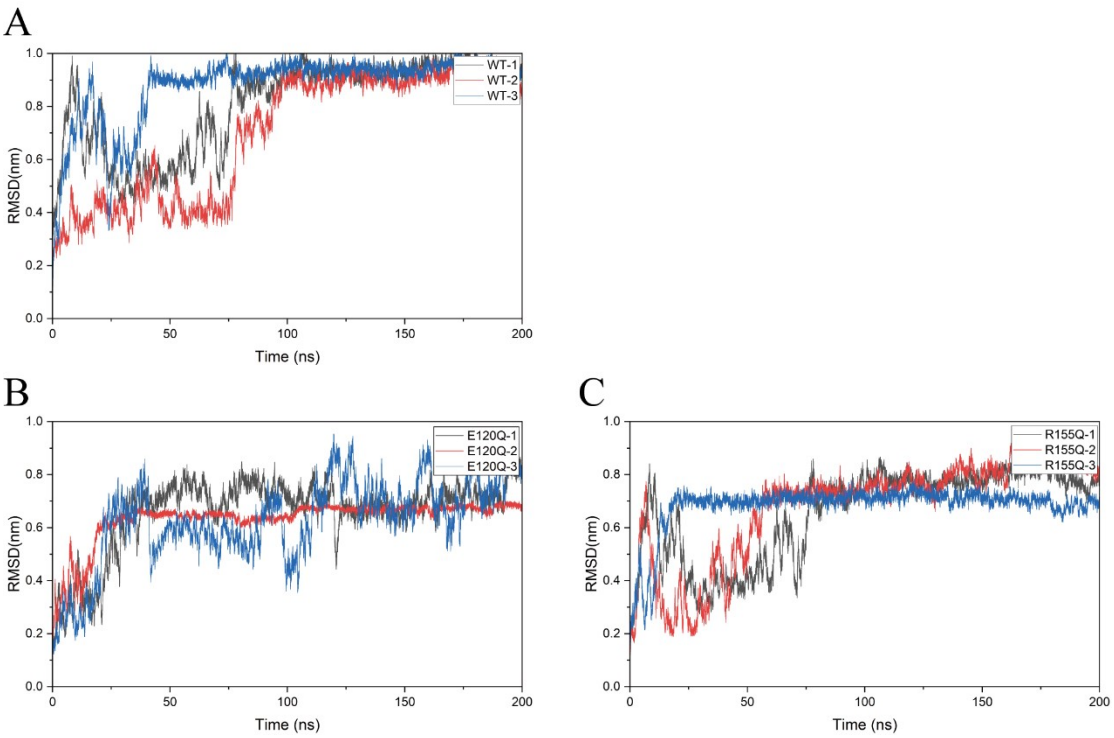



**Fig. S6** The results of three molecular dynamics simulations. (A) WT; (B) E120Q；(C)R155Q.

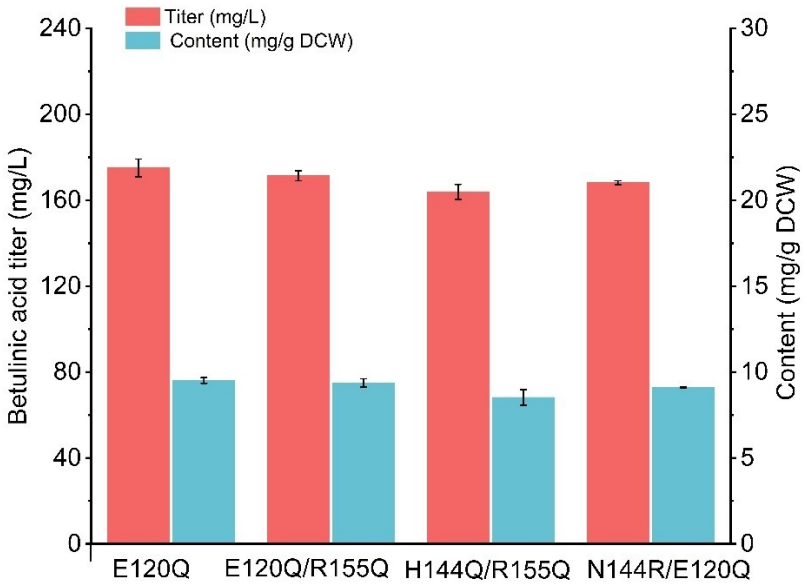



**Fig. S7** BA production of combinatorial double mutants E120Q/R155Q, H144Q/R155Q, and N144R/E120Q compared to the best-performing single mutant E120Q.

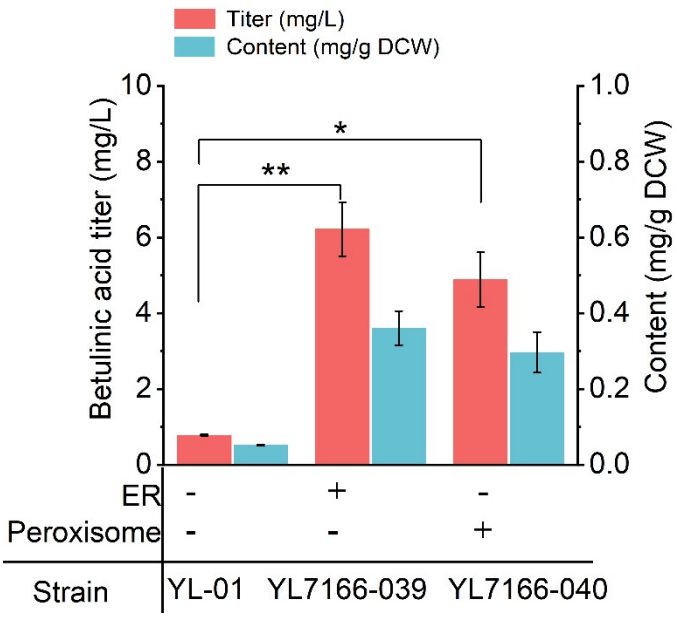
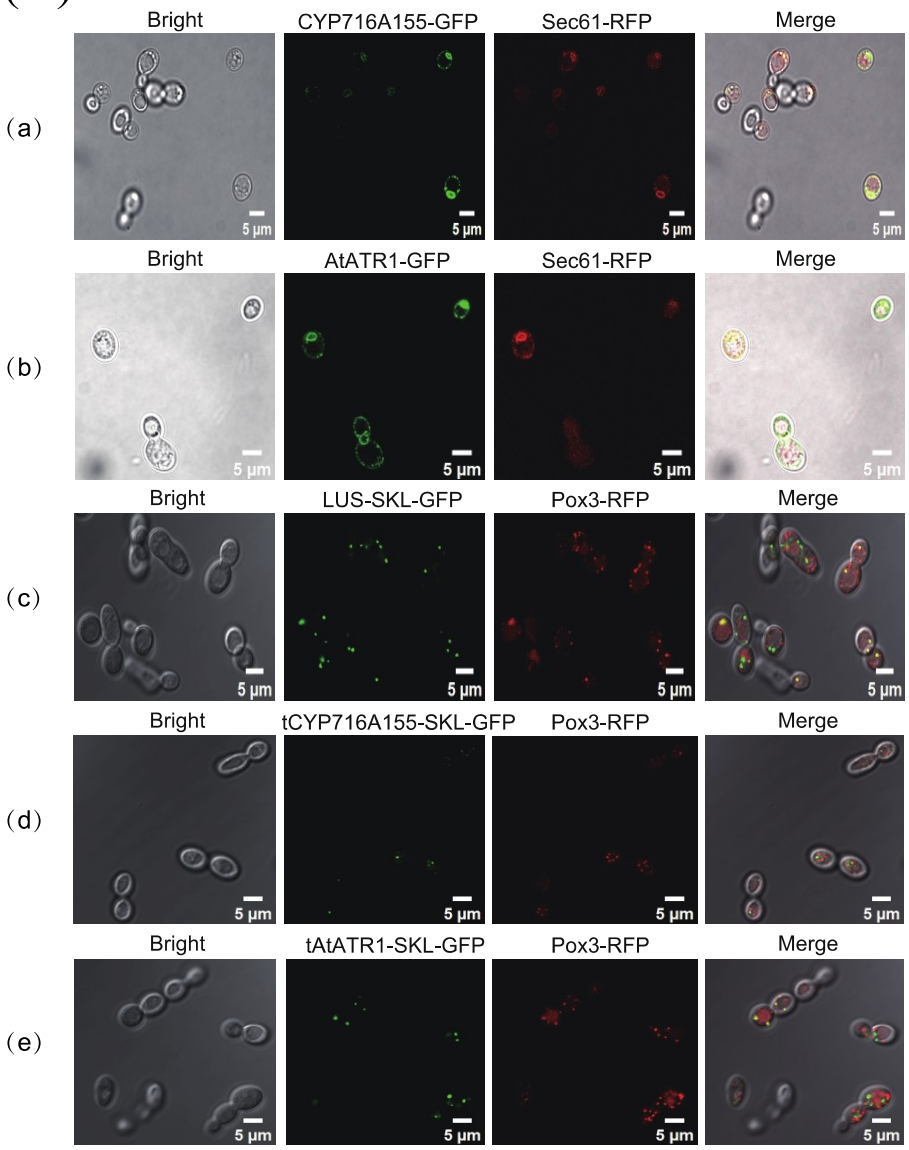
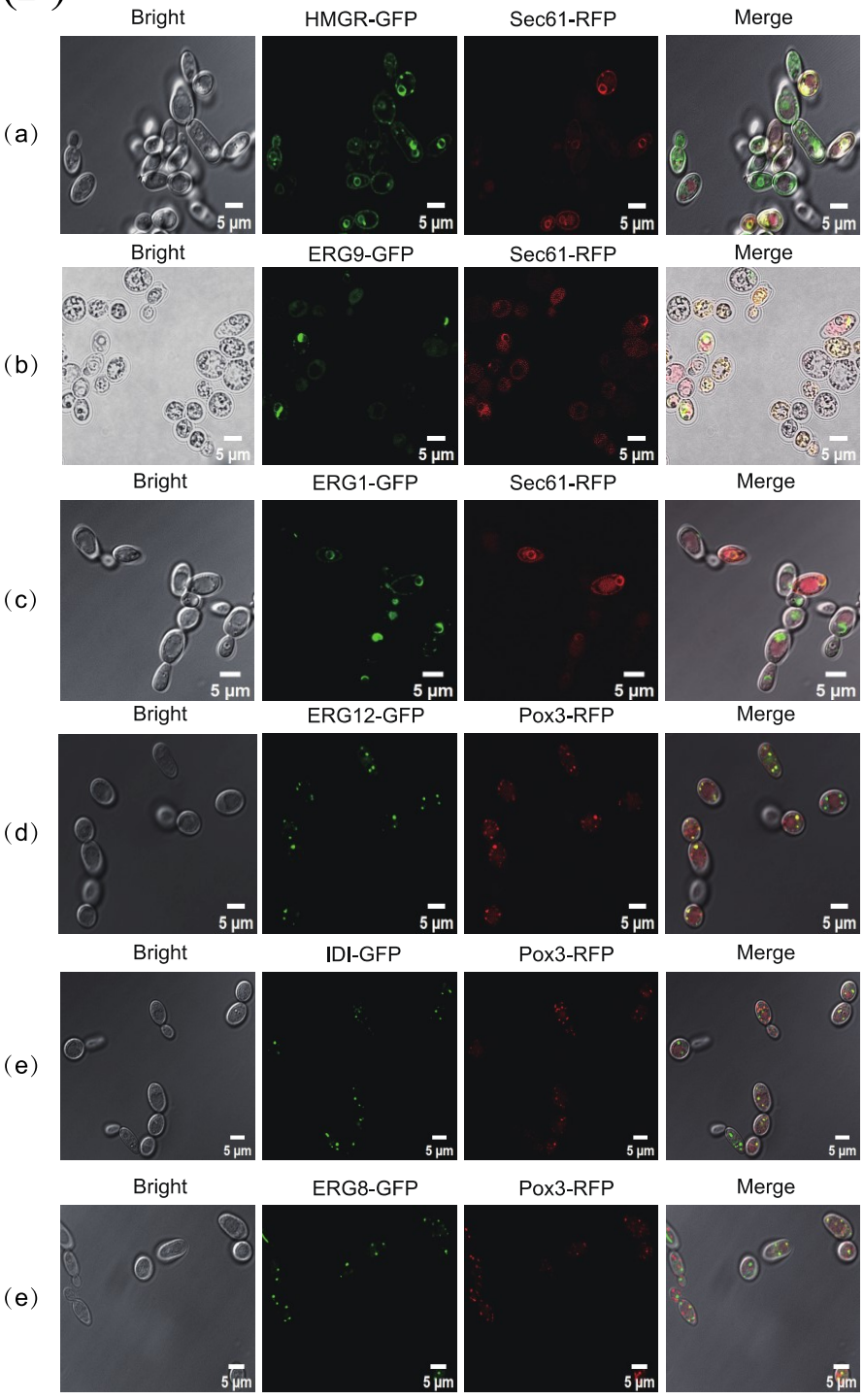



**Fig. S8** Comparison of the BA titer and content between strains with *LUS*, *CYP716A155*, and *ATR1* localized to the endoplasmic reticulum and peroxisomes versus unlocalized strains.

(A)


Bright
CYP716A155-GFP
Sec61-RFP
Merge
(a)
Bright
AtATR1-GFP
Sec61-RFP
Merge
(b)
Bright
LUS-SKL-GFP
Pox3-RFP
Merge
(c)
Bright
tCYP716A155-SKL-GFP
Pox3-RFP
Merge
(d)
Bright
tAtATR1-SKL-GFP
Pox3-RFP
Merge
(e)
5 μm

(B)


Bright
HMGR-GFP
Sec61-RFP
Merge
(a)
Bright
ERG9-GFP
Sec61-RFP
Merge
(b)
Bright
ERG1-GFP
Sec61-RFP
Merge
(c)
Bright
ERG12-GFP
Pox3-RFP
Merge
(d)
Bright
IDI-GFP
Pox3-RFP
Merge
(e)
Bright
ERG8-GFP
Pox3-RFP
Merge
(e)
5 μm

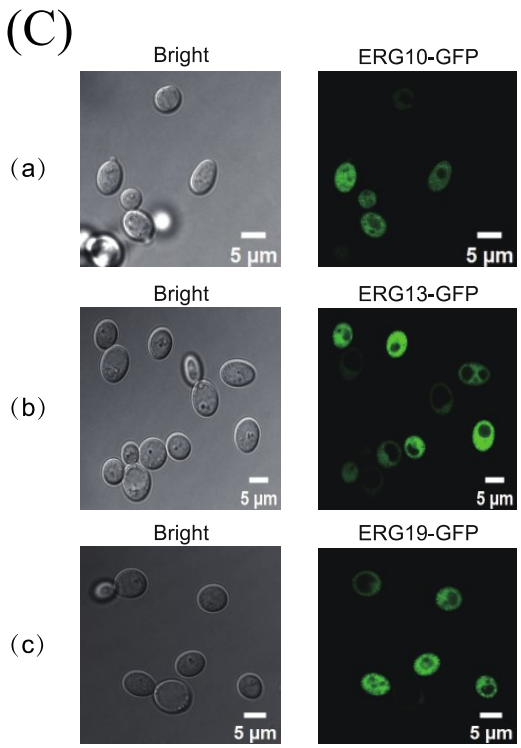



**Fig. S9** (A)and(B) Engineering *LUS*, t*CYP716A155* and t*AtATR1*to re-localize it to ER and peroxisome and subcellular localization of the enzymes *HMGR*, *ERG9*, *ERG1*, *ERG12*, *IDI*, and *ERG8* by fusion with GFP protein respectively. (C) Fluorescence microscopy images showing cytosolic localization of *ERG10*, *ERG13*, and *ERG19*, each fused to GFP.

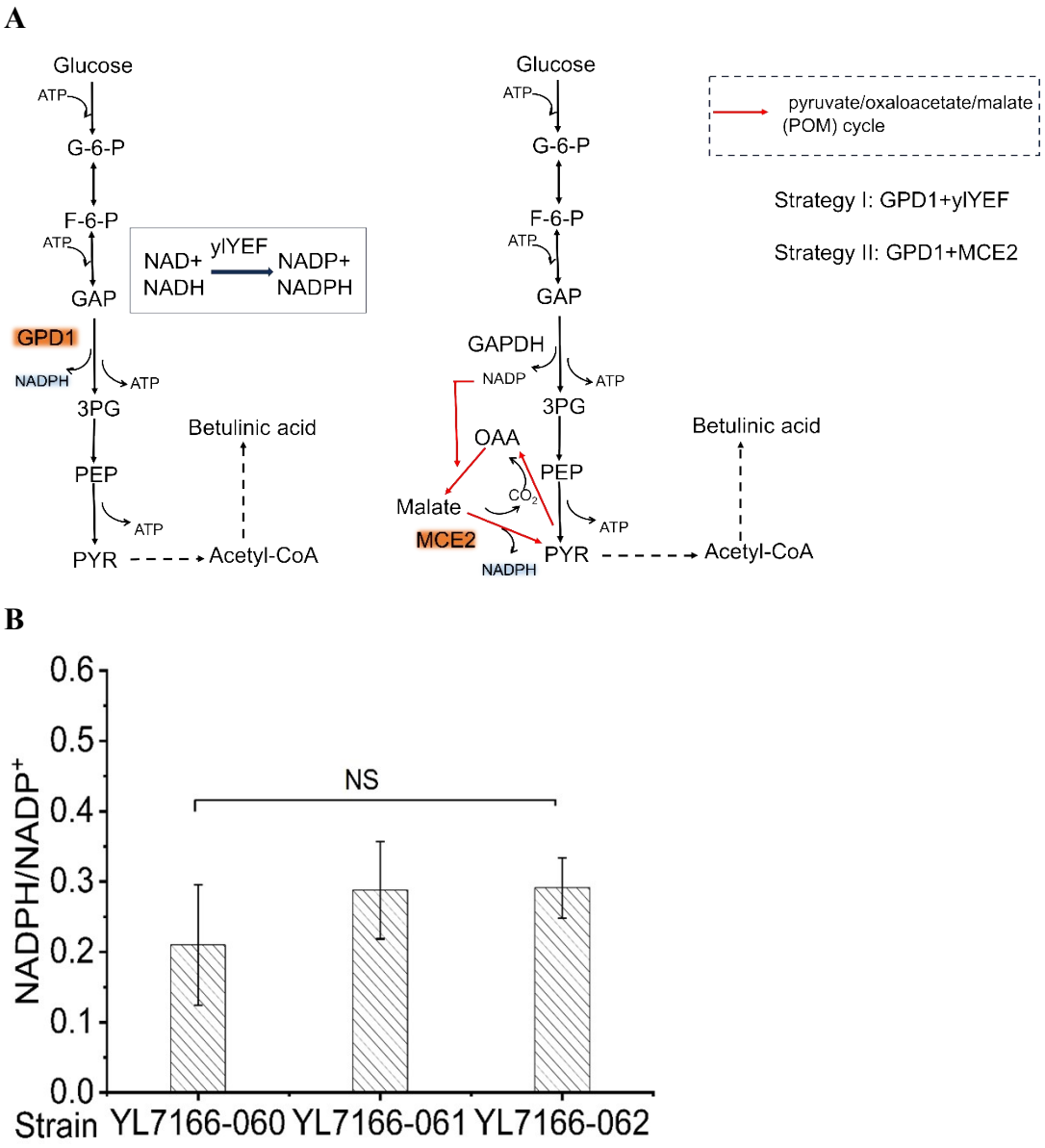
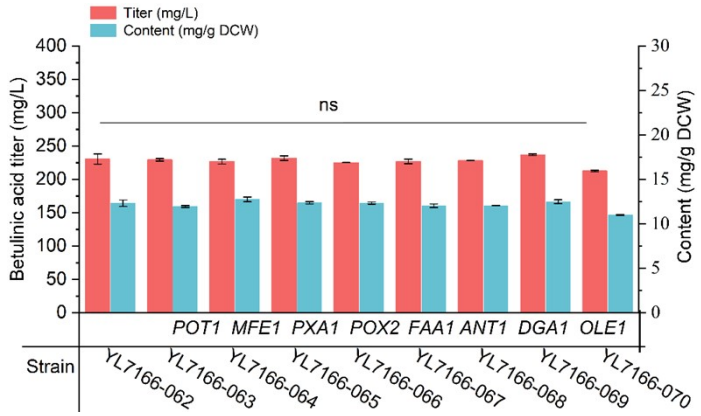
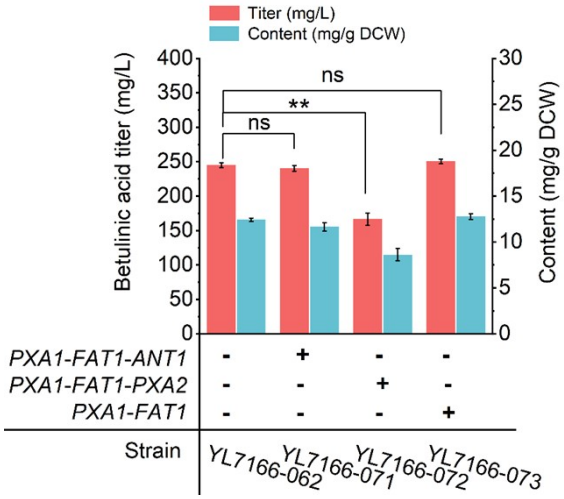



**Fig. S10** (A) Schematic diagram of two NADPH formation strategies: GPD1, glyceraldhyde-3-phosphate dehydrogenase from *Clostridium acetobutylicum*, MCE2, a cytosolic NADP+-dependent malic enzyme from *Mucor circinelloides*, ylYEF, a cytosolic $NAD^+$/NADH kinase from *Y. lipolytica*, GAPDH, glyceraldhyde-3-phosphate dehydrogenase from *Y. lipolytica*. and (B) The NADPH/NADP+ redox ratios were measured for strains YL7166-060，YL7166-061 and YL7166-062. Bars represent the mean ±standard deviation (SD), n=3.

**A**




**B**




**C**

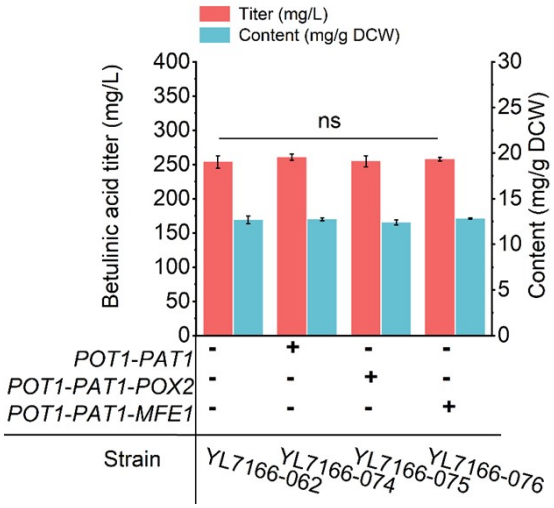



**Fig. S11** (A) (B) (C) The effect of overexpressing fatty acid metabolism-related genes on BA titer and content in engineered strains based on strain YL66-062.

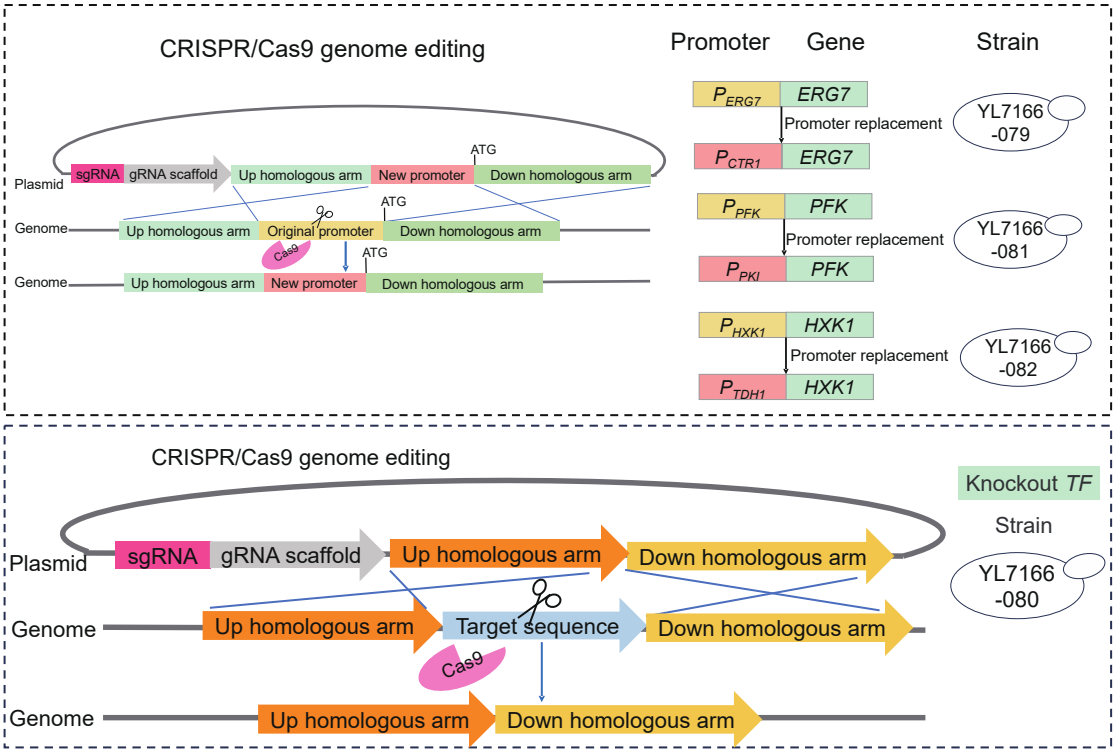



**Fig. S12** Diagram of the pURCyl-sg1HR plasmid designed for gene knockout and promoter replacement via CRISPR/Cas9.

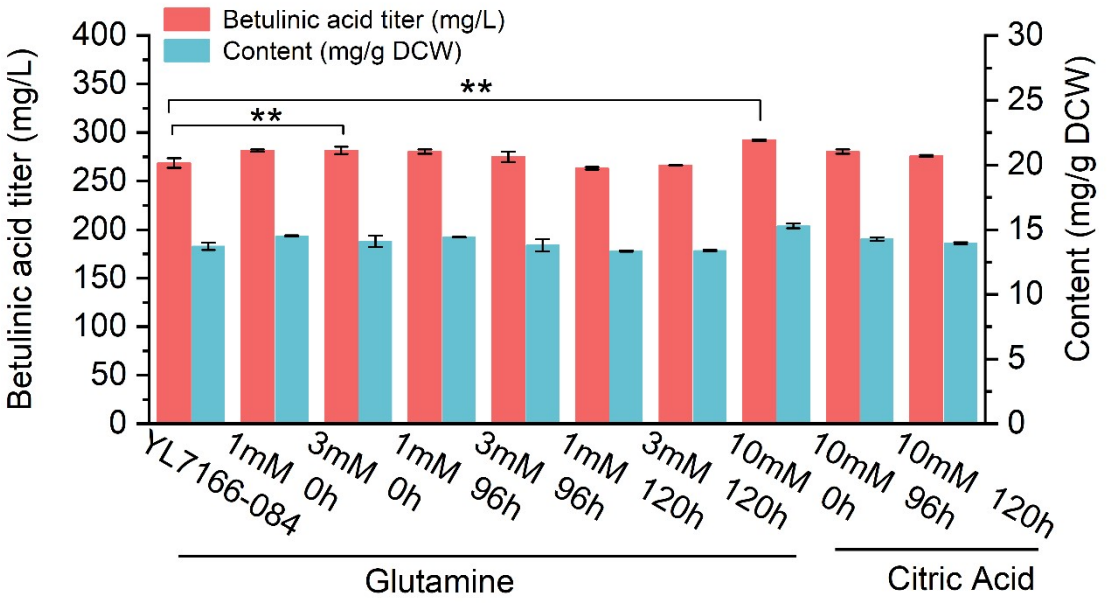
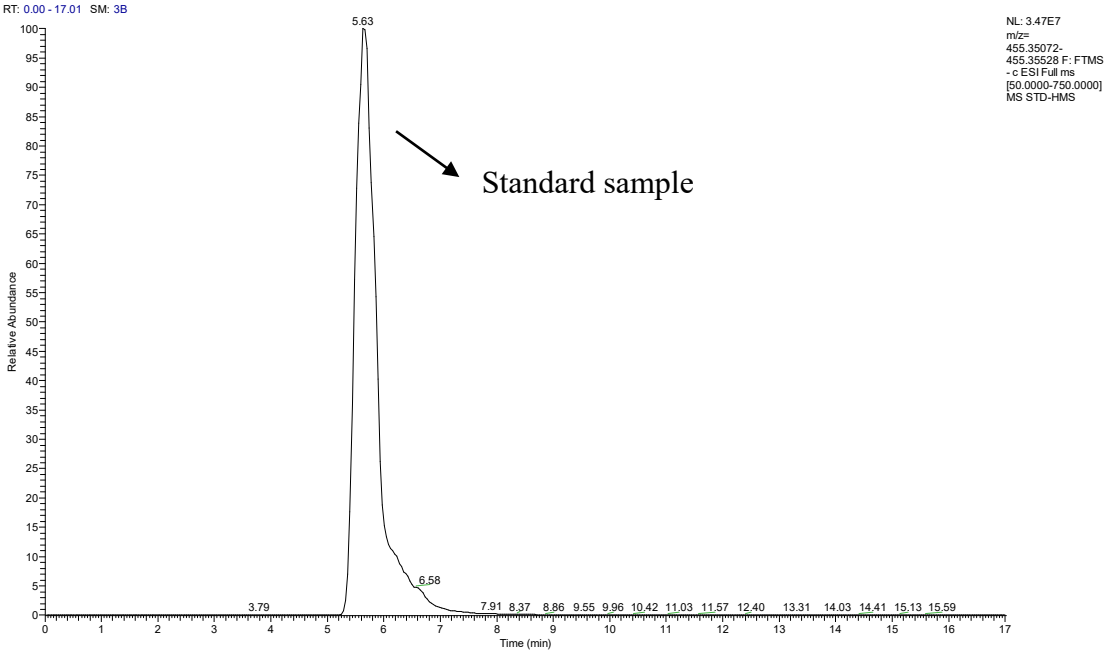
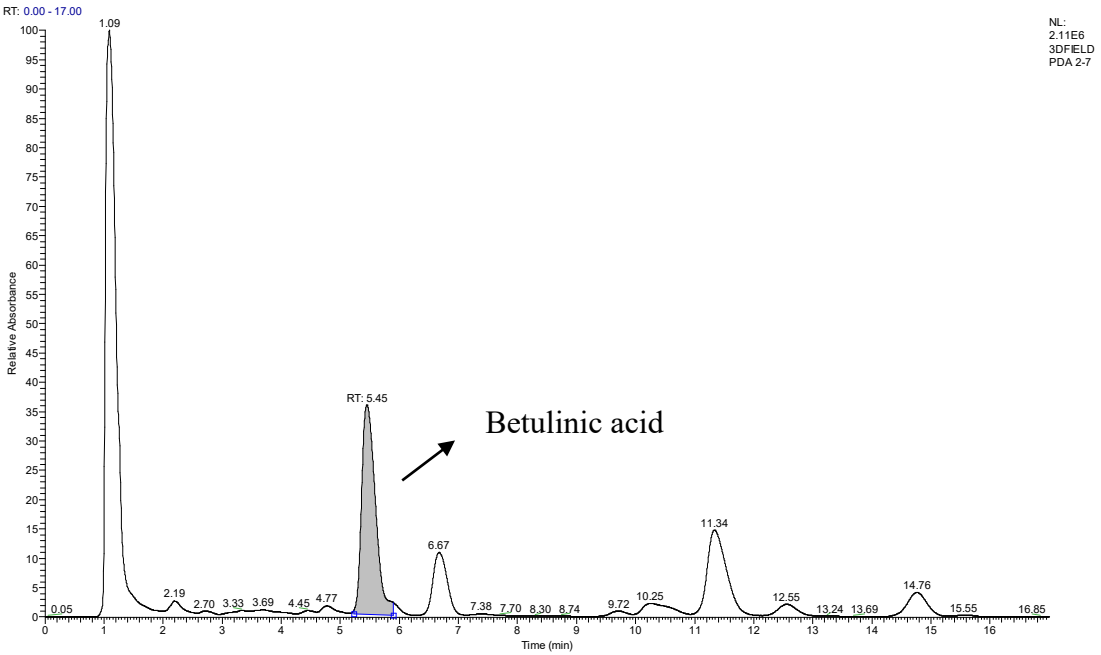



**Fig. S13** The effect of supplementing different concentrations of glutamine and citrate at various fermentation time points on BA titer and content in shake flask cultures.

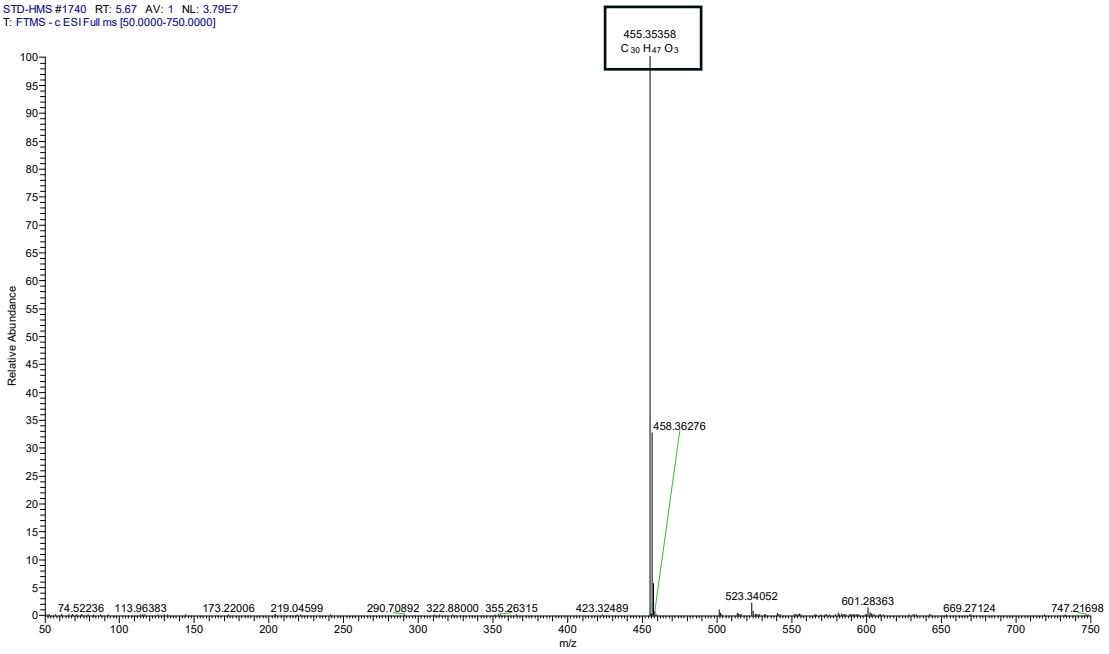
**A(a)**


RT: 0.00 - 17.01 SM: 3B
NL: 3.47E7
m/z=
455.35072-
455.35528 F: FTMS
- c ESI Full ms
[50.0000-750.0000]
MS STD-HMS
Standard sample
Relative Abundance
Time (min)
5.63
6.58
3.79
7.91
8.37
8.86
9.55
9.96
10.42
11.03
11.57
12.40
13.31
14.03
14.41
15.13
15.59

**A(b)**

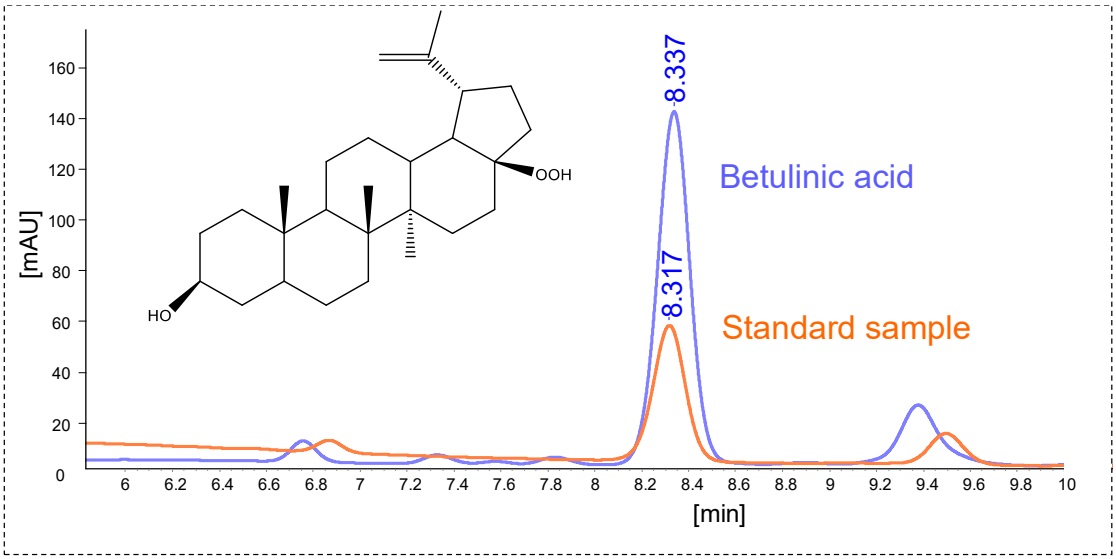

RT: 0.00 - 17.00
NL:
2.11E6
3DFIELD
PDA 2-7
Betulinic acid
Relative Absorbance
Time (min)
RT: 5.45
1.09
0.05
2.19
2.70
3.33
3.69
4.45
4.77
6.67
7.38
7.70
8.30
8.74
9.72
10.25
11.34
12.55
13.24
13.69
14.76
15.55
16.85

**B**


STD-HMS #1740 RT: 5.67 AV: 1 NL: 3.79E7
T: FTMS - c ESI Full ms [50.0000-750.0000]
455.35358
C 30 H47 O3
458.36276
Relative Abundance
m/z

**Mass(m/z)**

**C**


8.337
8.317
Betulinic acid
Standard sample
OOH
HO
[mAU]
[min]

**D**

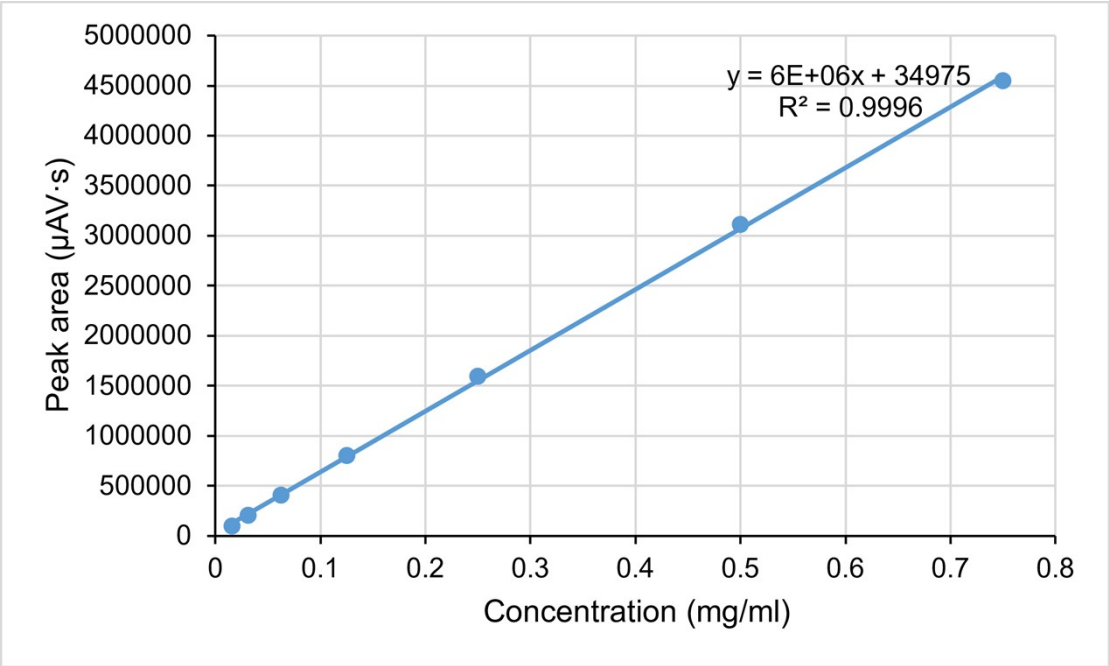



**Fig. S14 Qualitative and quantitative analysis of betulinic acid (BA) produced from fermentation.**

**A.** LC-MS-based qualitative identification of betulinic acid: (a) extracted ion chromatogram of the fermentation product; (b) chromatogram of a commercial BA standard.

**B.** Mass spectrum of the BA peak from the fermentation product, showing a dominant ion at *m/z* 455.35, consistent with the protonated molecular ion of BA.

**C.** HPLC-based quantification of BA: blue trace, chromatogram of the fermentation sample; orange trace, chromatogram of the commercial BA standard.

**D.** Calibration curve used for BA quantification.

## Codon-optimized sequences

Note: The yellow highlight represents the organelle-targeting sequence, and the green font indicates the T2A peptide sequence.

### 1. *CYP716A155*

ATGGAGTTCTTCTACGCCTCCCTGCTGTGTCTGTTCGTGTCTCTGGTCTTCCTGTCCCT
GCACCTGCTGTTCTACAAGACCAAGACCGGCTCTCTGCCTCCCGGCAAGACCGGTTGG
CCCGTGATTGGAGAGTCTCTGGAGTTCCTGTCCACCGGATGGAAGGGTCACCCCGAG
AAGTTCATCTTCGACCGAATGGCCCGATACTCTTCCCACGTCTTCCGAACCCACCTGC
TGGGAGAGCCCGCCGCTGTCCTGTGTGGCTCCGCCGGAAACAAGTTCCTGTTCTCTAA
CGAGAACAAGCTGGTGCAGGCTTGGTGGCCCTCTTCCGTCGAGAAGATTTTCCCCAAC
GACAACGCCGAGACCTCTTCCAAGGAGGAGTCCATCAAGATGCGACGAATGCTGCCC
ACCTTCTTCAAGCCCGAGGCCCTGCACCGATACGTCGGCATTATGGACCACATCGCCC
GACGACACTTCGCTGACGGTTGGGACGGCAAGCGAGAGGTGGTCGTGTTCCCCCTGG
CCAAGAACTACACCTTCTGGCTGGCCTGCCGACTGTTCCTGTCTGTGGAGGACCCCTC
CCAGGTCGAGAAGTTCGCCGCTCCCTTCAACCTGCTGGCCTCTGGACTGATCTCCATT
CCTATCGACCTGCCTGGCACCCCTTTCCACAAGGGCATTAAGGCTTCTGCCTACATCC
GAAAGGAGCTGGTGGCCATCATTAAGCAGCGAAAGGCTGACCTGGCTGACGGCACCG
CTTCCCCTACCCAGGACATTCTGTCCCACATGCTGCTGACCTCTAACGAGGACGGCAA
GTTCATGCAGGAGTCTGACATTGCCAACAAGATCCTGGGACTGCTGATCGGCGGACA
CGACACCGCTTCTTCCGCCTGTACCTTCGTCGTGAAGTACCTGGCTGAGCTGCCCCAG
GTGTACGAGGGCGTCTACAAGGAGCAGATGGAGATTGCCAAGTCCAAGGCCGCTGGA
GAGCTGCTGAACTGGGAGGACCTCCAGAAGATGAAGTACTCTTGGACGTGGCTTGTG
AGGTCCTGCGACTGGCCCCTCCCCTCCAGGGAGCTTTCCGAGAGGCTCTGGCCGACTT
CTCTTTCAACGGATTCTCCATCCCCAAGGGTTGGAAGCTGTACTGGTCTGCCAACTCC
ACCCACAAGAACTCCGAGTTCTTCCCCGAGCCCGAGAAGTTCGACCCTTCTCGATTCG
AGGGCTCCGGACCCGCTCCCTACACCTTCGTGCCTTTCGGTGGCGGACCCCGAATGTG
TCCCGGAAAGGAGTACGCCCGACTGGAGATTCTGGTGTTCATGCACCACCTGGTCAA
GCGATTCAAGTGGGAGAAGATGATTCCCGACGAGAAGATCGTCGTGGACCCCATGCC
TATCCCTGCTAACGGTCTGCCCGTCCGACTGTACCCCCACACCTCCTAG

### 2. *CYP716A180*

ATGGAGCACTTCTACCTGTCCCTGCTGCTGCTGTTCGTCTCTTTCGTGACCCTGTCCCT
GTTCACCCTGTTCTACAAGCACCGATCTCACTTCACCGGACCCAACCTGCCTCCCGGC
AAGACCGGCTACCCCATGATCGGAGAGTCTCTGGAGTTCCTGTCCACCGGATGGAAG
GGTCACCCCGAGAAGTTCATTTTCGACCGAATGACCAAGTTCTCTTCCGAGGTGTTCA
AGACCTCTCTGCTGGGTCAGCCCGCCGCTGTCTTCTGTGGTGCTGCTTGCAACAAGTT
CCTGTTCTCCAACGAGAACAAGCTGGTGACCGCCTGGTGGCCCGACTCTGTCAACAA
GATTTTCCCCTCTTCCACCCAGACCTCTAACTCCAAGGAGGAGGCTAAGAAGATCGCG
AAAGCTGCTGCCCCAGTTCCTGAAGCCCGAGGCCCTCCAGAAGTACATCTCCATTATG
GACACCATTGCTCAGCGACACTTCGCCTCTGGCTGGGAGGGACAGAAGGAGGTCACC

GTGTTCCCCCTGGCTAAGCGATACACCTTCTGGCTGGCCTGCCGACTGTTCCTGTCTCT
GGAGGACCCCAACCACATCGCCCGATTCGCTGACCCCTTCCACTCTGTCGCTTCCGGA
ATCATTTCTATCCCCATTGACCTGCCCGGCACCCCTTCAACCGAGGAATCAAGGCCTC
CAACTTCATTCGAAAGGAGCTGTCTCTGATCATTAAGCAGCGACGAGTGGACCTGGG
TGAAGGCAAGGCTTCCCCCACCCAGGACATCCTGTCTCACATGCTGCTGACCTCTGAC
GAGTCCGGCCAGTACATGACCGAGCTGGACATCGCCGACAAGATTCTGGGACTGCTG
ATTGGCGGACACGACACCGCTTCCGCCGCTTGTACCTTCATCGTGAAGTACCTGGCCG
AGCTGCCCCACATCTACGAGGGCGTCTACAACGAGCAGATGGAGATCGCTAACTCCA
AGGCCCCGGAGAGCTGCTGAACTGGGAGGACATTCAGAAGATGCGATACTCTTGGAC
GTGGCTTGTGAGGTGCTGCGACTGGCCCCTCCCCTCCAGGGTGCTTTCCGAGAGGCCA
TCAACGACTTCATTTTTCAACGGTTTCTCTATCCCCAAGGGCTGGAAGCTGTACTGGT
CTGCTAACTCCACCCACCGATCCGCCGAGTACTTCCCCGAGCCCGAGAAGTTCGACCC
CTCTCGATTCGAGGGACGAGGTCCCGCTCCCTACACCTTCGTGCCTTTCGGTGGCGGA
CCCCGAATGTGTCCCGGAAAGGAGTACGCCCGACTGGAGATTCTGGTCTTCATGCAC
AACCTGGTGAAGCGATTCCGATGGGAGAAGATGATCCCCGACGAGAAGATTGTGGTG
GACCCCATGCCTATGCCTGCTAAGGGTCTGCCCGTCCGACTGTACCCCCACAAGGCT

### 3. *AtATR1*

ATGACCTCTGCTCTGTACGCCTCCGACCTGTTCAAGCAGCTGAAGTCCATTATGGGCA
CCGACTCTCTGTCCGACGACGTCGTGCTGGTCATCGCTACCACCTCTCTCGCCCTGGT
GGCTGGTTTCGTCGTGCTGCTGTGGAAGAAGACCACCGCTGACCGATCCGGAGAGCT
GAAGCCTCTGATGATTCCCAAGTCTCTGATGGCCAAGGACGAGGACGACGACCTGGA
CCTGGGCTCTGGCAAGACCCGAGTGTCTATCTTCTTCGGCACCCAGACCGGCACCGCC
GAGGGATTCGCTAAGGCCCTGTCTGAGGAGATTAAGGCTCGATACGAGAAGGCCGCT
GTCAAGGTCATCGACCTGGACGACTACGCCGCTGACGACGACCAGTACGAGGAGAAG
CTGAAGAAGGAGACCCTGGCCTTCTTCTGTGTCGCTACCTACGGCGACGGAGAGCCT
ACCGACAACGCTGCTCGATTCTACAAGTGGTTCACCGAGGAGAACGAGCGAGACATC
AAGCTCCAGCAGCTGGCCTACGGAGTGTTCGCTCTGGGTAACCGACAGTACGAGCAC
TTCAACAAGATTGGCATCGTCCTGGACGAGGAGCTGTGCAAGAAGGGAGCTAAGCGA
CTGATTGAAGTGGGTCTGGGCGACGACGACCAGTCCATCGAGGATGACTTCAACGCC
TGGAAGGAGTCTCTGTGGTCCGAGCTGGACAAGCTGCTGAAGGACGAGGACGACAAG
TCTGTGGCCACCCCCTACACCGCTGTCATTCCCGAGTACCGAGTCGTGACCCACGACC
CCCGATTCACCACCCAGAAGTCTATGGAGTCCAACGTGGCTAACGGAAACACCACCA
TTGACATCCACCACCCCTGTCGAGTGGACGTGGCCGTCCAGAAGGAGCTGCACACCC
ACGAGTCTGACCGATCCTGCATCCACCTGGAGTTCGACATTTCTCGAACCGGCATCAC
CTACGAGACCGGTGACCACGTCGGCGTCTACGCCGAGAACCACGTCGAGATTGTCGA
GGAGGCTGGAAAGCTGCTGGGTCACTCCCTGGACCTGGTCTTCTCTATCCACGCCGAC
AAGGAGGACGGCTCTCCCCTGGAGTCTGCCGTGCCCCCTCCCTTCCCCGGACCCTGTA
CCCTGGGAACCGGTCTGGCTCGATACGCCGACCTGCTGAACCCTCCCCGAAAGTCCGC
TCTGGTGGCCCTGGCCGCTTACGCCACCGAGCCTTCTGAGGCTGAGAAGCTGAAGCA
CCTGACCTCCCCCGACGGAAAGGACGAGTACTCTCAGTGGATCGTGGCCTCTCAGCG
ATCCCTGCTGGAGGTCATGGCTGCTTTCCCCTCTGCTAAGCCTCCCCTGGGCGTCTTCT
TCGCTGCTATTGCTCCCCGACTCCAGCCCCGATACTACTCTATTTCCTCTTCCTCCCGA

CTGGCCCCTTCCCGAGTGCACGTCACCTCTGCTCTGGTGTACGGACCCACCCCCACCG
GACGAATCCACAAGGGTGTGTGTTCCACCTGGATGAAGAACGCTGTCCCCGCCGAGA
AGTCTCACGAGTGCTCCGGAGCCCCCATTTTCATCCGAGCTTCTAACTTCAAGCTGCC
CTCTAACCCCTCCACCCCTATCGTGATGGTGGGACCCGGAACCGGTCTGGCCCCCTTC
CGAGGCTTCCTCCAGGAGCGAATGGCTCTGAAGGAGGACGGCGAGGAGCTGGGCTCT
TCTCTGCTGTTCTTCGGCTGCCGAAACCGACAGATGGACTTCATCTACGAGGACGAGC
TGAACAACTTCGTGGACCAGGGTGTCATTTCCGAGCTGATCATGGCCTTCTCTCGAGA
GGGCGCTCAGAAGGAGTACGTGCAGCACAAGATGATGGAGAAGGCCGCTCAGGTCT
GGGACCTGATCAAGGAGGAGGGCTACCTGTACGTGTGTGGTGACGCTAAGGGCATGG
CCCGAGACGTCCACCGAACCCTGCACACCATTGTCCAGGAGCAGGAGGGAGTCTCCT
CTTCCGAGGCTGAGGCCATCGTGAAGAAGCTCCAGACCGA
GGTCGATACCTGCGAGACGTCTGGTGA

**4. *RcLUS***

ATGTGGCGAATTAAGATCGCCGAGGGTGGCAACAACCCCTACATCTACTCTACCAAC
AACTTCCAGGGCCGACAGATTTGGGTGTTCGACCCTAACGCTGGCACCCCTGAGGAG
CAGGCCGAGGTCGAGGAGGCTCGACAGAACTTCTGGAAGAACCGATTCCAGGTGAAG
CCCAACTCCGACCTGCTGTGGCAGCTCCAGTTCCTGCGAGAGAAGAACTTCAAGCAG
AAGATTCCCAAGGTGAAGGTCGAGGACGGAGAGGAGATTACCTCTGAGATCGCCGCT
GCCGCTCTGCGACGATCTGTGCACCTGTTCTCTGCCCTCCAGGCCTCTGACGGACACT
GGTGTGCTGAGGAACGGAGGTCTGCTGTTCTTCCTGCCCTCCCCTGGTGTTCGCCGTC
TACATCACCGGTCACCTGAACACCGTCTTCTCCCCCGAGCACCGAAAGGAGATTCTGC
GATACATCTACTGCCACCAGAACGAGGACGGCGGATGGGGCATTCACATCGAGGGAC
ACTCTACCATGTTCTGTACCGTGCTGAACTACATTTGCATGCGAATCCTGGGTGAAGC
CCGAGACGGTGGCATTGAGAACGCTTGTGAGCGAGGCCGAAAGTGGATTCTGGACCA
CGGAGGTGCCACCGGTATCTCTTCCTGGGGCAAGACCTGGCTGTCTATCCTGGGCGTG
TACGAGTGGGACGGAACCAACCCCATGCCTCCCGAGTTCTGGGCCTTCCCCTCTTCCT
TCCCCCTGCACCCCGCTAAGATGTTCTGTTACTGCCGAATTACCTACATGCCCATGTCC
TACCTGTACGGAAAGCGATTCGTCGGTCCCATCACCCCTCTGATTCTCCAGATCCGAG
AGGAAATCTACAACGAGCCCTACAACAAGATCAAGTGGAACTCTGTGCGACACCTGT
GTGCTAAGGAGGACAACTACTTCCCCACCCCACCATTCAGAAGCTGCTGTGGGACGC
CCTGTACACCTTCTCCGAGCCCCTGTTCTCTCGATGGCCCTTCAACAAGCTGCGAGAG
AAGGCCCTGAAGATTACCATGGACCACATCCACTACGAGGACCACAACTCCCGATAC
ATTACCATCGGCTGCGTCGAGAAGCCCCTGTGTATGCTGGCCTGCTGGATTGAGGACC
CCCACGGAGAGGCTTTCAAGAAGCACCTGGCTCGAATTGCCGACTACATCTGGGTGG
GAGAGGACGGTATCAAGATGCAGTCTTTCGGCTCTCAGACCTGGGACACCTCTCTGGC
TCTCCAGGCCCTGATTGCTTCTGACCTGTCCCACGAGATCGGTCCCACCTGAAGCAGG
GCCACGTCTTCAACCAAGAACCCCAGGCCACCGAGACCCCTCTGGCGACTTCCGAAA
GATGTTCCGACACATCTCCAAGGGAGCTTGGACCTTCTCTGACAAGGACCAGGGATG
GCAGGTGTCCGACTGCACCGCCGAGTCTCTGAAGTGTTGCCTGCTGTTCTCCATGATG
CCTCCCGAGATTGTGGGAGAGAAGATGGAGCCCGAGAAGGTCTACGACTCTGTGAAC
GTCATCCTGTCTCTCCAGTCCCAGAACGGCGGATTACCCGCTTGGGAGCCCGCTCGAG
CCGGCTCCTGGATGGAGTGGCTGAACCCCGTGGAGTTCATGGAGGACCTGGTCGTGG

AGCAGAGTGTACCTCTTCCGCCATTCAGGCTCTGGTCCTGTTCAAGAAGCTGTACCCC
CGACACCGAAACAAGGAGATCGAGAACTGCATCATTAACGCCGCTCAGTTCATTGAG
AACATCCAGGAGCCCGACGGCTCTTGGTACGGCAACTGGGGAATTTGTTTCTCTTACG
GCACCTGGTTCGCTCTGAAGGGCCTGGCCGCTGCCGGACGAACCTACGAGAACTGCT
CCGCCATCCGAAAGGGCGTGGACTTCCTGCTGAAGTCTCAGCGAGACGACGGTGGCT
GGGCCGAGTCTTACCTGTCCTGTCCCAAGAAGGTGTACGTCCCCTTCGAGGGAACCGA
TCCAACCTGGTGCAGACCGCTTGGGCCATGATGGGTCTGATCTACGGAGGTCAAGCT
AAGCGAGACCCCATGCCTCTGCACCGAGCTGCCAAGCTGCTGATCAACTCTCAGACC
GACCTGGGTGACTTCCCTCAGCAGGAGCTGACCGGAGCCTTCATGCGAAACTGCATG
CTGCACTACGCTCTGTTCCGAAACACCTTCCCCATCTGGGCTCTGGCCGAGTACCGAC
GACACGTCCTGTTCCCCTCTGCCGGTTTCGGCTTCGGATTCACCAACAACCTGTAG

## 5. *LUS-KDEL*

ATGTGGCGAATCAAGATCGCCGAGGGCGGTAACAACCCCTACATCTACTCCACCAAC
AACTTCCAGGGCCGACAGATTTGGGTCTTCGACCCTAACGCCGGCACCCCTGAGGAG
CAGGCTGAGGTTGAGGAGGCTCGACAGAACTTTTGGAAGAACCGATTCCAGGTCAAG
CCCAACTCCGACCTGCTGTGGCAGCTGCAATTCCTGCGAGAGAAGAACTTCAAGCAG
AAGATCCCTAAGGTGAAGGTGGAGGACGGCGAGGAGATTACCTCCGAGATTGCCGCT
GCCGCTCTGCGACGATCTGTTCACCTGTTCTCCGCCCTCCAGGCTTCTGACGGTCACTG
GTGTGCTGAGAACGGTGGTCTGCTGTTTTTTCTGCCCCCCCTGGTCTTTGCCGTTTACA
TCACCGGCCACCTGAACACCGTCTTCTCCCCTGAGCACCGAAAGGAGATTCTGCGATA
CATCTACTGTCACCAGAACGAGGACGGTGGCTGGGGTATTCACATCGAGGGTCACTC
CACCATGTTTTGCACCGTCCTCAACTACATCTGCATGCGAATCCTGGGTGAGGCTCGA
GACGGTGGTATCGAGAACGCTTGCGAGCGAGGTCGAAAGTGGATCTTGGACCACGGC
GGTGCTACCGGTATCTCTTCTTGGGGTAAGACCTGGCTGTCCATCCTGGGTGTGTACG
AGTGGGACGGTACTAACCCCATGCCCCCTGAGTTCTGGGCTTTTCCTTCTTCCTTTCCT
CTGCACCCCGCCAAGATGTTCTGTTACTGCCGAATCACCTACATGCCCATGTCCTACC
TGTACGGCAAGCGATTCGTCGGCCCTATTACCCCTCTGATTCTGCAAATCCGAGAGGA
GATTTATAACGAGCCCTACAACAAGATCAAGTGGAACTCCGTCCGACACCTGTGCGC
TAAGGAGGACAACTACTTCCCCCACCCCACCATCCAGAAGCTGCTGTGGGACGCTCT
GTACACCTTCTCCGAGCCTCTGTTCTCCCGATGGCCTTTCAACAAGCTCCGAGAGAAG
GCCCTGAAGATCACCATGGACCACATCCACTACGAGGACCACAACTCCCGATACATC
ACCATCGGCTGCGTTGAGAAGCCCCTGTGTATGCTCGCCTGTTGGATCGAGGACCCCC
ACGGTGAGGCTTTTAAGAAGCACCTGGCCCGAATCGCCGACTACATCTGGGTTGGTG
AGGACGGTATTAAGATGCAGTCCTTTGGCTCCCAGACCTGGGACACCTCTCTGGCTCT
GCAAGCTCTGATCGCCTCCGACCTGTCTCACGAGATCGGTCCTACCCTGAAGCAGGGT
CACGTTTTTACCAAGAACTCCCAGGCCACCGAGAACCCTTCTGGTGACTTCCGAAAGA
TGTTCCGACACATCTCCAAGGGCGCCTGGACCTTTTCCGACAAGGACCAGGGATGGC
AGGTCTCTGACTGCACCGCTGAGTCTCTGAAGTGCTGCCTGCTGTTTTCCATGATGCC
CCCCGAGATCGTCGGCGAGAAGATGGAGCCTGAGAAGGTGTACGACTCCGTCAACGT
CATCCTGTCCCTGCAATCCCAGAACGGCGGTTTCACCGCTTGGGAGCCTGCTCGAGCT
GGTTCTTGGATGGAGTGGCTGAACCCCGTGGAGTTCATGGAGGACCTGGTCGTTGAG
CACGAGTACGTTGAGTGCACCTCCTCCGCTATCCAGGCTCTGGTTCTGTTCAAGAAGC

TGTACCCCCGACACCGAAACAAGGAGATTGAGAACTGTATCATTAACGCTGCCCAGT
TTATCGAGAACATCCAGGAGCCCGACGGCTCTTGGTACGGTAACTGGGGTATCTGCTT
CTCCTACGGCACCTGGTTCGCCCTGAAGGGTCTGGCTGCTGCTGGTCGAACCTACGAG
AACTGCTCCGCTATCCGAAAGGGCGTGGACTTTCTGCTGAAGTCCCAGCGAGACGAC
GGCGGTTGGGCTGAGTCTTACCTGTCTTGCCCTAAGAAGGTGTACGTGCCCTTCGAGG
GTAACCGATCCAACCTGGTCCAGACCGCTTGGGCTATGATGGGACTGATTTACGGCG
GACAGGCTAAGCGAGACCCTATGCCTCTGCACCGAGCTGCTAAGCTGCTCATCAACTC
CCAGACCGACCTGGGAGACTTTCCCCAGCAGGAGCTGACCGGTGCTTTTATGCGAAA
CTGCATGCTGCACTACGCCCTCTTCCGAAACACCTTCCCCATCTGGGCCTTGGCTGAG
TACCGACGACACGTTCTGTTTCCCTCCGCCGGTTTTGGCTTTGGCTTCACCAACAACCT
G**AAGGACGAGCTC**TAA

## 6. *LUS-SKL*

ATGTGGCGAATCAAGATCGCCGAGGGCGGTAACAACCCCTACATCTACTCCACCAAC
AACTTCCAGGGCCGACAGATTTGGGTCTTCGACCCTAACGCCGGCACCCCTGAGGAG
CAGGCTGAGGTTGAGGAGGCTCGACAGAACTTTTGGAAGAACCGATTCCAGGTCAAG
CCCAACTCCGACCTGCTGTGGCAGCTGCAATTCCTGCGAGAGAAGAACTTCAAGCAG
AAGATCCCTAAGGTGAAGGTGGAGGACGGCGAGGAGATTACCTCCGAGATTGCCGCT
GCCGCTCTGCGACGATCTGTTCACCTGTTCTCCGCCCTCCAGGCTTCTGACGGTCACTG
GTGTGCTGAGAACGGTGGTCTGCTGTTTTTTCTGCCCCCCCTGGTCTTTGCCGTTTACA
TCACCGGCCACCTGAACACCGTCTTCTCCCCTGAGCACCGAAAGGAGATTCTGCGATA
CATCTACTGTCACCAGAACGAGGACGGTGGCTGGGGTATTCACATCGAGGGTCACTC
CACCATGTTTTGCACCGTCCTCAACTACATCTGCATGCGAATCCTGGGTGAGGCTCGA
GACGGTGGTATCGAGAACGCTTGCGAGCGAGGTCGAAAGTGGATCTTGGACCACGGC
GGTGCTACCGGTATCTCTTCTTGGGGTAAGACCTGGCTGTCCATCCTGGGTGTGTACG
AGTGGGACGGTACTAACCCCATGCCCCCTGAGTTCTGGGCTTTTCCTTCTTCCTTTCCT
CTGCACCCCGCCAAGATGTTCTGTTACTGCCGAATCACCTACATGCCCATGTCCTACC
TGTACGGCAAGCGATTCGTCGGCCCTATTACCCCTCTGATTCTGCAAATCCGAGAGGA
GATTTATAACGAGCCCTACAACAAGATCAAGTGGAACTCCGTCCGACACCTGTGCGC
TAAGGAGGACAACTACTTCCCCCACCCCACCATCCAGAAGCTGCTGTGGGACGCTCT
GTACACCTTCTCCGAGCCTCTGTTCTCCCGATGGCCTTTCAACAAGCTCCGAGAGAAG
GCCCTGAAGATCACCATGGACCACATCCACTACGAGGACCACAACTCCCGATACATC
ACCATCGGCTGCGTTGAGAAGCCCCTGTGTATGCTCGCCTGTTGGATCGAGGACCCCC
ACGGTGAGGCTTTTAAGAAGCACCTGGCCCGAATCGCCGACTACATCTGGGTTGGTG
AGGACGGTATTAAGATGCAGTCCTTTGGCTCCCAGACCTGGGACACCTCTCTGGCTCT
GCAAGCTCTGATCGCCTCCGACCTGTCTCACGAGATCGGTCCTACCCTGAAGCAGGGT
CACGTTTTTACCAAGAACTCCCAGGCCACCGAGAACCCTTCTGGTGACTTCCGAAAGA
TGTTCCGACACATCTCCAAGGGCGCCTGGACCTTTTCCGACAAGGACCAGGGATGGC
AGGTCTCTGACTGCACCGCTGAGTCTCTGAAGTGCTGCCTGCTGTTTTCCATGATGCC
CCCCGAGATCGTCGGCGAGAAGATGGAGCCTGAGAAGGTGTACGACTCCGTCAACGT
CATCCTGTCCCTGCAATCCCAGAACGGCGGTTTCACCGCTTGGGAGCCTGCTCGAGCT
GGTTCTTGGATGGAGTGGCTGAACCCCGTGGAGTTCATGGAGGACCTGGTCGTTGAG
CACGAGTACGTTGAGTGCACCTCCTCCGCTATCCAGGCTCTGGTTCTGTTCAAGAAGC

TGTACCCCCGACACCGAAACAAGGAGATTGAGAACTGTATCATTAACGCTGCCCAGT
TTATCGAGAACATCCAGGAGCCCGACGGCTCTTGGTACGGTAACTGGGGTATCTGCTT
CTCCTACGGCACCTGGTTCGCCCTGAAGGGTCTGGCTGCTGCTGGTCGAACCTACGAG
AACTGCTCCGCTATCCGAAAGGGCGTGGACTTTCTGCTGAAGTCCCAGCGAGACGAC
GGCGGTTGGGCTGAGTCTTACCTGTCTTGCCCTAAGAAGGTGTACGTGCCCTTCGAGG
GTAACCGATCCAACCTGGTCCAGACCGCTTGGGCTATGATGGGACTGATTTACGGCG
GACAGGCTAAGCGAGACCCTATGCCTCTGCACCGAGCTGCTAAGCTGCTCATCAACTC
CCAGACCGACCTGGGAGACTTTCCCCAGCAGGAGCTGACCGGTGCTTTTATGCGAAA
CTGCATGCTGCACTACGCCCTCTTCCGAAACACCTTCCCCATCTGGGCCTTGGCTGAG
TACCGACGACACGTTCTGTTTCCCTCCGCCGGTTTTGGCTTTGGCTTCACCAACAAC**CT**
**GTCCAAGCTG**TAA

### 7. *tHMGR*-T2A-*ERG1*-T2A-*ERG9*

ATGACCCAGTCTGTGAAGGTGGTCGAGAAGCACGTCCCCATCGTCATTGAGAAGCCC
TCCGAGAAGGAGGAGGACACCTCTTCCGAGGACTCTATTGAGCTGACCGTGGGAAAG
CAGCCCAAGCCCGTCACCGAGACCCGATCCCTGGACGACCTGGAGGCTATCATGAAG
GCCGGAAAGACCAAGCTGCTCGAGGACCACGAGGTGGTCAAGCTGTCTCTGGAGGGC
AAGCTGCCCCTGTACGCCCTGGAGAAGCAGCTGGGAGACAACACCCGAGCTGTGGGT
ATTCGACGATCTATCATTTCCCAGCAGTCTAACACCAAGACCCTCGAGACCTCCAAGC
TGCCCTACCTGCACTACGACTACGACCGAGTGTTCGGAGCTTGTTGCGAGAACGTCAT
CGGTTACATGCCTCTGCCTGTGGGTGTGGCCGGACCCATGAACATTGACGGCAAGAA
CTACCACATCCCCATGGCCACCACCGAGGGTTGTCTGGTGGCTTCTACCATGCGAGGC
TGCAAGGCTATTAACGCCGGCGGAGGTGTGACCACCGTCCTGACCCAGGACGGAATG
ACCCGAGGTCCCTGTGTCTCCTTCCCCTCTCTGAAGCGAGCCGGCGCCGCTAAGATTT
GGCTGGACTCCGAGGAGGGACTGAAGTCTATGCGAAAGGCCTTCAACTCTACCTCCC
GATTCGCTCGACTCCAGTCTCTGCACTCTACCCTGGCCGGAAACCTGCTGTTCATTCG
ATTCCGAACCACCACCGGCGACGCTATGGGAATGAACATGATCTCTAAGGGCGTGGA
GCACTCTCTGGCCGTGATGGTCAAGGAGTACGGTTTCCCCGACATGGACATCGTGTCT
GTCTCCGGCAACTACTGCACCGACAAGAAGCCCGCCGCTATTAACTGGATCGAGGGA
CGGGGCAAGTCTGTGGTGGCTGAGGCTACCATCCCTGCTCACATTGTGAAGTCCGTCC
TGAAGTCTGAGGTGGACGCTCTGGTCGAGCTGAACATCTCTAAGAACCTGATTGGCTC
TGCCATGGCTGGTTCTGTGGGCGGATTCAACGCTCACGCCGCTAACCTGGTCACCGCC
ATCTACCTGGCTACCGGTCAGGACCCCGCTCAGAACGTGGAGTCTTCCAACTGTATTA
CCCTGATGTCCAACGTGGACGGAAACCTGCTGATCTCTGTGTCCATGCCCTCTATCGA
GGTCGGCACCATTGGTGGCGGAACCATCCTGGAGCCTCAGGGAGCTATGCTGGAGAT
GCTGGGAGTGCGAGGTCCCCACATTGAGACCCCTGGTGCTAACGCTCAGCAGCTGGC
TCGAATCATTGCCTCCGGAGTGCTGGCCGCTGAGCTGTCCCTGTGCTCTGCCCTGGCC
GCTGGACACCTGGTCCAGTCCCACATGACCCACAACCGATCTCAGGCTCCTACCCCTG
CTAAGCAGTCCCAGGCCGACCTCCAGCGACTCCAGAACGGCTCCAACATCTGTATTCG
ATCT**CGAGCCGAGGGCCGAGGATCTCTGCTGACCTGCGGCGACGTGGAGGAGAA**
**CCCCGGACCC**ATGGTCACCCAGCAGTCCGCCGCTGAGACCTCTGCCACCCAGACCAA
CGAGTACGACGTGGTCATTGTGGGTGCTGGCATTGCCGGACCCGCCCTGGCTGTGGCC
CTGGGCAACCAGGGACGAAAGGTCCTGGTGGTCGAGCGAGACCTGTCTGAGCCCGAC

CGAATTGTGGGAGAGCTGCTCCAGCCCGGCGGCGTGGCCGCTCTGAAGACCCTGGGA
CTGGGTTCCTGTATTGAGGACATCGACGCCATTCCCTGCCAGGGTTACAACGTGATCT
ACTCTGGCGAGGAGTGTGTGCTGAAGTACCCCAAGGTCCCCCGAGACATTCAGCAGG
ACTACAACGAGCTGTACCGATCCGGAAAGTCTGCCGACATCTCCAACGAGGCTCCCC
GAGGTGTGTCTTTCCACCACGGCCGATTCGTCATGAACCTGCGACGAGCCGCTCGAGA
CACCCCTAACGTGACCCTGCTGGAGGCTACCGTCACCGAGGTGGTCAAGAACCCCTA
CACCGGTCACATCATTGGCGTGAAGACCTTCTCCAAGACCGGAGGTGCCAAAATCTA
CAAGCACTTCTTCGCTCCCCTGACCGTGGTCTGCGACGGAACCTTCTCCAAGTTCCGA
AAGGACTTCTCTACCAACAAGACCTCCGTCCGATCTCACTTCGCCGGACTGATTCTGA
AGGACGCTGTGCTGCCTTCTCCCCAGCACGGTCACGTCATCCTGTCCCCCAACTCTTG
TCCTGTGCTGGTCTACCAAGTGGGCGCCCGAGAGACCCGAATCCTGTGCGACATTCAG
GGACCCGTGCCCTCTAACGCTACCGGAGCCCTGAAGGAGCACATGGAGAAGAACGTG
ATGCCCCACCTGCCCAAGTCCATTCAGCCCTCTTTCCAGGCCGCTCTGAAGGAGCAGA
CCATCCGAGTCATGCCCAACTCCTTCCTGTCTGCCTCCAAGAACGACCACCACGGCCT
GATCCTGCTGGGAGACGCTCTGAACATGCGACACCCTCTGACCGGCGGAGGTATGAC
CGTGGCTCTGAACGACGCCCTGCTGCTGTCTCGACTGCTGACCGGCGTCAACCTGGAG
GACACCTACGCCGTGTCTTCCGTCATGTCTTCCCAGTTCCACTGGCAGCGAAAGCACC
TGGACTCCATCGTGAACATTCTGTCCATGGCCCTGTACTCTCTGTTCGCCGCTGACTCT
GACTACCTGCGAATCCTCCAGCTGGGTTGTTTCAACTACTTCAAGCTGGGCGGAATCT
GCGTGGACCACCCCGTCATGCTGCTGGCTGGAGTCCTGCCCCGACCCATGTACCTGTT
CACCCACTTCTTCGTGGTCGCCATCTACGGTGGCATTTGTAACATGCAGGCTAACGGA
ATTGCCAAGCTGCCCGCTTCCCTGCTCCAGTTCGTGGCCTCTCTGGTCACCGCTTGTAT
CGTGATTTTCCCCTACATCTGGTCTGAGCTGACC**CGAGCCGAGGGTCGTGGATCTCT**
**GCTGACCTGCGGCGACGTGGAAGAGAACCCCGGTCCC**ATGGGCAAGCTGATTGA
GCTGCTGCTGCACCCTTCTGAGCTGTCTGCCGCTATCCACTACAAGCTGTGGCGACAG
CCTCTGCACCCCCGAGACCTGTCCAAGGAGTCTACCGAGCTGCGACGATGTTACGAG
CTGCTGGACGTGTGCTCTCGATCCTTCGCCGCTGTCATTCGAGAGCTGCACCCTGAGG
TGCGAGACGCCGTCATGCTGTTCTACCTGATTCTGCGAGCTCTGGACACCATCGAGGA
CGACATGACCCTGTCCCGAGACATCAAGATTCCCATCCTGCGAGACTTCACCAAGTGT
ATGAAGACCCCCGGATGGAAGTTCACCGACTCTGACCCCAACGAGCGAGACCGAGTG
GTCCTCCAGGAGTTCCCCGTGGTCATGACCGAGTTCAACAAGCTGAAGCCCAAGTAC
CAGGAGGTCATCTACGACATTACCGACCGAATGGGTAACGGCATGGCCGACTACGTG
ATTGACGACGACTTCAACAACAACGGTGTGGACACCATCGCCGCTTACGACCTGTACT
GTCACCACGTCGCTGGTATTGTGGGAGAGGGTCTGACCCGAATTACCATCCTGGCTGG
CTTCGGAACCGACGTGCTGCACGAGAACCCCCGACTCCAGGAGTCTATGGGACTGTT
CCTCCAGAAGGTGAACATCATTCGAGACTACCGAGAGGACATTGACGTCAACCGAGC
CTTCTGGCCCCGAGAGATTTGGCACAAGTACGCTGAGGAGATGCGAGACTTCAAGGA
CCCCAAGTACTCCAAGAAGGCTCTGCACTGTACCTCTGACCTGGTGGCTAACGCCCTG
GGCCACGCTACCGACTGCCTGGACTACCTGGACAACGTCACCGACCCCTCCACCTTCA
CCTTCTGTGCCATTCCCCAGGTCATGGCTATCGCCACCCTGGACCTGGTCTACCGAAA
CCCCGACGTGTTCCAGAAGAACGTCAAGCTGCGAAAGGGCACCACCGTGTCTCTGAT
TCTGGAGGCCTCCAACGTGTCTGGAGTCTGTGACATCTTCACCCGATACGCTCGAAAG
GTCTACAAGAAGTCCGACCCCAACGACCCCAACTACTTCCGAGTGTCTGTCCTGTGCG
GCAAGATTGAGCAGCACGCCGCTCTGATCAAGCGACAGCGAGGCCCTCCCGCCAAGA

CCATCGCTCAGCTGGAGGGAGAGCGAAAGGAGATGGCCCTGTCCCTGATTGTGTGCC
TGGCTGTCATCTTCTCTATGTCCGGACTGATGGCTTACATTGCCTACGTGTCCGGTTTC
CGATGGTCTCCCCGAGAGATTTTCGACTCTAAGATGTTCCCCCTGCGAGACTAG

**8.** ***Pex30*-T2A-*Rtn1*-T2A-*Ypo1***

ATGTCCGGTAAACCACCACAACGTCCACGAGACCCGAGCTAAGTTCGCCGAGACCCT
GCAACCTCGAATCGGTGGTAACACCACCAAGGTCATCCGAGCCGCTCTGGAGAAGAA
CGAGGCTGAGTCTGGCGTCTCCGAGGACAACGACAACGGTTCTCTGGAGAAGGTGAA
CGTCGCCACCTCCCCTCTGCTGACCTCTACCCCTCCTACCATTTCGAAGGCCCTCGTTA
AGCTGTACCCCTACCTGATTCGATTGACGAGTTCCTGAACGTCGTCACCTGGACCGGA
AAGAACATCTGGTCCTCCGTCCTGATGCTCTGCCTCTTCATCACCGTCGAGTACTTCG
AGACCCTCGTCAAGTACTTCGGCCACCTGGCTATTATCGCCATTCGTGGGGCTACTCC
CTGCTCGACAACTACATCGAGGGCACCCTGTCTTCCTCCCCTACCCTGGAGGACATCG
CTCTGCTGATGAACCGAGTCTCCCTGAAGTCCGACATCCTGCTGTCCCCTATGGTCAA
CCTCGGCACCCAGGACATCCAGCGACTGCTGTACACCACCGTCATCCTCTCCCCTATC
TACGTCATGATCACCTGGCTGCTCCTGCCCCCTCGATCTCTGATGCTGATGGTCGGTAT
GTTTCTGCTGACCTACCACCCCCCTGGTCCAAGGTTGCTCGACGACTGCTGTGGAAGT
TCAAGATCGTCCGACTGCTCGTCTTCTACGTCACCGGTCTCGACCTCGGTGGTATCAA
CAAGGACCAGGGCATCTTCGCCACCGTCCAGAAGCAGGTTAAGAAGCTGGCCTCCAC
CGAGAACTCCAACGGAGTTCTGTCCGACTCCAAGCCCATCCGATTTACTTACGTGCTG
TACGAGAACCAGCGACGATGGCTGGGAATTGGCTGGAAGCCTTCCATGCTCTCCTAC
GAGCGAACCCCTTGGACCGACGAGTTCCTCAACGAGGCTCCTTCCCCTGAGAACTTCC
ACCTGCCTGAGGAGACTAATACTATGGTTTGGCGATGGGTCGATAAGACCTGGCGAC
TGGACATGACCAACGACGGAGCTATCCAGGTCCCCAACTCCAAGGCTCGAACCTCTG
CTGACCCCTCTCCTGACGAGGGTTTCATCTACTACGACAACACCTGGAAGAAGCCCTC
CAAGGAGGACTCCTTCTCTAAGTACACCCGACGACGACGATGGGTCCGAACCGCTGA
GCTGGTTAAGACCTCCGACTTTGACGAGTCCGTCATCAACTCCAACCGAAACTCCGCC
ATCGAGCAGAAGGTCGAGGAGAACTCCACCAACGGACTGACCGCTGAGCAGGAGCT
GGGTTCTAACAAGCAGGAGAAGGACAACGCCAAGAAGGTCGGAGAGCCACCACCGA
AGAGACCAAGGAGTTTGCTGAGGCCTCCAACATCAACGAGGGCGAGTTCGAGCGAAT
TTCTTCTACCGACGAGGAGGTCCTGAAGTCCCGAGCTCGAGACCGACTGGCTAAGGTT
CTGGACGACACCGAGGAGAAGGAGCAGTCTAACCCCACCATTGGTCGAGACTCCAAG
AAGGCCGTCTGA**CGAGCTGAGGGTCGAGGTTCTCTGCTGACCTGTGGAGACGTC**
**GAGGAGAACCCGGTCCT**GCTACCATGTCCGCTTCCGCTCAGCACTCTCAGGCTCAGC
AGCAGCAGCAGCAGAAGTCTTGCAACTGTGACCTGCTGCTCTGGCGAAACCCCGTTC
AGACCGGTAAGTACTTTGGCGGTTCCCTGCTGGCCCTCCTGATTCTGAAGAAGGTCAA
CCTCATCACCTTCTTCCTGAAGGTCGCCTACACCATCCTCTTCACCACCGGATCTATTG
AGTTCGTCTCCAAGCTGTTCCTCGGCCAGGGTCTGATCACCAAGTACGGTCCTAAGGA
GTGCCCCAACATCGCCGGTTTCATTAAGCCCCACATCGACGAGGCCCTGAAGCAGCT
GCCTGTTTTCCAGGCTCACATCCGAAAGACCGTCTTTGCCCAGGTGCCCAAGCACACC
TTTAAGACCGCCGTCGCTCTGTTCCTGCTGCACAAGTTCTTTTCCTGGTTTTCCATCTG
GACCATTGTCTTCGTTGCCGACATCTTCACCTTCACCCTGCCCGTTATCTACCACTCCT
ACAAGCACGAGATTGACGCCACCGTCGCCCAGGGTGTTGAGATCAGCAAGCAGAAGA

CCCAGGAGTTCTCCCAGATGGCCTGCGAGAAGACCAAGCCTTACCTGGACAAGGTCG
AGTCCAAGCTGGGTCCTATCTCCAACCTCGTCAAGTCCAAGACCGCCCCTGTCTCTTC
CACCGCTGGTCCTCAGACCGCTTCTACCTCTAAGCTCGCCGCTGACGTTCCCCTGGAG
CCTGAGTCTAAGGCTTACACCTCCTCCGCCCAGGTCATGCCTGAGGTTCCTCAGCACG
AGCCTTCCACCACCCAGGAGTTTAACGTCGATGAGCTCTCTAACGAGCTGAAGAAGT
CCACCAAGAACCTGCAAAACGAGCTCGAGAAGAACAACGC**CCGAGCCGAGGGTCG**
**AGGCTCTCTGCTGACCTGCGGTGACGTTGAGGAGAACCCCGGTCC**TGCTACCATG
TCTGAGTACGCTTCCTCCATCCACTCCCAGATGAAGCAGTTCGACACCAAGTACTCCG
GCAACCGAATCCTGCAACAGCTGGAGAACAAGACCAACCTGCCCAAGTCCTACCTCG
TCGCTGGTCTGGGTTTCGCTTACCTCCTGCTGATCTTCATCAACGTCGGCGGCGTCGGT
GAGATCCTGTCTAACTTCGCCGGCTTTGTCCTGCCCGCTTACCTGTCTCTGGTCGCTCT
GAAGACCCCCACCTCTACCGACGACACCCAGCTGCTGACCTACTGGATTGTTTTTTCC
TTTCTCTCCGTCATCGAGTTCTGGTCCAAGGCCATTCTGTACCTCATCCCCTTCTACTG
GTTCCTCAAGACCGTTTTCCTGATCTACATCGCCCTCCCCCCAGACCGGCGGTGCTCG
AATGATCTACCAGAAGATCGTCGCTCCCCTGACCGACCGATACATCCTGCGAGACGTC
TCTAAGACCGAGAAGGACGAGATCCGAGCCTCCGTTAACGAGGCTTCCAAGGCTACC
GGCGCTTCTGTTCACTAA

## 9. *GPD1*

ATGTTCGAGAACATCTCCTCCAACGGCGTCTACAAGAACCTGTTCGACGGCAAGTGG
GTCGAGTCCAAGACCAACAAGACCATCGAGACCCACTCCCCCTACGACGGTTCTCTG
ATCGGTAAGGTCCAGGCCCTGTCCAAGGAGGAGGTTGACGAGATTTTTAAGTCTTCCC
GAACCGCCCAGAAGAAGTGGGGAGAGACCCCTATCAACGAGCGAGCTCGAATCATG
CGAAAGGCCGCTGACATCCTGGACGACAACGCTGAGTACATTGCTAAGATCCTGTCC
AACGAGATCGCCAAGGACCTGAAGTCTTCCCTCTCCGAGGTCAAGCGAACCGCTGAC
TTCATCCGATTCACCGCCAACGAGGGAACCCACATGGAGGGTGAGGCTATCAACTCC
GACAACTTCCCCGGTTCCAAGAAGGACAAGCTCTCCCTGGTTGAGCGAGTCCCTCTCG
GTATCGTCCTGGCTATCTCCCCTTTTAACTACCCCGTCAACCTGTCCGGTTCCAAGGTT
GCTCCTGCCCTGATCGCTGGAAACTCCGTTGTCCTCAAGCCCTCCACCACCGGTGCTA
TTTCTGCCCTGCACCTGGCTGAGATTTTTAACGCCGCTGGCCTCCCCGCTGGTGTTCTG
AACACCGTTACCGGTAAGGGCTCCGAGATTGGTGACTACCTCATCACCCACGAGGAG
GTTAACTTTATCAACTTCACCGGCTCCTCCGCCGTTGGTAAGCACATTTCCAAGATTG
CCGGCATGATCCCCATGGTCCTCGAACTGGGTGGTAAGGACGCCGCTATCGTCCTGGA
GGACGCTAACCTGGAGACCACCGCTAAGTCCATCGTCTCCGGAGCTTACGGCTACTCC
GGTCAGCGATGTACCGCTGTTAAGCGAGTCCTGGTGATGGACAAGGTTGCCGACGAG
CTGGTGGAGCTGGTTACCAAGAAGGTGAAGGAGCTGAAGGTCGGCAACCCCTTTGAC
GACGTTACCATTACCCCCCTGATCGACAACAAGGCCGCCGACTACGTCCAGACCCTG
ATTGACGACGCCATCGAGAAGGGAGCCACCCTGATTGTCGGCAACAAGCGAAAGGA
GAACCTGATGTACCCCACCCTCTTTGACAACGTCACCGCCGACATGCGAATTGCCTGG
GAGGAGCCTTTCGGTCCTGTGCTGCCTATCATCCGAGTCAAGTCTATGGACGAGGCCA
TTGAGCTCGCTAACCGATCCGAGTACGGACTGCAATCCGCCGTTTTCACCGAGAACAT
GCACGACGCCTTCTACATCGCCAACAAGCTCGACGTTGGTACTGTCCAGGTGAACAA
CAAGCCCGAGCGAGGACCTGACCACTTCCCTTTTCTCGGCACCAAGTCCTCCGGCATG

GGTACTCAGGGTATCCGATACTCCATCGAGGCCATGACCCGACACAAGTCCATTGTCC
TGAACCTGTAAATCGAT

**10.** ***ylYEF***

ATGGCCCGCAACACAACGGACCGCCATCTCACCGTGCTTGTCCATGATCTGCTAAACA
TTGCCGACGAGCATACCGGCAGCTCGCTGCTGAGCACCAACCAGGCTCGCGCGGAGG
CGACAGGCCACATTCTGTGCGAAAAGTCGCGCCACTCTCGAGAGGAGCTCAACGAGT
TTGTCATGAACGTCCGGGGTCTGTCCAACCGGCTGAGCAACCTCAAGTTGAAGCCGC
AGCTGCGACAAGTGATGATTGTAGCGAAACTGCAGGATAAAGACATCATTGCCAAGA
CGCGCGACTTTGCGTCGCTGCTGATGAAACGTGGAATCTCCGTCTACGTGCAGAAAG
AGCTGGCGGCCCATCCTCTGTTCAACCTCAATGGACTTGAGGGAGACGCCAAAAACG
CCGACACAAAGTTCCACACTTGGTCCGAGGTGGCTCTGCCGGACCCCAACAAACTGG
ACCTGGTCGTGACCCTTGGGGGCGACGGAACGGTGCTATTTGTGTCCTGGCTGTTCCA
GCAGATTGTGCCACCGGTGGTCTCCTTTGGCCTGGGCTCTCTGGGATTCCTCACCGAG
TACGAGTGGGACAGACGTGAGGAGACGATCGATTCGATCGACAAAAACGGCATCTAT
CTGTCGTTGAGAATGCGGTTCGAGTGCCGCGTCATCCGAGCTGTCAAGGACGACGGA
GAGGACTGGATGACCCGAGACTTGGACGACGAAATTCGTTCCATGGTTACCTCCCAC
AACTCGACCGACAACCTGGACGAGTACTCGTACGACAAGCATTACGTGGACGCCACG
CACTCGATTCTCAACGACTTGGTGGTTGACCGAGGCACAAACTCCACCATGACCACCA
CAGAGCTGTACACGGACTTTGATCACCTGACCACCGTACAGGCCGATGGACTGGTGA
TTGCCACTCCTTCTGGATCCACGGCGTACTCCCTGTCCGCAGGAGGATCTCTTGTTCAC
CCCGATATCCCCGGCATTCTCATTTCCCCCATTTGTCCCCATACTCTGAGTTTCCGGCC
GGTTGTTGTGCCCGATAATACTACGATTCGAATCGGAGTGCCATACGATGCTCGGGCG
TCGGCGTACTGCTCGTTCGACGGCCGATCGAGGGTGGAACTGACGCCTGGAGACTTT
ATCACCGTCACCGCGTCGCGATTCCCATTCCCCAAGGTGCAGTCGGAGGCTGGGTCCG
AGTGGTATTCTGGTTTGTCCAATACGTTGAACTGGAACCAGCGAAAGCGACAGAAGC
GGTTCACCAACATTTAA

**11.** ***MCE2***

ATGTCCCCTATCATCGACTTTGTTCGACGACAGCTGTCCTCCACCAAGCTGCACGAGG
AGCAGCAGACCGCTACCACCAACGACCTCGTTTCCCGATCCGGATACCTCAACGAGG
GCAAGTACGAGGTCCGACTGAACTGCATCAACGCCGGATGCCTGCAAAAGAAGCTGA
ACTACATCGGAACCGCCATGGACCCCGCTAAGCGACAGCGACTGGGTCTGAACGGAC
TGCTGCCTGCTGGTGTTGAGACCCTGGAGATCCAGAAGGCCCGAGCTCTGCGAGTTCT
CCGATCTAAGCACAACCTGCTGGAGAAGTACATCCTGATGGCCCAGCTGCGAACCAC
CAACGTTCGACTGTTCTACAAGATCGTCATCGACGAGCTGGAGACCGTGCAGCTGGCT
CCTGTTATCTACACCCCCACCGTCGGAACCGCTTGTCTGGAGTACTCCACCATCTACC
CCTTCCTGGCCGCTCCTGGTGTTCCTGACGGTCTGTACCTCACCAAGGCTGAGCTGCC
TGAGCTGTGTCAGACCATCCGAAACTACCGACCCACCGACACCGAGGGTTTCGAGCC
TGAGATCGCTGTTATCTCCGACGGTTCCCGAATCCTCGGCCTGGGTGACCTGGGTACT
AACGGTATGGGTATCCCCATGGGCAAGCTGCAACTCTACGTCGCTGGTGCTGGCATTG
ACCCTCGACGAACCCTGCCTATTATTCTCGACCTGGGCACCAACAACGAGAAGCTGCT

GAACGACGAGTTCTACATCGGCCTGCGACAGAAGCGACCCAACGACGAGGAGTTCTA
CCAGACCGTCGATACCGTCCTCACCGCTCTGCACACCGTTTACCCTAACCTGCTGATC
CAGTTCGAGGACTGGTCCTCCGAGCACGCTTTCGGTCTGCTGGAGAAGTACCAGAAC
CAGATGCTGTGTTTTAACGACGACATCCAGGGCACCGGCGCTGTTATTCTGTCCGGTG
TTATCAACGCCATTCGAAAGGTCGAGAAGGAGAACCAGGTCTCTCCCCGAGACCACC
GAATCGTTTTCTACGGCGCCGGTTCCGCTGCTATCGGTGTTGCTCGACAGATCCAGTC
TTACTTCCAGATTGAGCACAACATGACCGAGGAGGAGGCCAAGCACGTTTTCTGGAT
TGTCGATTCCAAGGGTCTCGTTACCACCACCCGAGGAGACAAGCTGGCTCAGCACAA
GGTCTACTACGCCCGAGGTGACAACGAGGGACAGCAGTACAAGGAGCTCATTGACAT
TGTCAACTACAACCTGTACTCCCTGATCGGCCTGTCCTCTACCACCGGTGCTTTCAAC
ACCCAGGTCCTGGAGCGACTGGCTTCTCTGAACGAGCAGCCTATCGTTTTCCCCCTCT
CCAACCCCGCTACCCAGGCTGAGTGTACCTTTGAGCAGGCCATGGAGGCTACCAACA
ACAAGGTCATCTTCGCCTCCGGCACCGCTTTTCCTGCTTACACCATCAAGTCTACCGG
CGAGGTCAACACCCCCGGTCAGGGTAACAACATGTACATCTTCCCCGGTCTGGGCCTC
GGAGCTTGTCTGGCTAACCCTGCTCACTTCGACCGAATGATCTACGAGGCCTCTAAGG
CTCTGGCCGACTCTCTGACCGAGGAGGAGATTTCCAAGGCCTGGCTCTACCCCTCCCT
CAACTACCGATCCGTCTCCGCTATCGTCGCCGCTGCTGTTTGTCAGGAGACCCTGAAC
GAGAACCTGGCCACCTCTCAGGCTATGATGACCCAGTGCAAGTCTCACGAGGACATC
CTGGACTACGTTTCCGCCCACATGTGGTCTCCCGACTACGGTAACAACAACTCCAACC
AGCAGGCCGGCAAGCTGTAA

**12.** ***CkPTA***

ATGAAGCTGATGGAGAACATCTTCGGCCTGGCCAAGGCCGACAAGAAGAAGATCGTC
CTGGCCGAGGGCGAGGAGGAGCGAAACATCCGAGCTTCTGAGGAGATCATCCGAGA
CGGCATCGCCGACATCATTCTCGTCGGTTCTGAGTCCGTCATCAAGGAGAACGCCGCC
AAGTTCGGTGTGAACCTGGCTGGTGTCGAGATCGTCGATCCCGAGACCTCTTCCAAGA
CCGCTGGTTACGCCAACGCCTTCTACGAGATCCGAAAGAACAAGGGAGTCACCCTGG
AGAAGGCCGACAAGATCGTCCGAGACCCTATCTACTTCGCTACTATGATGGTTAAGCT
GGGCGACGCCGACGGACTGGTTTCTGGTGCTATCCACACCACCGGCGACCTGCTGCG
ACCTGGTCTGCAAATTGTCAAGACCGTCCCCGGCGCTTCTGTTGTTTCCTCTGTGTTCC
TGATGTCCGTTCCCGACTGCGAGTACGGTGAGGACGGTTTTCTGCTGTTTGCCGACTG
CGCCGTGAACGTTTGCCCTACCGCTGAGGAGCTGTCCTCTATCGCTATCACCACCGCC
GAGACCGCTAAGAACCTGTGTAAGATCGAGCCCCGAGTGGCCATGCTCTCTTTCTCTA
CCATGGGCTCCGCCTCCCACGAGCTGGTTGACAAGGTTACCAAGGCCACCAAGCTCG
CTAAGGAGGCTCGACCTGACCTGGACATTGACGGTGAGCTGCAACTGGACGCCTCCC
TGGTTAAGAAGGTTGCCGACCTGAAGGCTCCCGGATCTAAGGTTGCCGGTAAGGCTA
ACGTTCTGATTTTTCCCGACATTCAGGCCGGCAACATCGGCTACAAGCTGGTTCAGCG
ATTCGCCAAGGCCGAGGCTATTGGTCCTATCTGCCAGGGTTTCGCCAAGCCTATCAAC
GACCTGTCCCGAGGTTGCTCCGTCGATGACATCGTCAAGGTCGTCGCCGTCACCGCTG
TTCAGGCTCAGGCTCAGGGTTAA

**13.** ***BbPK***

ATGACCTCCCCTGTGATCGGTACTCCCTGGAAGAAGCTGAACGCCCCTGTCTCTGAGG
AGTCCCTGGAGGGTGTTGACAAGTACTGGCGAGTGGCCAACTACCTGTCCATCGGTC
AGATTTATCTGCGATCTAACCCCCTGATGAAGGCCCCTTTTACCCGAGAGGACGTCAA
GCACCGACTGGTTGGTCACTGGGGAACCACCCCTGGTCTGAACTTCCTGATCGGCCAC
ATCAACCGATTCATTGCCGACCACGGCCAGAACACCGTCATCATTATGGGACCCGGC
CACGGAGGTCCTGCTGGTACTTCTCAGTCCTACCTCGACGGCACCTACACCGAGACCT
TTCCTAAGATTACCAAGGACGAGGCCGGTCTCCAGAAGTTTTTCCGACAGTTTTCCTA
CCCCGGCGGCATCCCTTCTCACTTTGCTCCTGAGACCCCCGGTTCCATCCACGAGGGC
GGTGAGCTGGGTTACGCTCTGTCTCACGCTTACGGCGCTATCATGGACAACCCCTCCC
TGTTTGTCCCCGCTATTGTCGGTGACGGTGAGGCTGAGACCGGTCCTCTGGCTACCGG
TTGGCAGTCTAACAAGCTGGTTAACCCCCGAACCGACGGCATCGTTCTGCCTATTCTG
CACCTGAACGGCTACAAGATTGCTAACCCCACCATTCTGTCCCGAATCTCCGACGAGG
AGCTCCACGAGTTCTTCCACGGTATGGGCTACGAGCCCTACGAGTTTGTCGCCGGTTT
TGACGACGAGGACCACATGTCCATCCACCGACGATTCGCCGAGCTGTGGGAGACCAT
TTGGGACGAGATTTGCGACATCAAGGCCGCTGCCCAGACCGACAACGTTCACCGACC
TTTTTACCCCATGCTGATCTTCCGAACCCCCAAGGGTTGGACCTGTCCTAAGTACATC
GACGGCAAGAAGACCGAGGGCTCCTGGCGAGCTCACCAGGTTCCTCTGGCTTCTGCT
CGAGACACCGAGGCTCACTTCGAGGTTCTGAAGAACTGGCTGGAGTCCTACAAGCCC
GAGGAGCTGTTCGACGCTAACGGTGCTGTGAAGGACGACGTCCTGGCTTTCATGCCC
AAGGGTGAGCTGCGAATCGGCGCTAACCCTAACGCTAACGGCGGTGTTATCCGAGAC
GACCTGAAGCTCCCCAACCTGGAGGACTACGAGGTCAAGGAGGTGGCTGAGTACGGC
CACGGTTGGGGTCAGCTGGAGGCTACCCGAACCCTGGGTGCTTACACCCGAGACATC
ATTCGAAACAACCCCCGAGACTTCCGAATTTTCGGCCCCGACGAGACCGCTTCTAACC
GACTGCAAGCTAGTTACGAGGTGACCAACAAGCAGTGGGACGCCGGTTACATTTCCG
ACGAGGTCGATGAGCACATGCACGTTTCCGGCCAGGTGGTTGAGCAGCTGTCTGAGC
ACCAGATGGAGGGATTCCTGGAGGCTTACCTGCTGACCGGTAGACATGGCATCTGGT
CCTCCTACGAGTCCTTTGTCCACGTCATTGACTCCATGCTGAACCAGCACGCCAAGTG
GCTCGAGGCTACCGTTCGAGAGATTCCCTGGCGAAAGCCCATCGCTTCCATGAACCTG
CTGGTCTCCTCCCACGTCTGGCGACAGGACCACAACGGTTTCTCCCACCAGGACCCCG
GTGTTACCTCTGTTCTGCTGAACAAGTGCTTTCACAACGACCACGTTATCGGCATCTA
CTTCGCCACCGACGCCAACATGCTGCTCGCTATTGCCGAGAAGTGCTACAAGTCTACC
AACAAGATCAACGCCATCATCGCCGGCAAGCAGCCTGCTGCTACCTGGCTGACCCTG
GACGAGGCTCGAGCTGAGCTGGCTAAGGGTGCTGCTGCTTGGGACTGGGCTTCTACC
GCTAAGAACAACGACGAGGCCGAGGTGGTTCTGGCTGCTGCTGGTGACGTTCCTACC
CAGGAGATTATGGCCGCCTCCGACAAGCTGAAGGAGCTGGGTGTTAAGTTTAAGGTC
GTCAACGTGGCCGACCTGCTGTCTCTGCAATCCGCTAAGGAGAACGACGAGGCTCTG
TCCGACGAGGAGTTCGCTGACATCTTCACCGCCGACAAGCCCGTTCTGTTCGCTTACC
ACTCCTACGCCCACGACGTCCGAGGTCTGATTTACGACCGACCCAACCACGACAACTT
CAACGTGCACGGCTACGAGGAGGAGGGTTCTACCACCACCCCTTACGACATGGTTCG
AGTCAACCGAATTGACCGATACGAGCTGACCGCTGAGGCCCTGCGAATGATTGACGC
TGACAAGTACGCCGACAAGATCGACGAGCTGGAGAAGTTCCGAGACGAGGCTTTTCA
GTTCGCCGTCGATAAGGGATACGACCACCCCGACTACACCGACTGGGTTTACTCCGGT
GTCAACACCGACAAGAAGGGCGCTGTGACCGCTACCGCTGCTACCGCTGGTGACAAC
GAGTAA

**14.** ***LmPK***

ATGGCTGACTTCGACTCCAAGGAGTACCTGGAGCTGGTGGACAAGTGGTGGCGAGCT
ACCAACTACCTGTCCGCTGGTATGATCTTCCTGAAGTCCAACCCCCTGTTCTCCGTGA
CCAACACCCCTATCAAGGCCGAGGACGTCAAGGTGAAGCCCATCGGTCACTGGGGTA
CTATCTCCGGTCAGACCTTTCTCTACGCCCACGCTAACCGACTGATCAACAAGTACGG
ACTCAACATGTTCTACGTCGGTGGCCCCGGTCACGGTGGTCAGGTTATGGTTACCAAC
GCCTACCTCGACGGCGCTTACACCGAGGACTACCCTGAGATTACCCAGGACATCGAG
GGCATGTCCCACCTGTTCAAGCGATTCTCCTTTCCCGGCGGCATCGGATCTCACATGA
CCGCTCAGACCCCCGGTTCTCTGCACGAGGGTGGTGAGCTGGGTTACTCTCTGTCCCA
CGCTTTCGGTGCCGTCCTGGACAACCCTGACCAGGTTGCTTTTGCTGTGGTCGGCGAC
GGAGAGGCTGAGACCGGTCCTTCTATGGCTTCCTGGCACTCTATCAAGTTCCTCAACG
CCAAGAACGACGGTGCCGTCCTCCCTGTTCTGGACCTGAACGGTTTTAAGATTTCCAA
CCCCACCATTTTCTCCCGAATGTCCGACGAGGAGATCACCAAGTTCTTCGAGGGCCTG
GGTTACTCCCCCCGATTTATCGAGAACGACGACATTCACGACTACGCCACCTACCACC
AGCTCGCTGCTAACATCCTGGACCAGGCTATTGAGGACATCCAGGCTATCCAGAACG
ACGCCCGAGAGAACGGTAAGTACCAGGACGGAGAGATCCCCGCTTGGCCTGTTATCA
TCGCCCGACTGCCTAAGGGCTGGGGTGGTCCTACCCACGACGCTTCTAACAACCCCAT
CGAGAACTCCTTCCGAGCCCACCAGGTTCCTCTGCCTCTGGAGCAGCACGACCTGGCT
ACCCTGCCTGAGTTTGAGGACTGGATGAACTCCTACAAGCCCGAGGAGCTCTTCAAC
GCCGACGGTTCTCTCAAGGACGAGCTGAAGGCCATCGCCCCTAAGGGTGACAAGCGA
ATGTCCGCTAACCCCATCACCAACGGCGGTGCTGACCGATCTGACCTGAAGCTGCCTA
ACTGGCGAGAGTTCGCCAACGACATTAACGACGACACCCGAGGCAAGGAGTTCGCTG
ACTCTAAGCGAAACATGGACATGGCCACCCTGTCCAACTACCTCGGAGCTGTTTCCCA
GCTGAACCCCACCCGATTCCGATTTTTCGGCCCCGACGAGACCATGTCCAACCGACTG
TGGGGACTGTTCAACGTCACCCCTCGACAGTGGATGGAGGAGATCAAGGAGCCTCAG
GACCAGCTGCTGTCCCCTACCGGTCGAATTATTGACTCCCAGCTGTCCGAGCACCAGG
CTGAGGGTTGGCTGGAGGGTTACACCCTGACCGGTCGAGTTGGAATCTTCGCCTCCTA
CGAGTCCTTCCTCCGAGTCGTTGACACCATGGTCACCCAGCACTTCAAGTGGCTCCGA
CACGCTTCTGAGCAGGCTTGGCGAAACGACTACCCCTCTCTGAACCTGATCGCCACCT
CCACCGCTTTCCAGCAGGACCACAACGGATACACCCACCAGGACCCCGGTATGCTGA
CCCACCTGGCTGAGAAGAAGTCCAACTTCATCCGAGAGTACCTCCCCGCCGACGGTA
ACTCTCTGCTGGCTGTTCAGGAGCGAGCCTTTTCTGAGCGACACAAGGTCAACCTGCT
GATCGCCTCTAAGCAGCCCCGACAGCAGTGGTTTACCGTTGAGGAGGCCGAGGTGCT
GGCTAACGAGGGTCTGAAGATCATTGACTGGGCTTCCACCGCCCCCTCTTCTGACGTT
GACATCACCTTCGCCTCCGCCGGTACTGAGCCTACCATTGAGACCCTCGCCGCTCTGT
GGCTGATCAACCAGGCTTTCCCCGACGTGAAGTTTCGATACGTCAACGTCGTCGAGCT
CCTGCGACTGCAAAAGAAGTCCGAGCCCAACATGAACGACGAGCGAGAGCTGTCCGC
CGAGGAGTTCAACAAGTACTTCCAGGCCGACACCCCCGTTATCTTCGGATTTCACGCC
TACGAGAACCTCATCGAGTCCTTTTTCTTCGAGCGAAAGTTCACCGGCGACGTCTACG
TGCACGGATACCGAGAGGACGGTGACATCACCACCACCTACGACATGCGAGTCTACT
CCCACCTCGACCGATTCCACCAGGCTAAGGAGGCTGCTGAGATCCTGTCCGCTAACG
GTAAGATCGACCAGGCCGCTGCTGACACCTTCATCGCTAAGATGGACGACACCCTCG

CCAAGCACTTTCAGGTCACCCGAAACGAGGGCCGAGACATTGAGGAGTTCACCGACT
GGACCTGGTCCCCTCTGAAGTAA

**15.** ***AtIPK*-T2A-*ScCK***

ATGGAGCTGAACATCTCGGAGAGCCGATCGCGAAGCATCCGATGTATTGTGAAGCTG
GGCGGCGCCGCTATCACCTGTAAGAACGAGCTGGAGAAGATTCACGACGAGAACCTG
GAGGTGGTGGCCTGTCAGCTCCGACAGGCTATGCTGGAGGGCTCTGCTCCTTCTAAGG
TTATTGGTATGGACTGGTCTAAGCGACCCGGCTCCTCGGAGATCAGCTGTGATGTCGA
CGACATTGGCGACCAGAAGTCTTCCGAGTTTTCCAAGTTTGTGGTCGTCCACGGCGCC
GGTTCTTTTGGTCACTTTCAGGCCTCCCGATCCGGCGTTCACAAGGGTGGTCTGGAGA
AGCCCATCGTGAAGGCCGGTTTTGTTGCCACTCGAATCTCTGTTACCAACCTCAACCT
TGAGATCGTCCGAGCCCTCGCCCGAGAGGGTATTCCTACTATCGGAATGTCGCCCTTC
TCCTGCGGCTGGTCTACCTCTAAGCGAGATGTCGCCTCCGCCGATCTGGCTACTGTCG
CTAAGACCATCGACTCCGGCTTCGTCCCCGTTCTGCACGGTGATGCTGTCCTCGACAA
CATTCTGGGCTGTACCATCCTCTCCGGCGACGTTATCATCCGACACCTTGCCGACCAC
CTCAAGCCCGAGTACGTCGTTTTCCTCACCGATGTTCTCGGCGTGTACGATCGACCCC
CCTCTCCTTCTGAGCCCGATGCTGTTCTGCTGAAGGAGATCGCCGTCGGCGAGGATGG
TTCCTGGAAGGTTGTCAACCCCCTGCTCGAGCACACCGACAAGAAGGTCGACTACTCC
GTCGCCGCCCACGATACTACCGGTGGTATGGAGACCAAGATCAGCGAGGCCGCCATG
ATCGCCAAGCTGGGTGTTGACGTCTACATCGTGAAGGCCGCCACCACCCACTCTCAGC
GAGCTCTTAACGGCGACCTCCGAGATTCGGTCCCCGAGGATTGGCTGGGAACCATTAT
CCGATTTTCCAAG**CGAGCCGAGGGCCGAGGTTCTCTTCTGACTTGCGGCGACGTC**
**GAGGAGAACCCTGGTCCT**GCCACCATGGTGCAGGAGTCCCGACCTGGTTCTGTGCG
ATCCTACTCCGTTGGCTACCAGGCCCGATCTCGATCCTCTTCCCAGCGACGACATTCC
CTGACCCGACAGCGATCTTCCCAGCGACTGATTCGAACTATTTCCATCGAGTCCGATG
TCTCCAACATCACCGACGACGACGATCTCCGAGCCGTCAACGAGGGAGTCGCTGGTG
TTCAGCTGGACGTCTCCGAGACCGCTAACAAGGGCCCTCGACGAGCTTCCGCTACCG
ATGTTACCGACTCCCTGGGTAGCACCTCCTCCGAGTACATTGAGATTCCCTTCGTCAA
GGAGACCCTCGACGCCTCTCTGCCTTCCGATTACCTGAAGCAGGACATTCTCAACCTG
ATTCAGTCCCTGAAGATCAGCAAGTGGTACAACAACAAGAAGATCCAGCCCGTTGCC
CAGGACATGAACCTGGTTAAGATCAGCGGCGCCATGACCAACGCCATTTTCAAGGTC
GAGTACCCCAAGCTCCCCTCCCTGCTTCTGCGAATCTACGGCCCTAACATCGACAACA
TCATCGACCGAGAGTACGAGCTGCAGATCCTGGCCCGACTGTCCCTTAAGAACATCG
GCCCCTCCCTCTACGGCTGCTTCGTTAACGGCCGATTCGAGCAGTTCCTGGAGAACAG
CAAGACCCTGACCAAGGACGACATCCGAAACTGGAAGAACTCCCAGCGAATCGCCCG
ACGAATGAAGGAGCTGCATGTCGGCGTGCCCCTGCTTTCTTCCGAGCGAAAGAACGG
CTCCGCCTGCTGGCAGAAGATCAACCAGTGGCTGCGAACCATCGAGAAGGTCGACCA
GTGGGTCGGCGACCCTAAGAACATCGAGAACTCCCTGCTGTGCGAGAACTGGTCCAA
GTTCATGGATATCGTCGACCGATACCACAAGTGGCTGATCTCCCAGGAGCAGGGAAT
CGAGCAGGTGAACAAGAACCTGATCTTCTGTCACAACGACGCTCAGTACGGCAACCT
GCTGTTCACCGCCCCTGTCATGAACACCCCCTCCCTTTACACCGCCCCCTCTTCTACCA
GCCTGACCTCTCAGTCCTCCTCCCTCTTTCCCTCCTCCTCCAACGTCATCGTCGACGAC
ATCATCAACCCCCCCAAGCAGGAGCAGTCCCAGGATTCTAAGCTCGTCGTCATCGACT

TCGAGTACGCCGGCGCTAACCCCGCTGCTTACGATCTTGCCAACCACCTGTCCGAGTG
GATGTACGACTACAACAACGCCAAGGCCCCCCACCAGTGCCATGCTGATAGATATCC
CGACAAGGAGCAGGTCCTTAACTTCCTGTACTCCTACGTCTCCCACCTGCGAGGCGGT
GCTAAGGAGCCTATCGACGAGGAGGTCCAGCGACTCTACAAGTCCATCATTCAGTGG
CGACCCACCGTGCAGCTGTTCTGGTCTCTCTGGGCTATCCTGCAGTCCGGCAAGCTGG
AGAAGAAGGAGGCCTCCACCGCCATCACCCGAGAGGAGATTGGCCCTAACGGCAAG
AAGTACATCATCAAGACCGAGCCCGAGTCCCCCGAGGAGGATTTTGTCGAGAACGAC
GACGAGCCCGAGGCCGGTGTTTCTATCGACACCTTTGACTACATGGCTTACGGCCGAG
ACAAGATCGCCGTCTTCTGGGGTGACCTGATCGGCCTTGGTATTATCACCGAGGAGGA
GTGCAAGAACTTTTCCTCCTTCAAGTTCCTGGACACCTCCTACCTGTAA

**16. E. coli-*IDI***

ATGCAGACCGAGCACGTCATCCTGCTGAACGCCCAGGGTGTGCCTACCGGTACTCTG
GAGAAGTACGCCGCCCACACCGCTGATACCCGACTTCATCTGGCCTTCTCGTCCTGGC
TGTTTAACGCCAAGGGCCAGCTGCTTGTCACCCGACGAGCTCTTTCCAAGAAGGCCTG
GCCCGGAGTCTGGACTAACTCCGTTTGCGGCCACCCCCAGCTCGGTGAGTCTAACGAG
GACGCTGTCATTCGACGATGCCGATACGAGCTGGGAGTCGAGATCACCCCCCCTGAG
TCTATCTACCCCGACTTCCGATACCGAGCCACTGACCCTTCCGGCATTGTGGAGAACG
AGGTGTGCCCCGTCTTTGCCGCTCGAACTACCTCTGCCCTGCAGATCAACGACGACGA
GGTCATGGACTACCAGTGGTGCGACCTGGCCGACGTTCTTCACGGTATCGACGCTACC
CCCTGGGCTTTCTCCCCTTGGATGGTCATGCAGGCCACCAACCGAGAGGCCCGAAAG
CGACTTTCCGCCTTCACCCAGCTTAAGTAA

**17. *tCYP* (*CYP716A155* with truncated endoplasmic reticulum targeting sequences)**

ATGGAGTTCTTCTACAAGACCAAGACCGGCTCTCTGCCTCCCGGCAAGACCGGTTGGC
CCGTGATTGGAGAGTCTCTGGAGTTCCTGTCCACCGGATGGAAGGGTCACCCCGAGA
AGTTCATCTTCGACCGAATGGCCCGATACTCTTCCCACGTCTTCCGAACCCACCTGCT
GGGAGAGCCCGCCGCTGTCCTGTGTGGCTCCGCCGGAAACAAGTTCCTGTTCTCTAAC
GAGAACAAGCTGGTGCAGGCTTGGTGGCCCTCTTCCGTCGAGAAGATTTTCCCCAACG
ACAACGCCGAGACCTCTTCCAAGGAGGAGTCCATCAAGATGCGACGAATGCTGCCCA
CCTTCTTCAAGCCCGAGGCCCTGCACCGATACGTCGGCATTATGGACCACATCGCCCG
ACGACACTTCGCTGACGGTTGGGACGGCAAGCGAGAGGTGGTCGTGTTCCCCCTGGC
CAAGAACTACACCTTCTGGCTGGCCTGCCGACTGTTCCTGTCTGTGGAGGACCCCTCC
CAGGTCGAGAAGTTCGCCGCTCCCTTCAACCTGCTGGCCTCTGGACTGATCTCCATTC
CTATCGACCTGCCTGGCACCCCTTTCCACAAGGGCATTAAGGCTTCTGCCTACATCCG
AAAGGAGCTGGTGGCCATCATTAAGCAGCGAAAGGCTGACCTGGCTGACGGCACCGC
TTCCCCTACCCAGGACATTCTGTCCCACATGCTGCTGACCTCTAACGAGGACGGCAAG
TTCATGCAGGAGTCTGACATTGCCAACAAGATCCTGGGACTGCTGATCGGCGGACAC
GACACCGCTTCTTCCGCCTGTACCTTCGTCGTGAAGTACCTGGCTGAGCTGCCCCAGG
TGTACGAGGGCGTCTACAAGGAGCAGATGGAGATTGCCAAGTCCAAGGCCGCTGGAG
AGCTGCTGAACTGGGAGGACCTCCAGAAGATGAAGTACTCTTGGACGTGGCTTGTGA

GGTCCTGCGACTGGCCCCTCCCCTCCAGGGAGCTTTCCGAGAGGCTCTGGCCGACTTC
TCTTTCAACGGATTCTCCATCCCCAAGGGTTGGAAGCTGTACTGGTCTGCCAACTCCA
CCCACAAGAACTCCGAGTTCTTCCCCGAGCCCGAGAAGTTCGACCCTTCTCGATTCGA
GGGCTCCGGACCCGCTCCCTACACCTTCGTGCCTTTCGGTGGCGGACCCCGAATGTGT
CCCGGAAAGGAGTACGCCCGACTGGAGATTCTGGTGTTCATGCACCACCTGGTCAAG
CGATTCAAGTGGGAGAAGATGATTCCCGACGAGAAGATCGTCGTGGACCCCATGCCT
ATCCCTGCTAACGGTCTGCCCGTCCGACTGTACCCCCACACCTCCTAG

**18. *tCPR* (*AtATR1* with truncated endoplasmic reticulum targeting sequences)**

ATGACCTCTGCTCTGTACGCCTCCGACCTGTTCAAGCAGCTGAAGTCCATTATGGGCA
CCGACTCTCTGTCCGACGACAAGAAGACCACCGCTGACCGATCCGGAGAGCTGAAGC
CTCTGATGATTCCCAAGTCTCTGATGGCCAAGGACGAGGACGACGACCTGGACCTGG
GCTCTGGCAAGACCCGAGTGTCTATCTTCTTCGGCACCCAGACCGGCACCGCCGAGG
GATTCGCTAAGGCCCTGTCTGAGGAGATTAAGGCTCGATACGAGAAGGCCGCTGTCA
AGGTCATCGACCTGGACGACTACGCCGCTGACGACGACCAGTACGAGGAGAAGCTGA
AGAAGGAGACCCTGGCCTTCTTCTGTGTCGCTACCTACGGCGACGGAGAGCCTACCG
ACAACGCTGCTCGATTCTACAAGTGGTTCACCGAGGAGAACGAGCGAGACATCAAGC
TCCAGCAGCTGGCCTACGGAGTGTTCGCTCTGGGTAACCGACAGTACGAGCACTTCA
ACAAGATTGGCATCGTCCTGGACGAGGAGCTGTGCAAGAAGGGAGCTAAGCGACTGA
TTGAAGTGGGTCTGGGCGACGACGACCAGTCCATCGAGGATGACTTCAACGCCTGGA
AGGAGTCTCTGTGGTCCGAGCTGGACAAGCTGCTGAAGGACGAGGACGACAAGTCTG
TGGCCACCCCCTACACCGCTGTCATTCCCGAGTACCGAGTCGTGACCCACGACCCCCG
ATTCACCACCCAGAAGTCTATGGAGTCCAACGTGGCTAACGGAAACACCACCATTGA
CATCCACCACCCCTGTCGAGTGGACGTGGCCGTCCAGAAGGAGCTGCACACCCACGA
GTCTGACCGATCCTGCATCCACCTGGAGTTCGACATTTCTCGAACCGGCATCACCTAC
GAGACCGGTGACCACGTCGGCGTCTACGCCGAGAACCACGTCGAGATTGTCGAGGAG
GCTGGAAAGCTGCTGGGTCACTCCCTGGACCTGGTCTTCTCTATCCACGCCGACAAGG
AGGACGGCTCTCCCCTGGAGTCTGCCGTGCCCCCTCCCTTCCCCGGACCCTGTACCCT
GGGAACCGGTCTGGCTCGATACGCCGACCTGCTGAACCCTCCCCGAAAGTCCGCTCTG
GTGGCCCTGGCCGCTTACGCCACCGAGCCTTCTGAGGCTGAGAAGCTGAAGCACCTG
ACCTCCCCCGACGGAAAGGACGAGTACTCTCAGTGGATCGTGGCCTCTCAGCGATCC
CTGCTGGAGGTCATGGCTGCTTTCCCCTCTGCTAAGCCTCCCCTGGGCGTCTTCTTCGC
TGCTATTGCTCCCCGACTCCAGCCCCGATACTACTCTATTTCCTCTTCCTCCCGACTGG
CCCCTTCCCGAGTGCACGTCACCTCTGCTCTGGTGTACGGACCCACCCCCACCGGACG
AATCCACAAGGGTGTGTGTTCCACCTGGATGAAGAACGCTGTCCCCGCCGAGAAGTC
TCACGAGTGCTCCGGAGCCCCCATTTTCATCCGAGCTTCTAACTTCAAGCTGCCCTCT
AACCCCTCCACCCCTATCGTGATGGTGGGACCCGGAACCGGTCTGGCCCCCTTCCGAG
GCTTCCTCCAGGAGCGAATGGCTCTGAAGGAGGACGGCGAGGAGCTGGGCTCTTCTC
TGCTGTTCTTCGGCTGCCGAAACCGACAGATGGACTTCATCTACGAGGACGAGCTGA
ACAACTTCGTGGACCAGGGTGTCATTTCCGAGCTGATCATGGCCTTCTCTCGAGAGGG
CGCTCAGAAGGAGTACGTGCAGCACAAGATGATGGAGAAGGCCGCTCAGGTCTGGG
ACCTGATCAAGGAGGAGGGCTACCTGTACGTGTGTGGTGACGCTAAGGGCATGGCCC
GAGACGTCCACCGAACCCTGCACACCATTGTCCAGGAGCAGGAGGGAGTCTCCTCTT

CCGAGGCTGAGGCCATCGTGAAGAAGCTCCAGACCGA
GGTCGATACCTGCGAGACGTCTGGTGA

## 19. *Vhb* (Remove the MluI restriction site through synonymous mutation)

ATGTTGGATCAACAGACCATTAACATCATCAAAGCCACTGTTCCTGTATTGAAGGAGC
ATGG
CGTTACCATTACCACGACTTTTTATAAAAACTTGTTTGCCAAACACCCTGAAGTACGT
CCTT
TGTTTGATATGGGTCGCCAAGAATCTTTGGAGCAGCCTAAGGCTTTGGCGATGACGGT
ATTG
GCGGCAGCGCAAAACATTGAAAATTTGCCAGCTATTTTGCCTGCGGTCAAAAAAATT
GCAG
TCAAACATTGTCAAGCAGGCGTGGCAGCAGCGCATTATCCGATTGTCGGTCAAGAAT
TGTT
GGGTGCGATTAAAGAAGTATTGGGCGATGCCGCAACCGATGACATTTTGGACGCCTG
GGGC
AAGGCTTATGGCGTGATTGCAGATGTGTTTATTCAAGTGGAAGCAGATTTGTACGCTC
AAGC
TGTTGAATAA

## Supplementary references


[1] Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all [J]. Nat Methods, 2022, 19(6): 679-82.

[2] Tian C, Kasavajhala K, Belfon K A A, Raguette L, Huang H, Migues A N, . . . Simmerling C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution [J]. J Chem Theory Comput, 2020, 16(1): 528-52.

[3] Huang L, Li J, Ye H, Li C, Wang H, Liu B, Zhang Y. Molecular characterization of the pentacyclic triterpenoid biosynthetic pathway in *Catharanthus roseus* [J]. Planta, 2012, 236(5): 1571-81.

[4] Li J, Zhang Y. Increase of betulinic acid production in *Saccharomyces cerevisiae* by balancing fatty acids and betulinic acid forming pathways [J]. Appl Microbiol Biotechnol, 2014, 98(7): 3081-9.

[5] Li J, Zhang Y. Modulating betulinic acid production in *Saccharomyces cerevisiae* by managing the intracellular supplies of the co-factor NADPH and oxygen [J]. J Biosci Bioeng, 2015, 119(1): 77-81.

[6] Zhou C, Li J, Li C, Zhang Y. Improvement of betulinic acid biosynthesis in yeast employing multiple strategies [J]. BMC Biotechnol, 2016, 16(1): 59.

[7] Czarnotta E, Dianat M, Korf M, Granica F, Merz J, Maury J, . . . Blank L M. Fermentation and purification strategies for the production of betulinic acid and its lupane-type precursors in *Saccharomyces cerevisiae* [J]. Biotechnol Bioeng, 2017, 114(11): 2528-38.

[8] Tang M, Xu X, Liu Y, Li J, Du G, Lv X, Liu L. Combinatorial Metabolic Engineering for Improving Betulinic Acid Biosynthesis in *Saccharomyces cerevisiae* [J]. ACS Synth Biol, 2024, 13(6): 1798-808.

[9] Tang S, Ji W, Zhao Y, Zhang J, Wei D, Wang F Q. De novo biosynthesis of betulinic acid in engineered *Saccharomyces cerevisiae* [J]. Bioorg Chem, 2024, 152: 107737.

[10] Jin C C, Zhang J L, Song H, Cao Y X. Boosting the biosynthesis of betulinic acid and related triterpenoids in *Yarrowia lipolytica* via multimodular metabolic engineering [J]. Microb Cell Fact, 2019, 18(1): 77.

[11] Sun J, Zhang C B, Nan W H, Li D S, Ke D, Lu W Y. Glycerol improves heterologous biosynthesis of betulinic acid in engineered [J]. Chem Eng Sci, 2019, 196: 82-90.

[12] Madzak C, Beckerich J-M. Heterologous Protein Expression and Secretion in Yarrowia lipolytica [M]//BARTH G. Yarrowia lipolytica: Biotechnological Applications. Berlin, Heidelberg; Springer Berlin Heidelberg. 2013: 1-76.

[13] Nicaud J M, Madzak C, van den Broek P, Gysler C, Duboc P, Niederberger P, Gaillardin C. Protein expression and secretion in the yeast *Yarrowia lipolytica* [J]. FEMS Yeast Res, 2002, 2(3): 371-9.