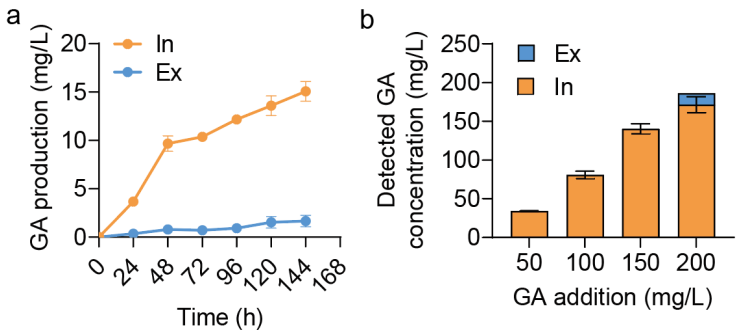
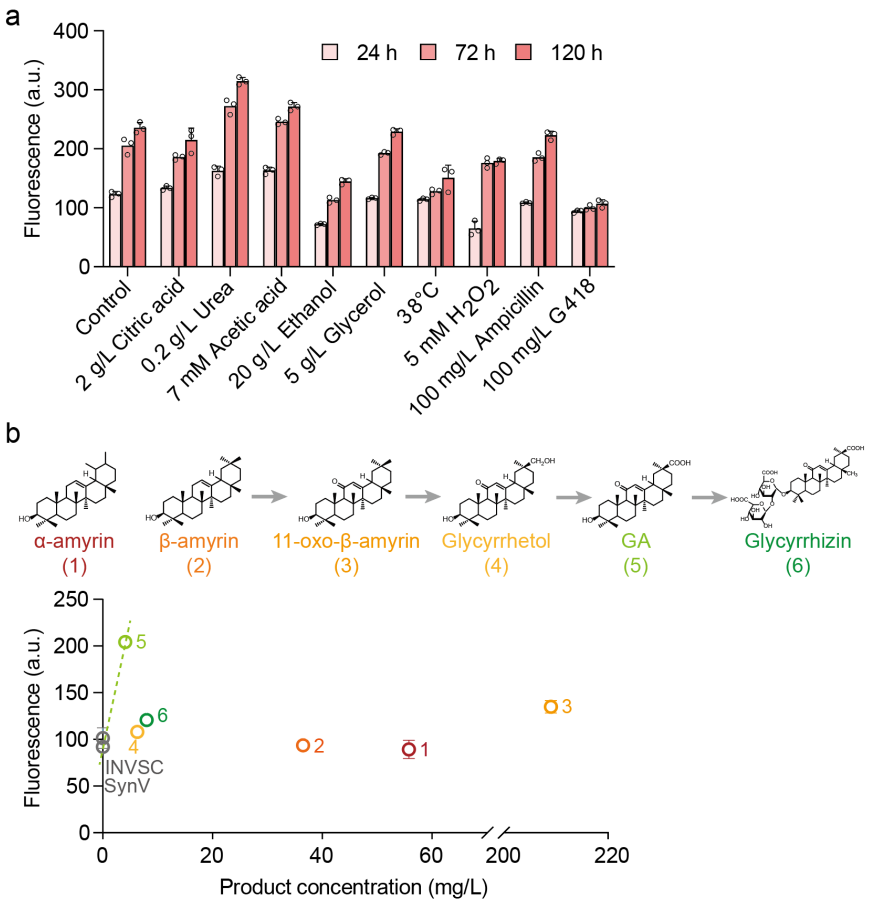
Design of the biosensor-dependent coupling system stabilizes the high-synthesis phenotype of cell factory

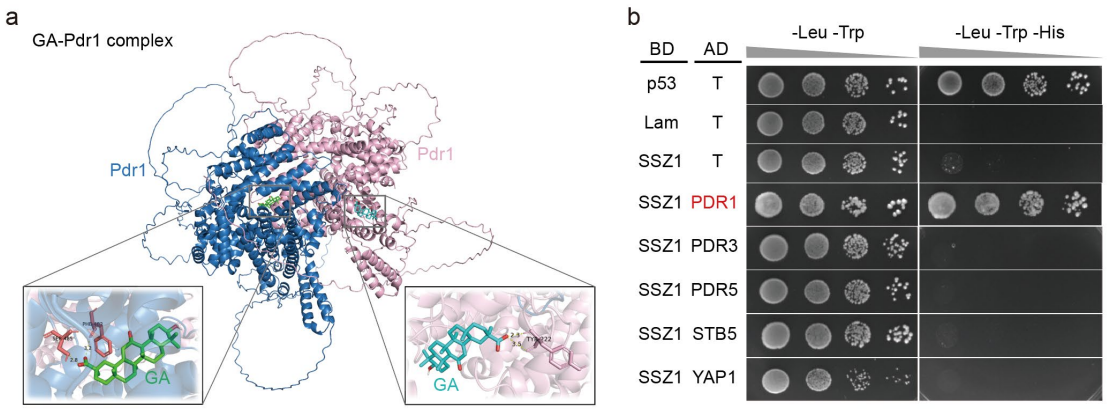
Qin *et al.*



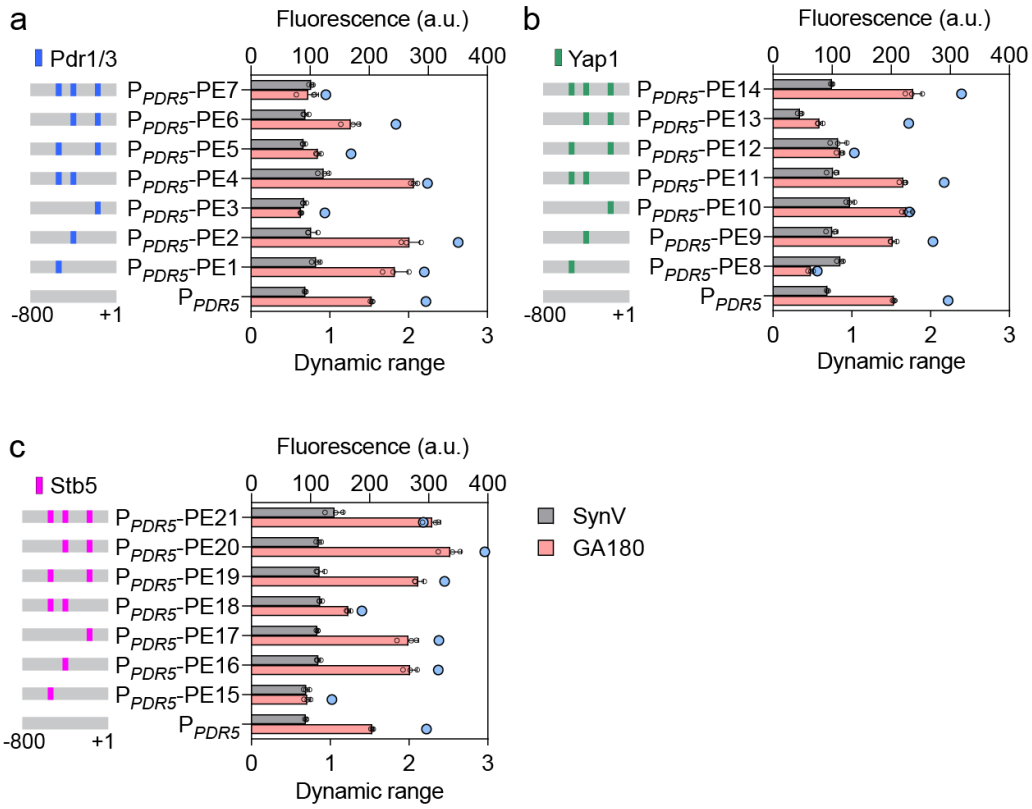
Supplementary Fig. 1 Distribution of GA intracellular and extracellular. (a) Intracellular (In) and extracellular (Ex) amount of glycyrrhetinic acid (GA) synthesized by strain GA180. (b) Intracellular and extracellular amount of GA after adding GA in yeast SynV cultures. Three biological replicates were performed and data are presented as mean values ± SD. Source data for this figure is available in the Source Data file.



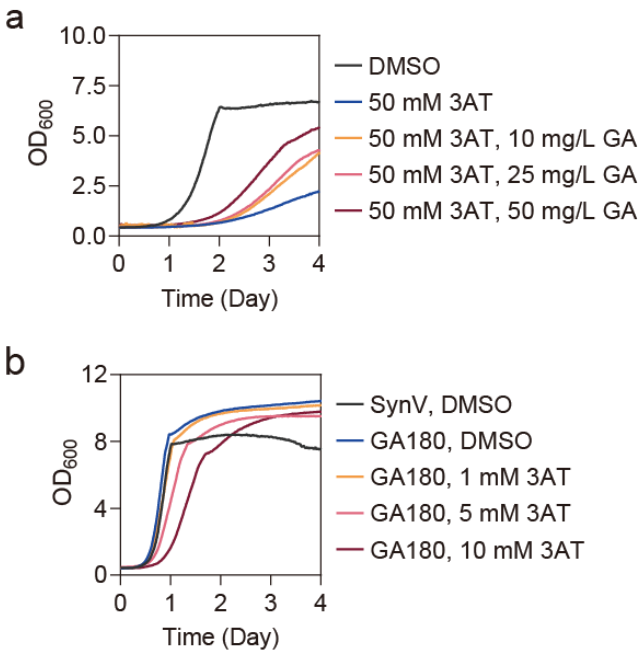
Supplementary Fig. 2 Potential interference factors affecting the GA responsive element P_{PDR5}. (a) Effects of additional metabolites and exerting stress conditions on the response of the P_{PDR5} biosensor. The compound was added to the yeast culture media in a reasonable amount (based on the concentration of cell metabolites under normal conditions). **(b)** The response of the P_{PDR5} biosensor to GA synthesis intermediates and analogues in the corresponding engineered strains. The strain information for the endogenous biosynthesis of the products is listed in Table S2. Their parent strains *S. cerevisiae* INVSc1 or SynV with the P_{PDR5} biosensor are tested as controls. The strains were fermented in 20 mL of YPD medium with designated condition at 200 rpm for 120 h, and the response of P_{PDR5} biosensor was detected. Three biological replicates were performed and data are presented as mean values ± SD. Source data for this figure is available in the Source Data file.



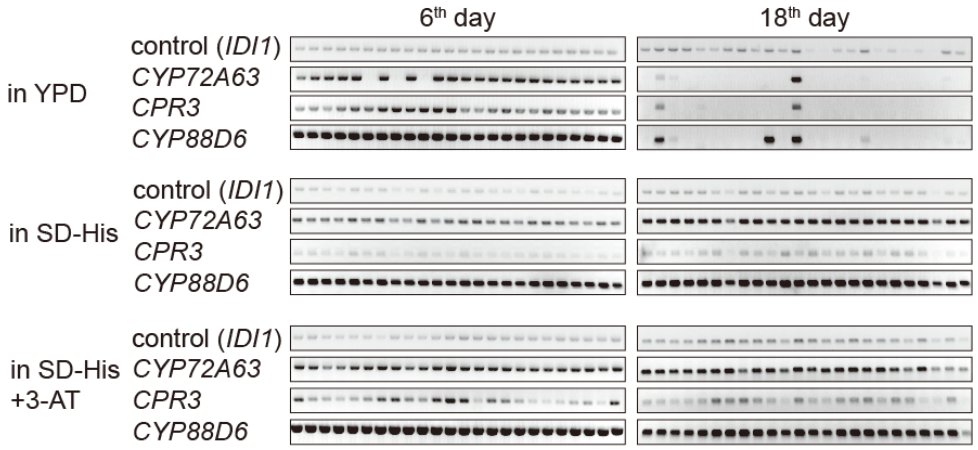
Supplementary Fig. 3 Exploration of the mechanism by which transcription factor Pdr1 regulates *PDR5*. (a) The simulation of the GA-Pdr1 complex. AlphaFold3 was employed to predict the Pdr1 homodimer structure and molecular docking with GA was performed using AutoDock Vina. The binding energies between GA and Pdr1 are -9.4 and -9.3 kcal/mol, respectively. The hydrogen bond interactions between GA and the Pdr1 homodimer were further analyzed. Notably, the hydrogen bond distance between GA and the residue Ser-485 was measured at 2.8 Å, highlighting the strong binding interactions. These findings indicate that the GA forms a stable complex with the Pdr1, facilitated by hydrogen bond interactions, which contribute to the overall stability of the docked structure. **(b)** Yeast two-hybrid monitoring of interaction proteins of Ssz1. Source data for this figure is available in the Source Data file.



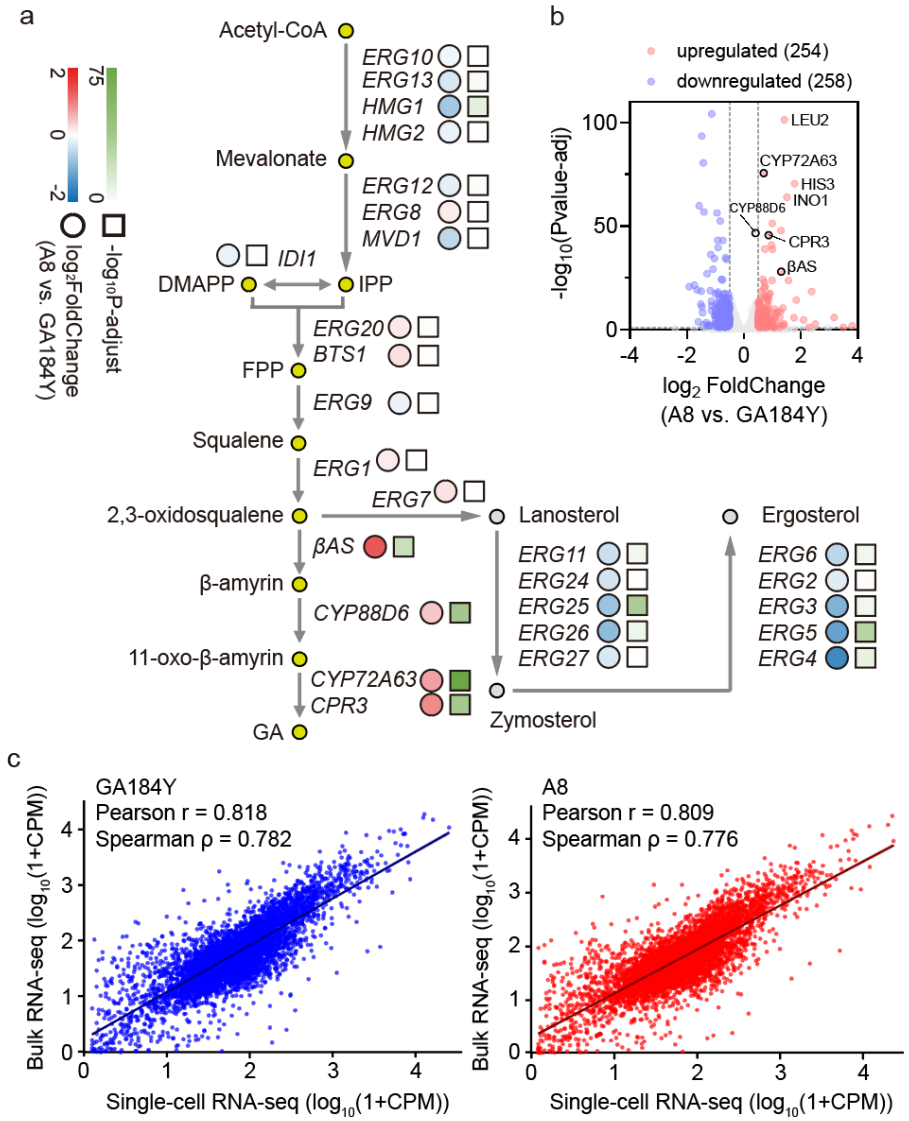
Supplementary Fig. 4 Effect of inserting transcriptional factor binding sites on the response to GA. (a) Pdr1/3. **(b)** Yap1. **(c)** Stb5. Three biological replicates were performed and data are presented as mean values ± SD. Source data for this figure is available in the Source Data file.



Supplementary Fig. 5 The addition of 3-AT exerted stress on the growth-production coupling system. (a) GA addition promoting the cell growth proved the effectiveness of the growth-production coupling system. **(b)** Effect of 3-AT concentration on cell growth. Source data for this figure is available in the Source Data file.



Supplementary Fig. 6 Detection of the genes in the synthetic pathway of the evolutionary strains by colony PCR. The yeast culture was spread onto the solid culture medium, after which the genes of individual colonies were identified using PCR. The quantified results of the gene lost rates were shown in Fig. 3e. Source data for this figure is available in the Source Data file.

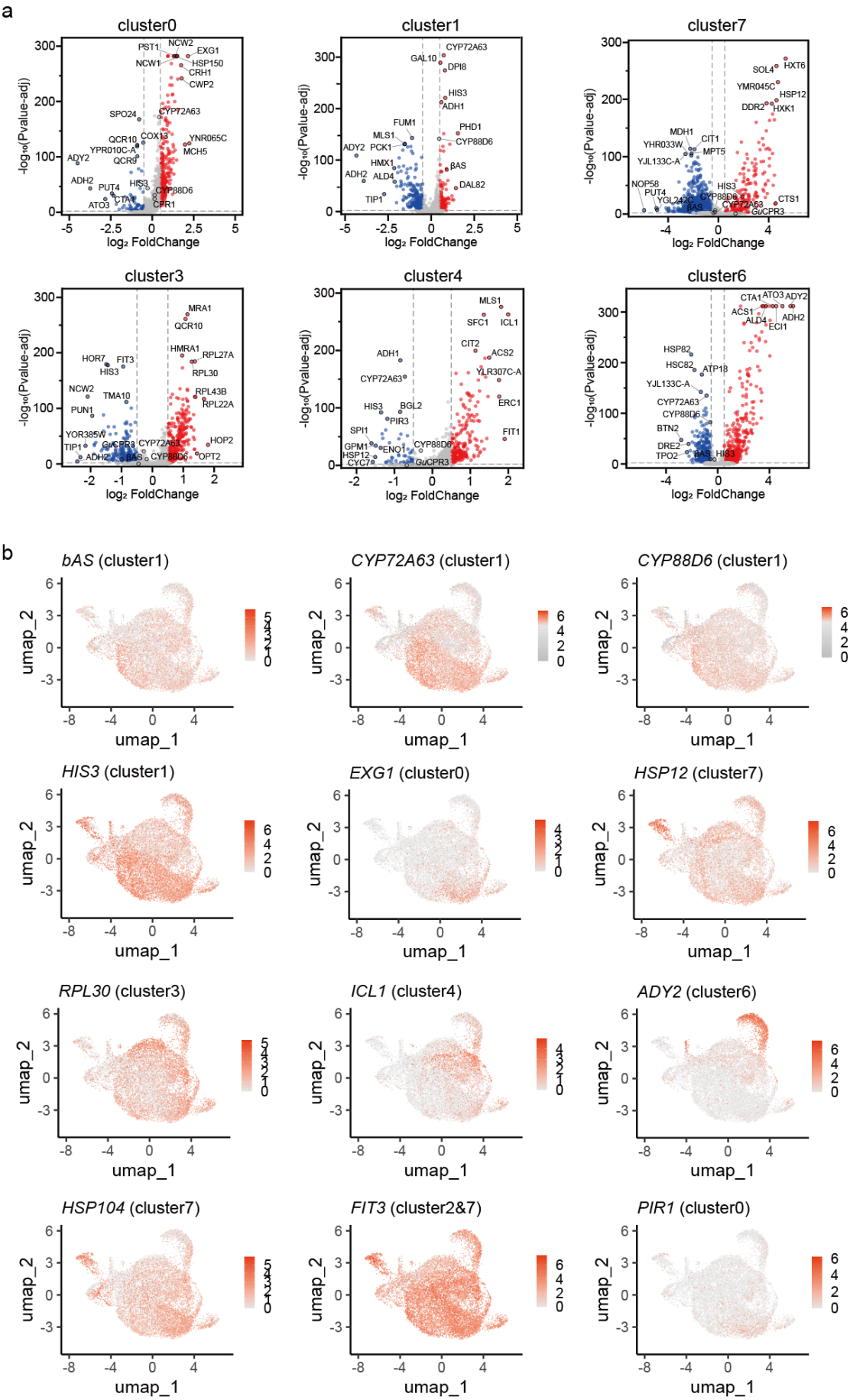


Supplementary Fig. 7 The transcriptomic analysis of the evolved strain and the control strain. (a) Expression difference of genes involving GA synthesis pathways. **(b)** Volcano map showed the significant differential expression genes. The control strain, named GA184Y, was obtained from the strain GA184-PE23 evolved in YPD medium at 6th day. Subsequently, the evolved strain A8 and the control strain GA184Y were cultivated in GM1 media for a period of 5 days. Thereafter, the cells were collected and subjected to RNA-seq. These two strains represent the biosensor-dependent evolved high-production strain and the normal passaged strain, respectively, utilized to explore the gene expression differences of high-production strain. Three

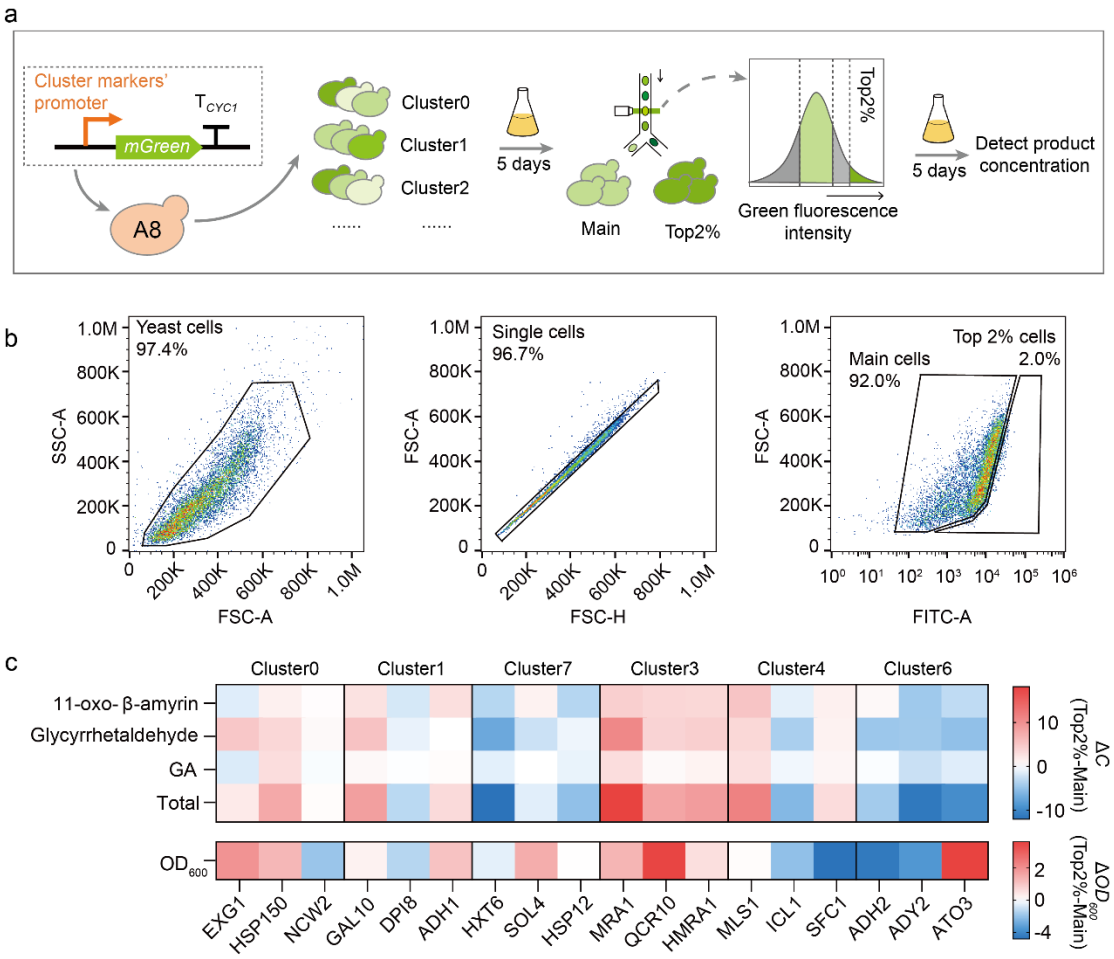
biological replicates were performed. Statistical significance was determined using two-sided Student's *t*-test. P values were adjusted for multiple comparisons using the Benjamini-Hochberg procedure to control the false discovery rate (FDR). Source data for this figure is available in the Source Data file. **(c)** Relationship between single-cell RNA-seq and bulk RNA-seq. For single-cell RNA-seq data, raw count matrices were aggregated by summing the counts of all cells within each condition to generate pseudo-bulk profiles. Gene expression was normalized to counts per million (*CPM*), calculated as:

$$CPM_i = \frac{raw\ counts_i}{\sum_j raw\ counts_j} \times 10^6 \quad (1)$$

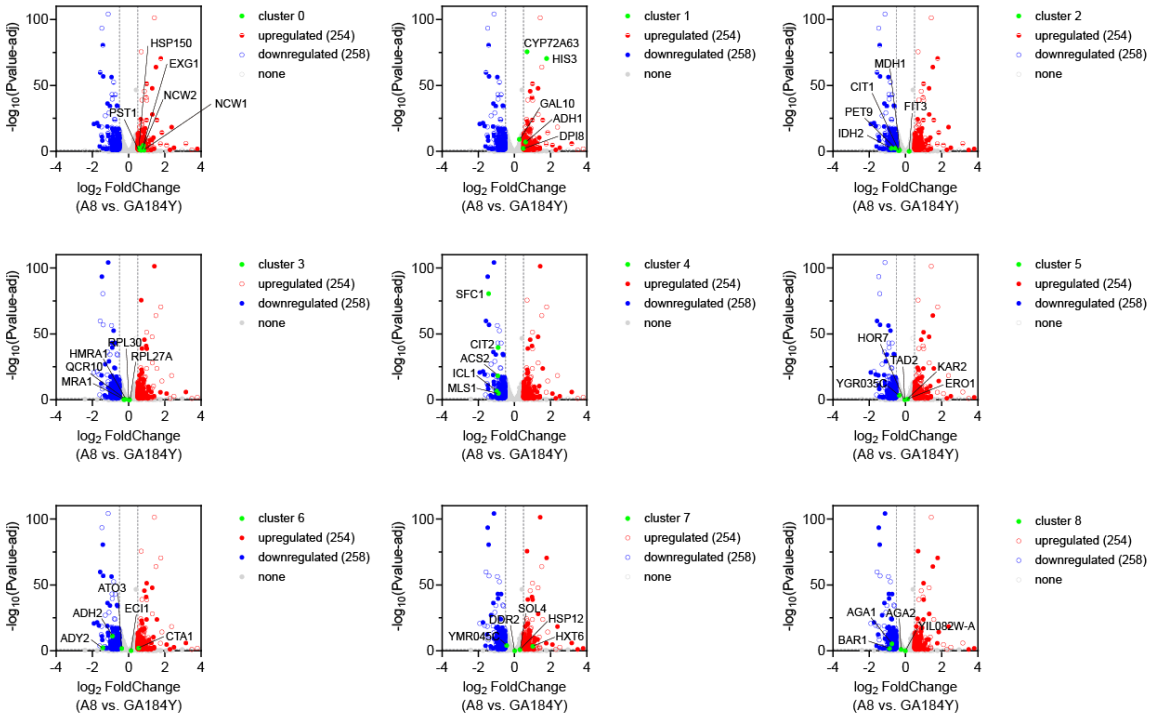
where *i* denotes a given gene and *j* indexes all genes. To stabilize variance and reduce the impact of highly expressed genes, *CPM* values were further transformed as $log_{10}(1 + CPM)$. Correlation between pseudo-bulk and bulk RNA-seq expression values was quantified using both Pearson's correlation coefficient (r) and Spearman's rank correlation coefficient (ρ), based on the log-transformed *CPM* values. All analyses were performed in R (v4.3) using the packages Seurat, Matrix, dplyr and ggplot2.


a
cluster0
cluster1
cluster7
cluster3
cluster4
cluster6
-log10(Pvalue-adj)
log2 FoldChange
b
bAS (cluster1)
CYP72A63 (cluster1)
CYP88D6 (cluster1)
HIS3 (cluster1)
EXG1 (cluster0)
HSP12 (cluster7)
RPL30 (cluster3)
ICL1 (cluster4)
ADY2 (cluster6)
HSP104 (cluster7)
FIT3 (cluster2&7)
PIR1 (cluster0)
umap_1
umap_2

Supplementary Fig. 8 Marker gene analysis in single-cell transcriptome. (a) Significant differential genes between the concerned clusters and the entire cells. Source data for this figure is available in the Source Data file. **(b)** The distribution of marker genes of interest in UMAP, including synthetic pathway genes (*CYP72A63*, *CYP88D6*, and *βAS*), aging marker genes (*HSP104*, *FIT3*, and *PIR1*), and hypervariable genes in each cluster (*EXG1*, *HIS3*, *HSP12*, *RPL30*, *ICL1*, and *ADY2*).



Supplementary Fig. 9 Validation of the production in cells with high expression of marker genes. (a) Sorting process of cells with high expression of marker genes. **(b)** Gating strategy for yeast flow cytometry analysis and cell sorting. **(c)** The production of cells with high expression of marker genes in the main differential clusters. The selection of marker genes was based on the top-3 genes of each cluster in Fig. 4g, excluding the gene directly involved in the GA synthesis. Strain A8 was used as the initial strain. Three biological replicates were performed and data are presented as mean values. Source data for this figure is available in the Source Data file.



Supplementary Fig. 10 The top-5 genes of each cluster in single-cell RNA-seq are shown in the bulk RNA-seq volcano map. The green dots represent the top-5 genes of each cluster. It is evident that a significant proportion of top-5 genes are not significant in bulk RNA-seq, thus emphasizing the necessity of single-cell RNA-seq to identify differences in cell differentiation subgroups. Source data for this figure is available in the Source Data file.

Supplementary Table 1 Selected genes encoding transporters or hypothetical proteins with upregulated expression

No.	Gene	Protein function	log_2FoldChange (GA180/SynV)
1	*YLR154C-G*	Hypothetical protein	3.20
2	*MPC2*	Integral membrane protein of the mitochondrial inner membrane with pyruvate transporter activity	1.75
3	*PUN1*	Plasma membrane protein with a role in cell wall integrity	0.82
4	*YNK1*	Nucleoside diphosphate kinase involved in DNA damage response	1.48
5	*TIM17*	Subunit of TIM23 mitochondrial import inner membrane translocase complex that typically transports proteins that possess a matrix-targeting N-terminal presequence	0.97
6	*LGE1*	Nuclear protein involved in histone methylation, histone ubiquitination, and premeiotic DNA replication	1.35
7	*YJR085C*	Hypothetical protein	1.33
8	*REE1*	Cytoplasmic protein involved in the regulation of enolase	1.41
9	*PDR5*	Plasma membrane ATP-binding cassette (ABC) transporter involved in resistance to multiple drugs	3.39
10	*FIT3*	Protein involved in siderophore transport	1.02
11	*YLR346C*	Hypothetical protein	4.05
12	*FIT2*	Protein involved in siderophore transport	1.92

Supplementary Table 2 The strains and plasmids used in this study

Strain or plasmid	Genotype	Source
pRS415	CEN/ARS, *leu2*, *Amp*	Lab
pRS415-P_{PDR5}	CEN/ARS, *BleR*, pRS415-insulator-P_{PDR5}-*mRFP*-T_{CYC1}	This study
pRS415-PE23	CEN/ARS, *LEU2*, pRS415-($P_{PDR5\text{-PE23}}$)-*HIS3*-T_{CYC1}	This study
SynV	*S. cerevisiae* BY4741, *MATa his3*Δ1*, leu2*Δ0*, met15*Δ0*, ura3*Δ0, synthetic chromosome V	[45]
GA180 (producing GA)	*S. cerevisiae* SynV, *HO*::P_{GAL10}-*βAS*-T_{CYC1}-P_{GAL2}-*CYP72A63*^{T338S}-T_{ADH1}-P_{GAL7}-*unigene25647*-T_{TYS1}-P_{PGK1}-*GuCPR3*-T_{PGK1}, Δ*GAL80*	[46]
SynV-PDR5	SynV, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	This study
GA180-PDR5	GA180, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	This study
SynV-pRS415-PDR5	SynV, CEN/ARS, *BleR,* pRS415-insulator-P_{PDR5}-*mRFP*-T_{CYC1}	This study
GA180-pRS415-PDR5	GA180, pRS415-insulator-P_{PDR5}-*mRFP*-T_{CYC1}	This study
SynV-PE23	SynV, *NRT1*::insulator-$P_{PDR5\text{-PE23}}$-*mRFP*-T_{CYC1}-*BleR*	This study
GA180-PE23	GA180, *NRT1*::insulator-$P_{PDR5\text{-PE23}}$-*mRFP*-T_{CYC1}-*BleR*	This study
GA184 (producing GA)	GA180, Δ*LEU2*::P_{GAL2}-*NHMGR-NAT*	[47]
GA184-PE23	GA184, pRS415-PE23	This study
aAM11 (producing α-amyrin)	*S. cerevisiae* INVSc1, *rDNA*::P_{TDH3}-*tHMG1*-P_{ALA1}-*ERG20*-P_{GPM1}-*ERG9*-P_{TYS1}-*ERG1*-*HIS3*, *Delta*:: P_{GAL1}-*MdOSC1*$^{N11T/P250H/P373A}$-P_{GAL10}-*MdOSC1*$^{N11T/P250H/P373A}$-P_{ADH1}-*DGA1*-T_{DGA1}-*URA3*	[54]
Sgib (producing β-amyrin)	*S. cerevisiae* INVSc1, *rDNA*::P_{ADH1}-*IDI1*-T_{ADH1}-P_{ALA1}-*ERG20*-T_{ALA1}-P_{GPM1}-*ERG9*-T_{GPM1}-P_{TYS1}-*ERG1*-T_{TYS1}-P_{FBA1}-*bAS*-T_{FBA1}-*hphNT1*, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	[55]
11O-99 (producing 11-O-β-amyrin)	*S. cerevisiae* SynV, *HO*::P_{GAL10}-*βAS*-T_{CYC1}-P_{GAL7}-*unigene25647*-T_{TYS1}-P_{PGK1}-*GuCPR3*-T_{PGK1}, Δ*GAL80*, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	This study
GOH-1	*S. cerevisiae* SynV, *HO*::P_{GAL10}-*βAS*-T_{CYC1}-	[46]

Strain or plasmid	Genotype	Source
(producing glycyrrhetol)	P_{GAL2}-*CYP72A63*^{L509I}-T_{ADH1}-P_{GAL7}-*unigene25647*-T_{TYS1}-P_{PGK1}-*GuCPR3*-T_{PGK1}, Δ*GAL80*, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	
GL-P01 (producing glycyrrhizin)	*S. cerevisiae* GA180, *rDNA*::P_{TDH3}-*tHMG1*-T_{ECM10}-P_{TEF1}-*ERG20*-T_{AIP1}-P_{HTA2}-*ERG9*-T_{YHI9}-P_{CHO1}-*ERG1*-T_{RPL41B}-P_{FBA1}-*IDI1*-T_{TDH1}- *hphNT1*, *YPRC3*::P_{GAL10}-*HsUGDH*-T_{TPI1}-P_{GAL2}-*GuCSyGT*-T_{ADH1}-P_{GAL7}-*UGT73P12*-T_{PGK1}-*LEU2*, *NRT1*::insulator-P_{PDR5}-*mRFP*-T_{CYC1}-*BleR*	[36]
G04 (producing medicarpin)	*S. cerevisiae* BY4742, *int1*::P_{TPI1}-*GuCPR3*-T_{CYC1}, *int16*::P_{SeGal2}-*CHI1*-T_{FBA1}-P_{SeGal2}-*CHR1*-T_{TEF2}-P_{SeGal2}-*CHS1*-T_{PGK1}, *int17*::$P_{SptGal2}$-*PAL1*-T_{ADH1}-$P_{SptGal2}$-*StsTAL*-T_{FBA1}-P_{SkGal1}-*C4H1*-T_{TEF2}-$P_{SkGal10}$-*4CL1*-T_{PGK1}, *int5*::P_{SmGal2}-*IFS1*-T_{ADH1}-P_{SbGal2}-*HID1*-T_{CYC1}-P_{SmGal2}-*OMT1*-T_{FBA1}, *int21*::$P_{SpdGal2}$-*IFR1*-T_{TEF2}-P_{ScGal2}-*I2H1*-T_{PGK1}-P_{ScGal2}-*VR1*-T_{ADH1}-P_{SeGal2}-*PTS1*-T_{FBA1}	[56]