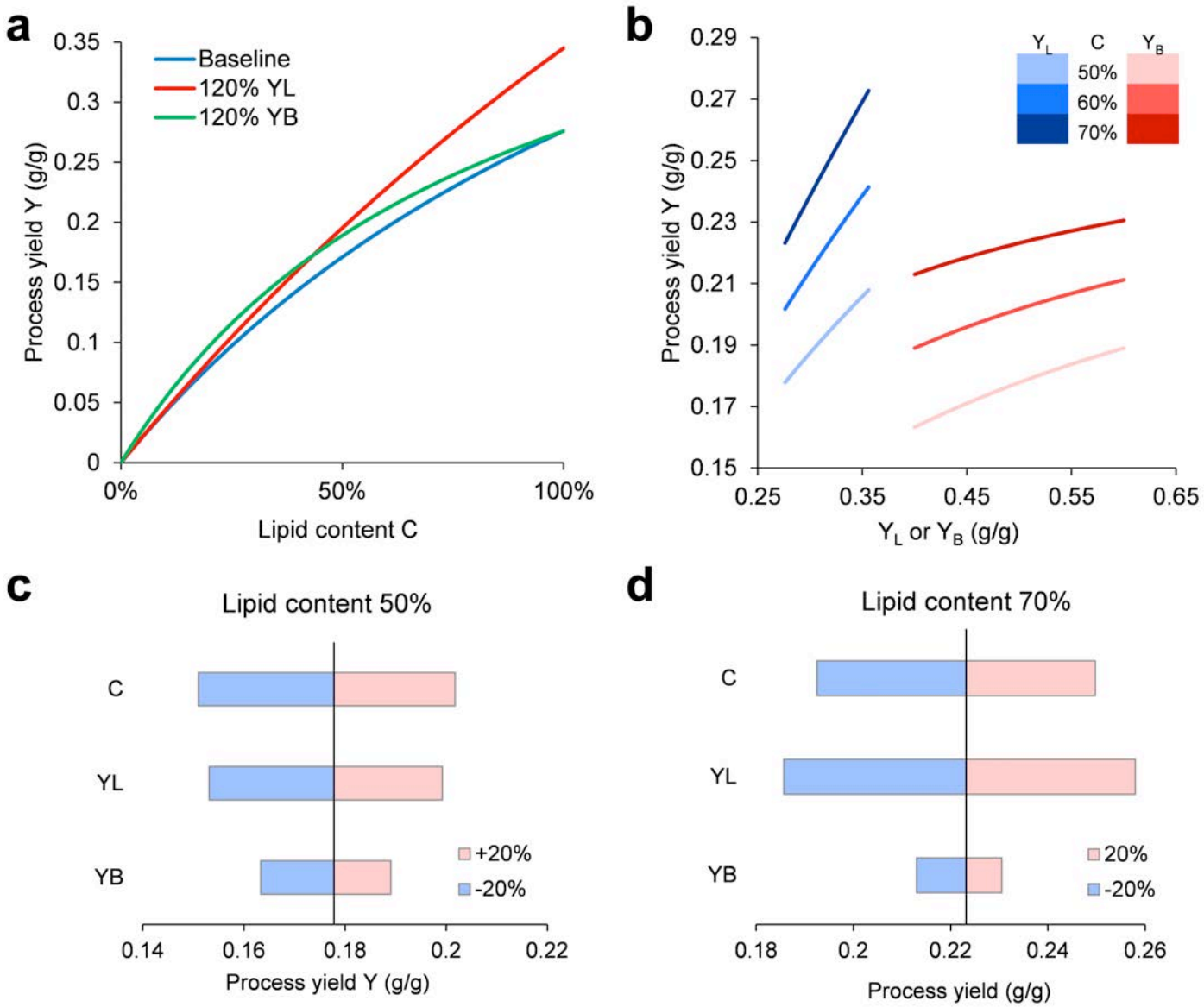
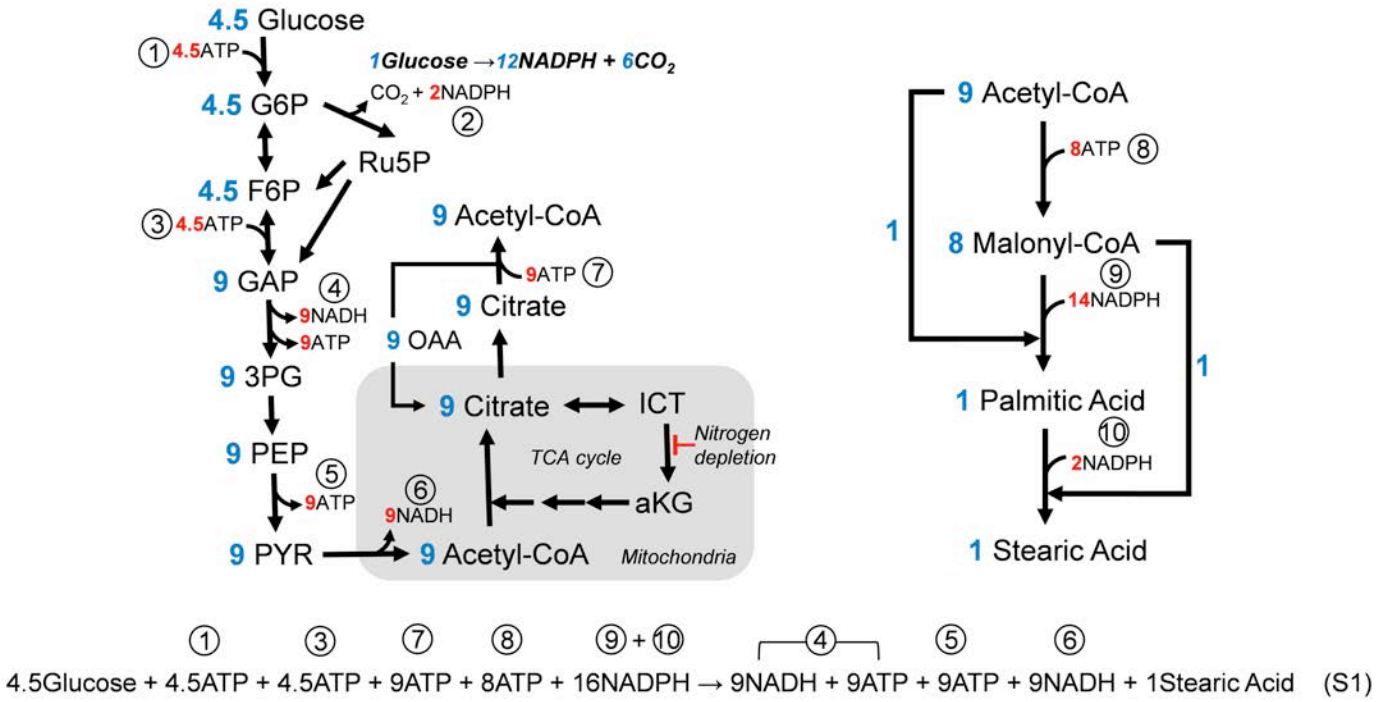


Supplementary Figure 1

In silico mathematical optimization of Y by varying the three parameters- Y_B, Y_L and C.

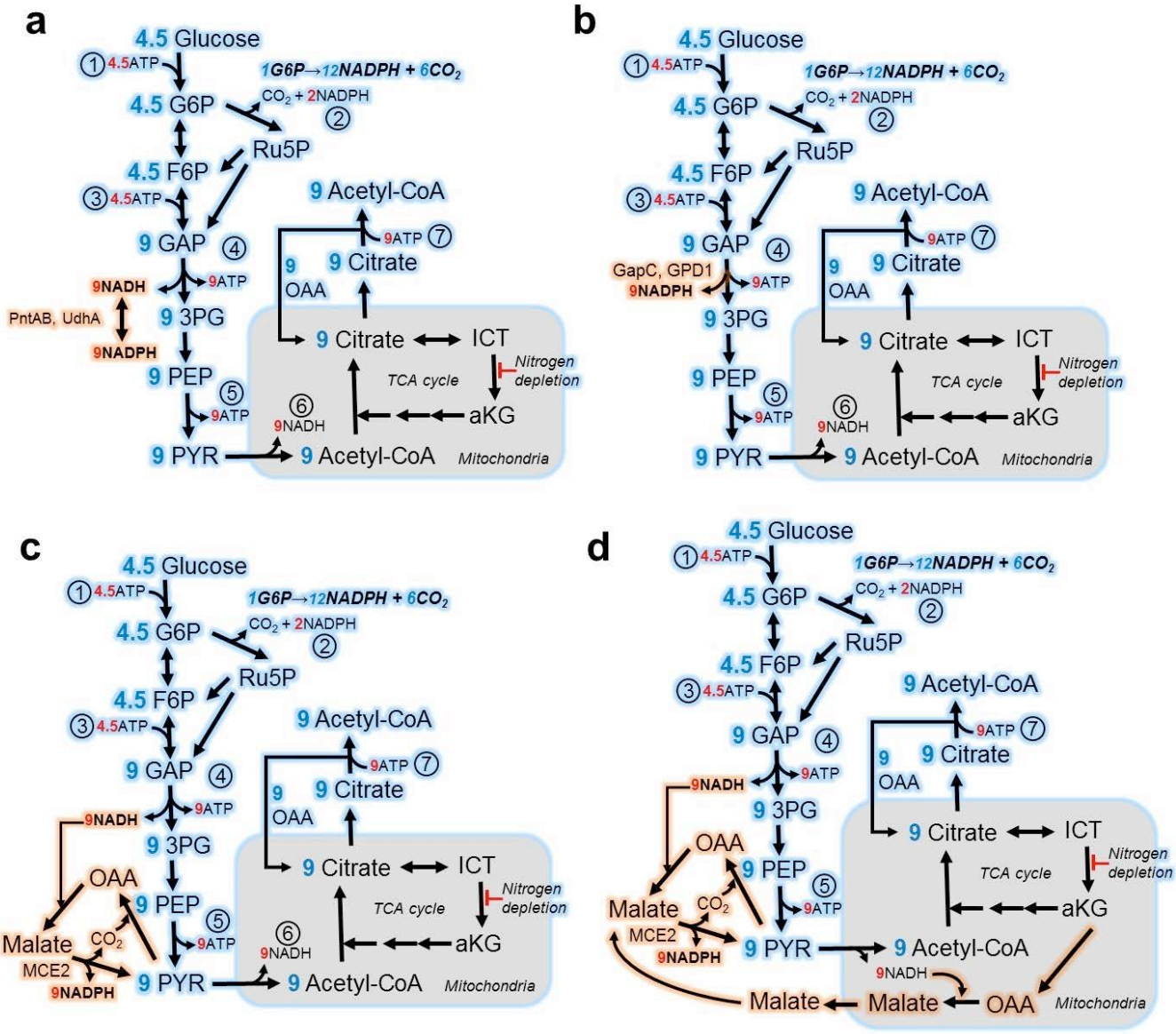
(**a**) Process yield Y as function of lipid content C for three sets of Y_L and Y_B: baseline (Y_L, 0.271 g/g, YB, 0.55 g/g), highlighted in blue; 120% Y_L (Y_L, 0.325 g/g, YB, 0.55 g/g), highlighted in red; 120% Y_B (Y_L, 0.271 g/g, YB, 0.66 g/g), highlighted in green. (**b**) Process yield as a function of the lipid yield Y_L (blue colors) and non-lipid biomass yield Y_B (red colors) for three different lipid contents C = 50%, 60% and 70% (as indicated by the color intensity of the heat map). Single-point sensitivity tornado charts for (**c**) process yield Y with C at the baseline of 50% and (d) process yield Y with C at its baseline of 70% (see **Supplementary Note 3** for detailed analysis).



Supplementary Figure 2

Pathways and overall stoichiometry of *de novo* fatty acid biosynthesis from glucose in *Y. lipolytica* under nitrogen starvation conditions.

(**a**) Central carbon metabolic pathways involved in supplying cytosolic acetyl-CoA (EMP, part of TCA cycle) and NADPH (pentose phosphate pathway) from glucose in support of lipogenesis. Different from S. cerevisiae, with sugar carbon sources cytosolic acetyl-CoA is exclusively synthesized from citrate in a reaction catalyzed by Acetyl-CoA Lyase (ACL). Under nitrogen starvation conditions, metabolic flux is interrupted at the citrate node due to inactivation of isocitrate dehydrogenase. (**b**) Biosynthesis of fatty acids in Y. lipolytica, as exemplified using stearic acid. Formation of SA is an energy intensive process, requiring 1 acetyl-CoA, 8 malonyl-CoA and 16 NADPH to afford 1 SA (See Supplementary Note 1 for details). The stoichiometric number of each metabolite (e. g. glucose) and cofactor (e. g. NADH) are respectively highlighted in blue and red color.

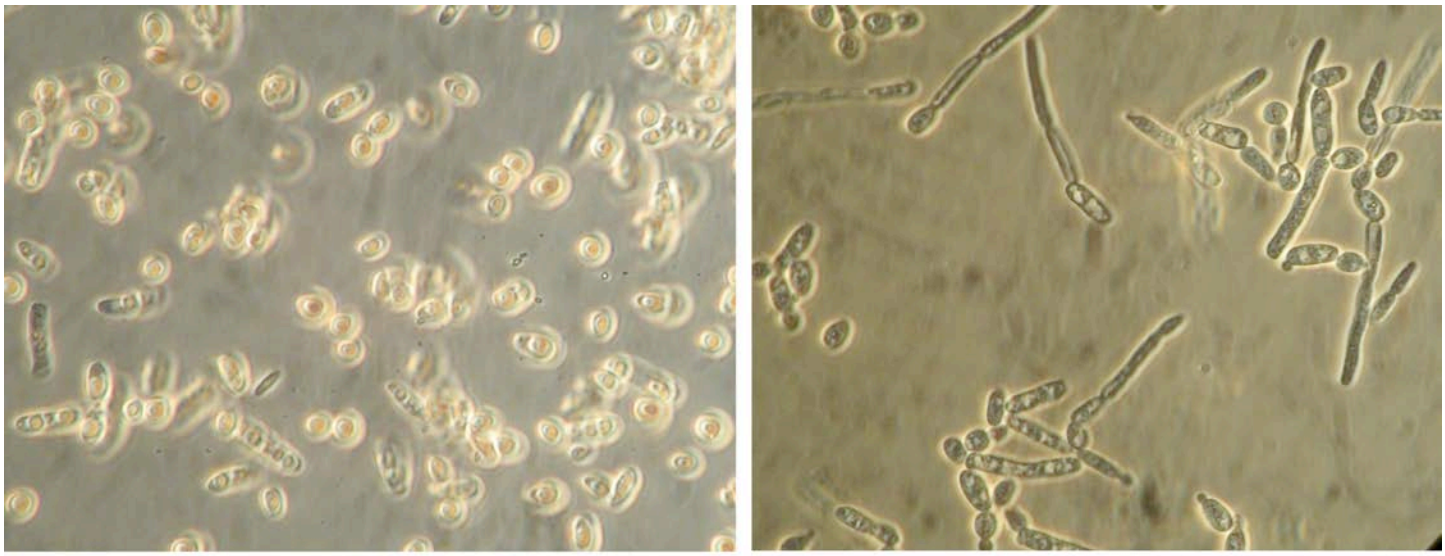


Supplementary Figure 3

Central carbon pathways and engineered synthetic networks used to trade cytosolic NADH with NADPH.

Three strategies are presented that would allow, in the cytosol of yeast cell, (**a**) recycling NADPH directly through heterologous expression of *E. coli* transhydrogenases PntAB and/or UdhA, (**b**) generation of NADPH via introduction of NADP+-dependent glyceraldehyde-3-phosphate dehydrogenases, and, (**c**) regeneration of NADPH by the activation of pyruvate-oxaloacetate-malate (POM) cycle (see **Supplementary Note 2** for detailed discussions).

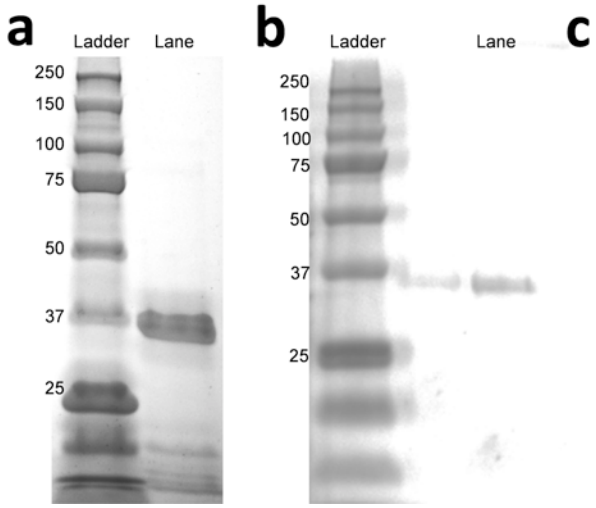
Nature Biotechnology: doi:10.1038/nbt.3763



Supplementary Figure 4

Microscopic images of engineered *Y. lipolytica* strain AD and strain ADpntAB expressing the two *E. coli* membrane-bound proteins PntA and PntB.

Both strains were cultured in shake flasks containing fermentation medium for 72 h before sampled for imaging. Clearly, ADpntAB cells exhibited elongated or filamentous morphology, while most AD cells remain spherical with identifiable lipid bodies (bright) at the center of individual cells.

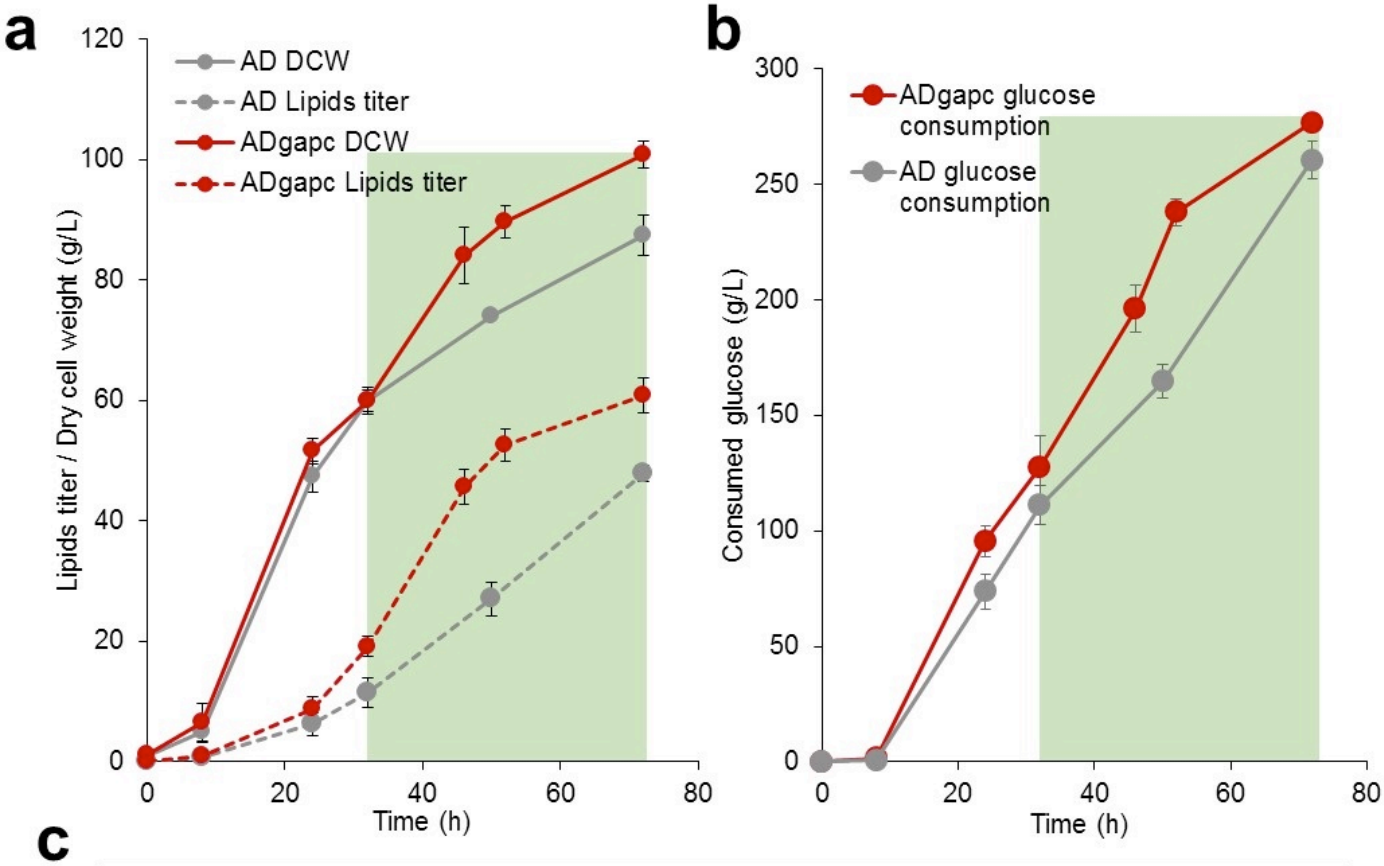


Enzyme	Cofactor	K_m (mM)	K_{cat} (min^{-1})
ylGPD	NAD^+	0.0505	416.60
	$NADP^+$	0.2583	259.82
GapC	NAD^+	0.0568	376.64
	$NADP^+$	0.0278	392.15

Supplementary Figure 5

Purification and in vitro characterizations of GAP dehydrogenases – ylGPD and GapC.

(**a**) 12% SDS-PAGE gel of C-terminal hexahistidine tagged ylGPD purified from *Y. lipolytica* strain YL-GPD (35.3 kDa) using Ni-NTA chromatography. (**b**) 12% SDS-PAGE gel of C-terminal hexahistidine tagged GapC (35.8 kDa) purified from *Y. lipolytica* strain YL-GPD using Ni-NTA chromatography. (**c**) Enzyme kinetics of ylGPD and GapC. The enzyme purification and assay conditions are available in Methods section.



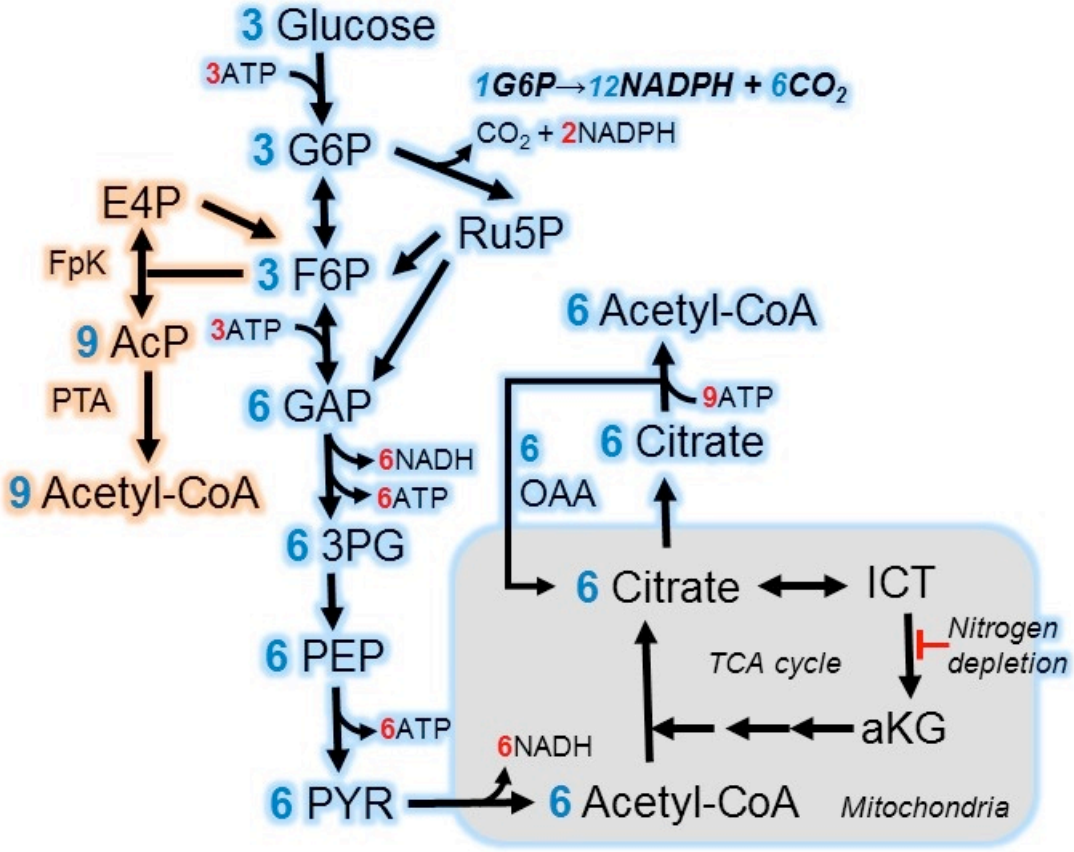
c

	AD	ADgapc
Titer (g/L)	47.8 ± 1.4	63.2 ± 2.4
Lipid content	54.7% ± 1.0%	62.4% ± 2.0%
Process yield (g-FAME/g-Glucose)	0.184 ± 0.003	0.229 ± 0.007
Maximum yield of FAME (g-FAME/g-Glucose)	0.244 ± 0.009	0.291 ± 0.005
Maximum Yield of total fatty acids (g FA/g Glucose)	0.234 ± 0.009	0.279 ± 0.005

Supplementary Figure 6

Time course fermentation profiles of the strains AD and ADgapC.

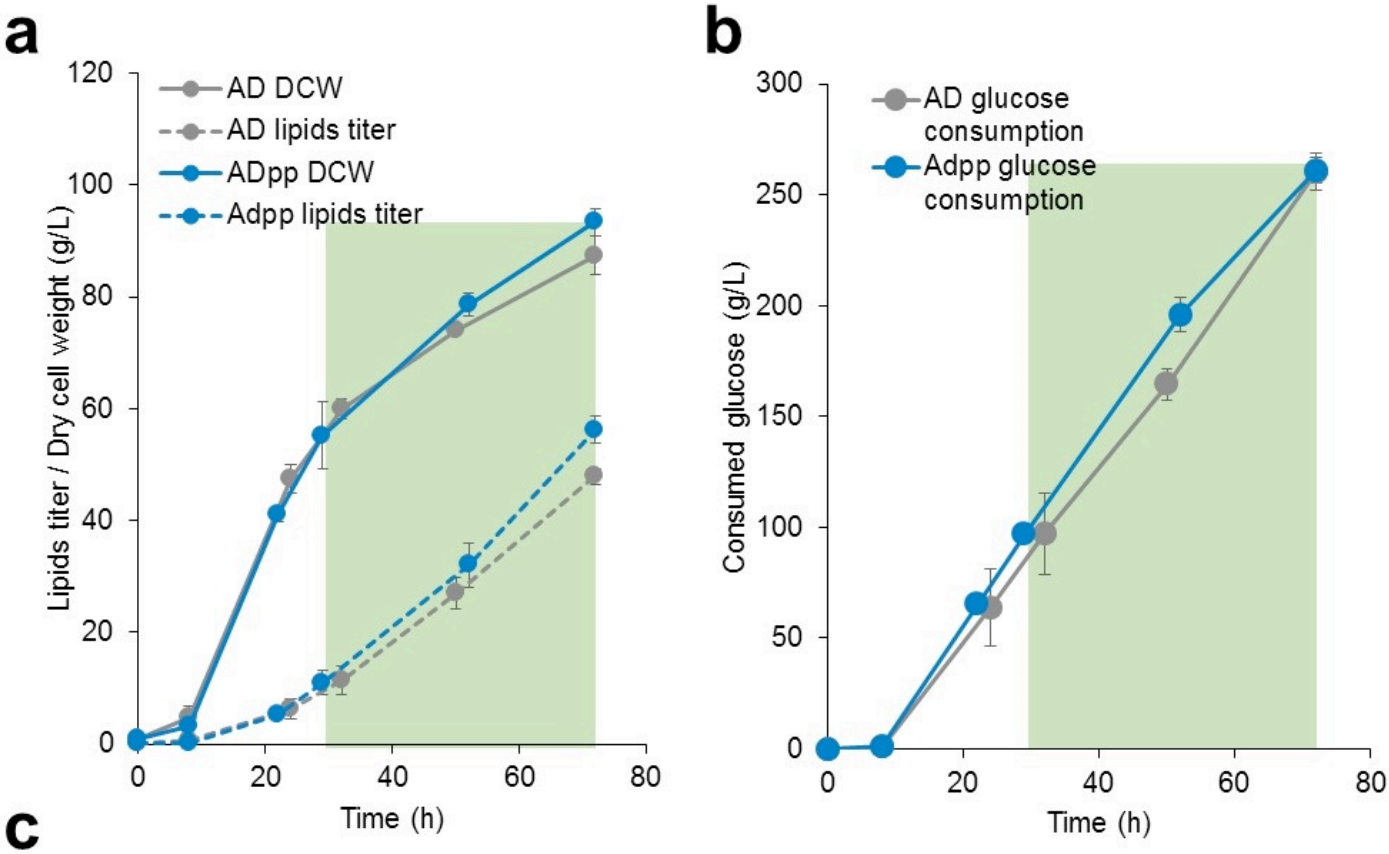
Time course profile of (**a**) cell growth and lipids production and (**b**) glucose consumption of AD and ADgapc in fed-batch fermentations. The fermentation characteristics including lipids titer, lipid content, process yield and maximum FAME and FA yield are shown in (**c**). The maximum yield is the yield calculated by dividing the lipids made during lipid production phase (highlighted in green in **a**) by the glucose consumed in the same phase (highlighted in green in **b**). The maximum FA yield of ADgapc is higher than the stoichiometric maximum yield obtained in wild type *Y. lipolytica*. Error bars mean ± s. d., n = 2, biological replicates.



Supplementary Figure 7

Illustration of NOG pathway that branching out from EMP pathway from F6P (see **Supplementary Note 2** for detailed discussions).

Not Applicable



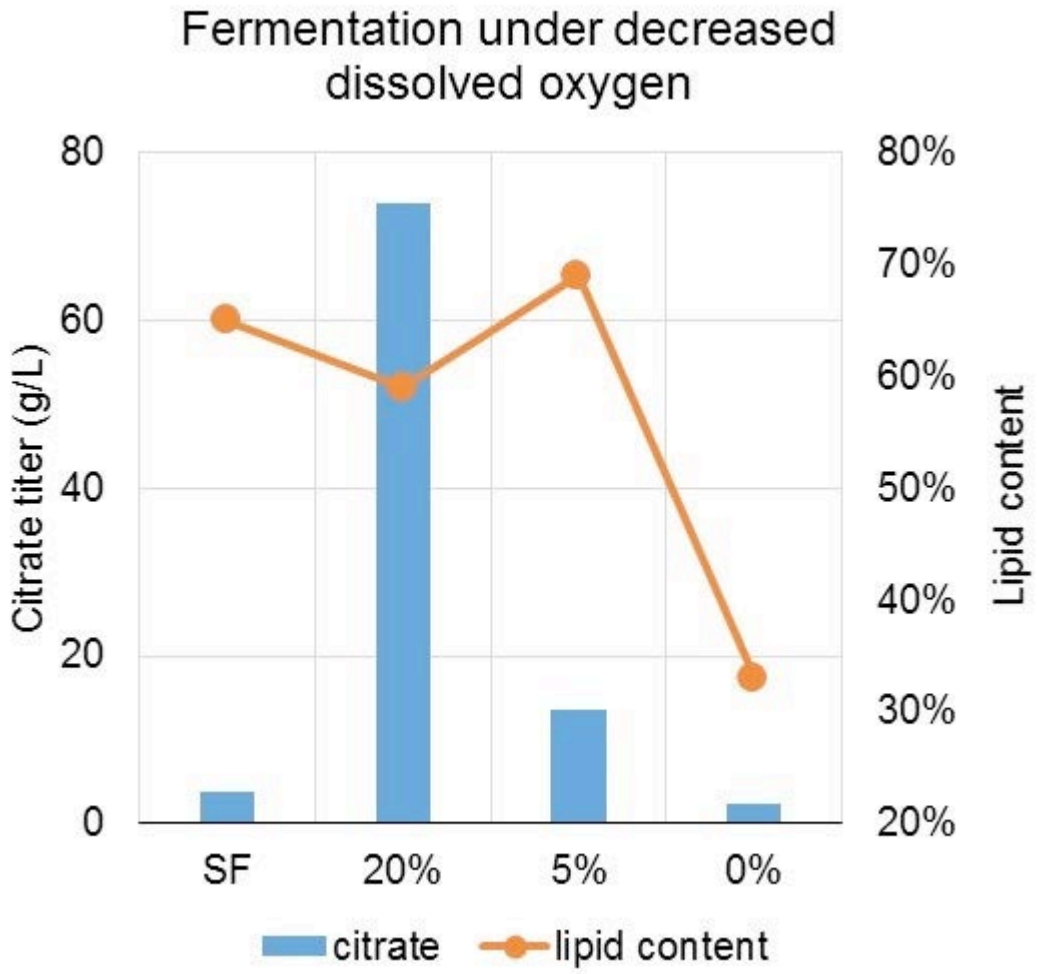
c

	AD	ADpp
Titer (g/L)	47.8 ± 1.4	56.2 ± 0.7
Lipid content	54.7% ± 1.0%	60.2% ± 2.3%
Process yield (g-FAME/g-Glucose)	0.184 ± 0.003	0.216 ± 0.003
Maximum yield of FAME (g-FAME/g-Glucose)	0.244 ± 0.009	0.276 ± 0.007
Maximum Yield of total fatty acids (g FA/g Glucose)	0.234 ± 0.009	0.261 ± 0.007

Supplementary Figure 8

Time course fermentation profiles of the strains AD and ADpp.

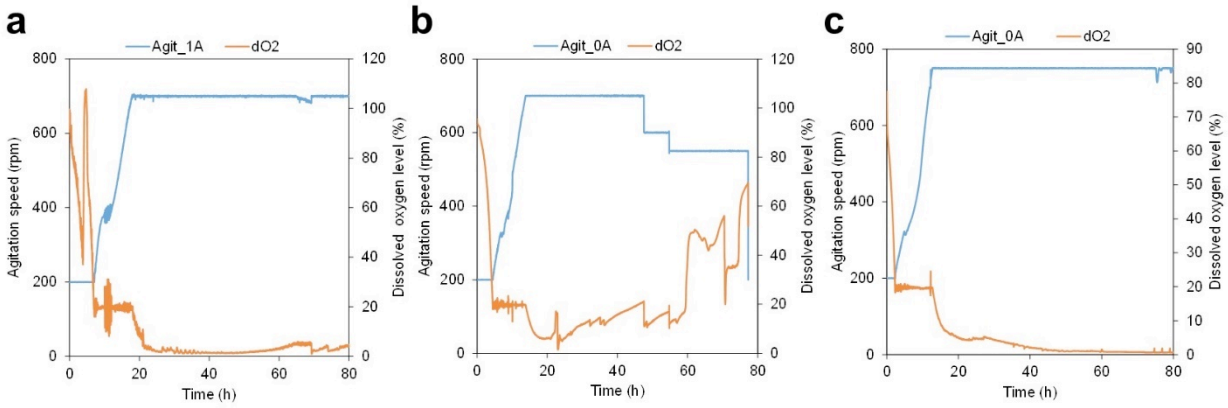
Time course profile of (**a**) cell growth and lipids production and (**b**) glucose consumption of AD and ADpp in fed-batch fermentations. The fermentation characteristics including lipids titer, lipid content, process yield and maximum FAME and FA yield are shown in (**c**). The maximum yields are calculated by dividing the lipids synthesized during lipid production phase (highlighted in green in **a**) by the glucose consumed in the same phase (highlighted in green in **b**). Error bars mean ± s. d., n = 2, biological replicates.



Supplementary Figure 9

The accumulation of lipids and citrate of strain ADgm in the presence of different dissolved oxygen levels during lipid production phase.

The cells are harvested at the end of the fermentation in shake flask (SF) culturing (144 h) and 1.5 liter bioreactor (72 h). Under low dissolved oxygen levels, including SF and 5% in bioreactor operation, optimum production of lipids is observed as indicated by the high lipids contents (above 60%). However, the high oxygen level (20% or above) completely induce citrate production while to some extent compromising lipids accumulation, while the lipids production ceased when cells are incubated at anaerobic conditions (0% or below). Morphological changes (elongated cell shapes or hyphae formation) of *Y. lipolytica* were observed at 0% oxygen level.

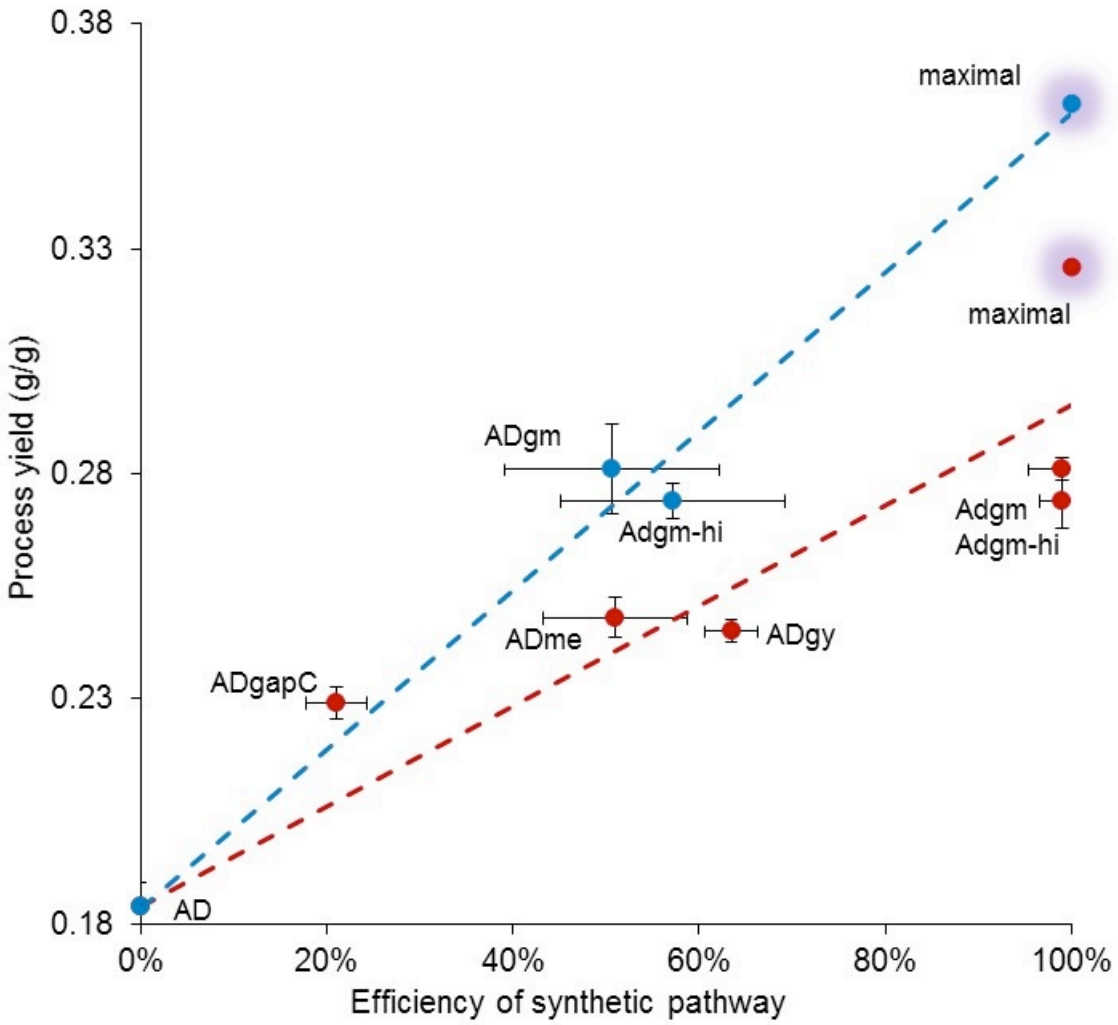


Supplementary Figure 10

Time course of agitation (Agit_0A) and dissolved oxygen (dO2) of fed-batch fermentations of AD, ADgm and ADgm under high nitrogen condition.

The maximum agitation speed is 750 rpm and the aeration rate is fixed to 5 vlm. As illustrated in (**b**), the dO2 level is gradually increased after 20 h. The spikes in dO2 level led to accumulation of byproduct citrate. In comparison, the fermentation of AD features no increase of dO_2 level throughout the first 60 hours as shown in (**a**). Double the initial concentration of ammonium in the starting fermentation medium increased the cell number and recreate the micro-aerobic conditions that prohibit citrate production without significantly compromising the lipids biosynthesis.

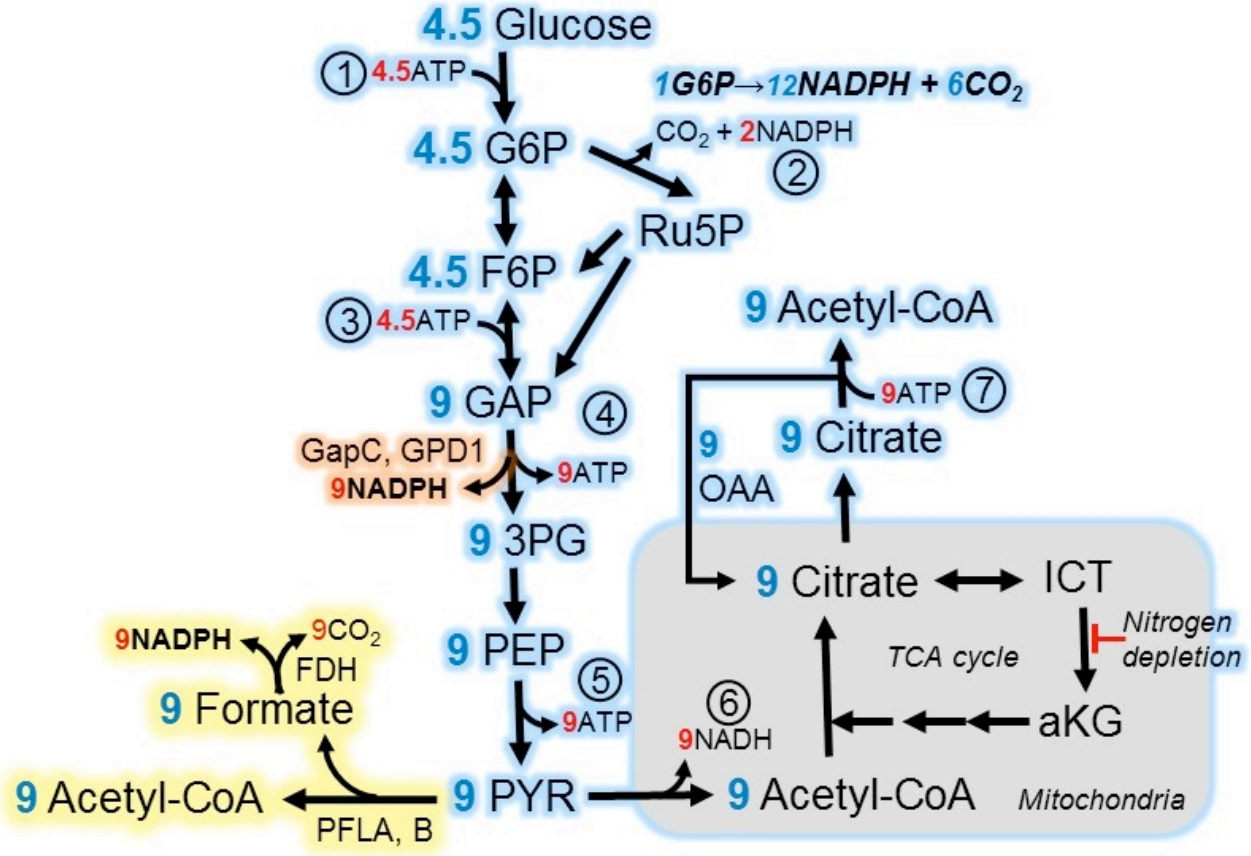
Nature Biotechnology: doi:10.1038/nbt.3763



Supplementary Figure 11

The process yields of lipids from different engineered *Y. lipolytica* strains are positively correlated to efficiencies of the introduced synthetic pathways.

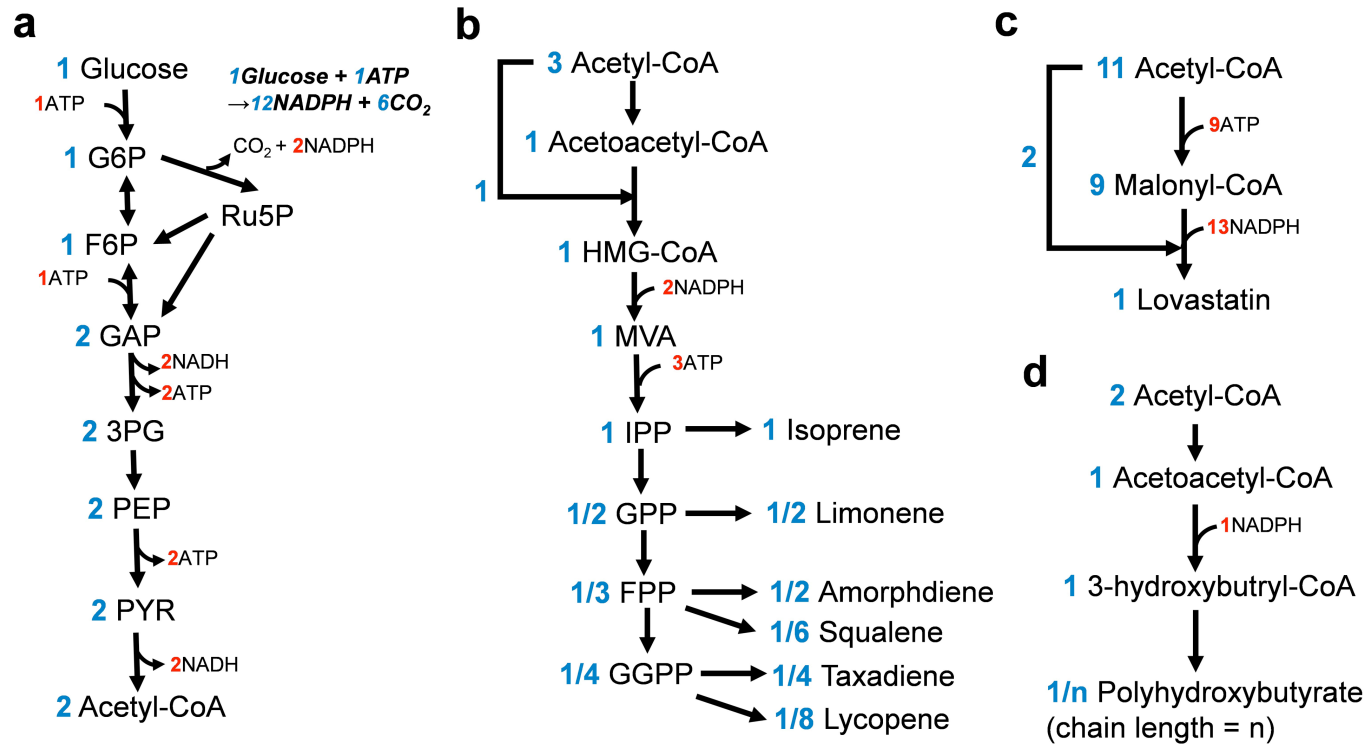
Measured process yields for different microbial constructs are plotted against their individual pathway efficiency when Y_L = 0.311 g/g. Only cytosolic NADH conversion to NADPH is shown ion red. The trend-line (dash line, red) was added and the maximum yield highlighted in purple. When both NADH.c and NADH.m are recycled by POM cycle, the plot and its trend-line are shown in blue color with maximal highlighted in purple. Error bars mean ± s. d., n = 2.



Supplementary Figure 12

Intercepting pyruvate from entering into mitochondrion by introduction of the competitive synthetic pathway that converts 1 pyruvate to 1 acetyl-CoA, 1 CO2 and 1 NADPH.

The synthetic pathway was proposed to be enabled by expressing E. coli pyruvate-formate lyase PFLB, its cognate activating enzyme PFLA and the $NADP^+$-dependent formate dehydrogenase (see **Supplementary Note 2** for detailed discussion).



Supplementary Figure 13

Overall stoichiometry of metabolic pathways involving in the biosynthesis of representative molecules belonging to terpenoids, polyketides and polyhydroxyalkanoates.

Not Applicable

Supplementary Information for

Lipid production in *Yarrowia lipolytica* is maximized by engineering cytosolic redox metabolism

Kangjian Qiao, Thomas M. Wasylenko, Kang Zhou, Peng Xu and Gregory Stephanopoulos [*]

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

[*]Corresponding author. Mailing address: Department of Chemical Engineering, Massachusetts Institute of Technology, Room 56-469, 77 Massachusetts Avenue, Cambridge, MA 02139. Phone: (617) 253-4583. Fax: (617) 253-3122. E-mail: gregstep@mit.edu

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 1. Validation of the quantitative yield model using reported fermentation metrics of three model microorganisms - *E. coli*, *Y. lipolytica* and *R. toruloides*.

	Y_B (g/g)[a]	Y_L (g/g)	Content	Process Y (g/g)	reported yield (g/g)	discrepancy	error	Ref
E. coli	0.50	0.35	78.0%	0.292	0.280	0.0123	***4.4%***	[1]
E. coli	0.50	0.35	21.0%	0.097	0.092	0.0041	***4.4%***	[2]
Y. lipolytica	0.55	0.27	69.0%	0.221	0.223	-0.0026	***-1.2%***	[3]
Y. lipolytica	0.325	0.27	75.0%	0.212	0.205	0.0071	***3.4%***	[4]
Y. lipolytica	0.50	0.27	61.9%	0.203	0.195	0.0082	***4.2%***	[5]
Y. lipolytica	0.325	0.27	77.6%	0.222	0.232	-0.0109	***-4.7%***	[6]
Y. lipolytica	0.325	0.27	79.0%[b]	0.222	0.213	0.0088	***4.2%***	[6]
R. toruloides	0.59	0.27	67.5%	0.222	0.230	-0.0081	***1.2%***	[7]

[a]The reported Maximum biomass yield Y_B value is used if it is available. Otherwise, the maximum biomass yield Y_B on glucose for *E. coli* or *Y. lipolytica* is 0.50 g-dry cell weight/g-glucose in rich medium (1 x Yeast nitrogen base – amino acids supplemented with 2 g/L yeast extract) and 0.325 g/g for *Y. lipolytica* in minimal medium (1 x yeast nitrogen base + glucose). [b]. this content value is estimated from the time-course plot of fed-batch fermentation in the original publication.

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 2. Proteins in *Y. lipolytica* catalyzing the phosphorylation of NADH to form NADPH.

Y. lipolytica accession number	protein name	*S. cerevisiae* homolog	EC no.	% amino acid identity/similarity	Deduced function	Ref
YALI0E17963p	ylPOS5	POS5; YPL188W	2.7.1.86	47/60	Mitochondrial NADH kinase	[8]
YALI0E27874p	ylUTR1	UTR1; YJR049C	2.7.1.23	50/68	ATP-NADH kinase, cytosol	[9, 10]
YALI0E23991p	ylYEF1	YEF1 ; YEL041W	2.7.1.86	37/58	ATP-NADH kinase, cytosol	[10]

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 3. Model calculated efficiencies of synthetic pathways introduced to *Y. lipolytica.*

Strain	Batch	Cell growth	Glucose consumed	Lipids (FAMEs) produced[a]			Synthetic Pathway
		DCW (g/L)	(g/L)	Titer (g/L)	Content (%)	Yield (g/g)	Efficiency
AD	1	84.6	252.1	47.1	55.9	0.187	*n/a*
AD	2	90.2	268.5	48.5	53.8	0.181	*n/a*
ADgapc	1	100.8	270.4	60.9	60.4	0.225	*25.1%*
ADgapc	2	101.9	282.6	65.7	64.5	0.232	*18.4%*
ADme	1	97.3	247.8	62.5	64.2	0.252	*57.4%*
ADme	2	95.4	248.4	60.2	63.1	0.242	*42.9%*
ADpp	1	95.85	255.0	55.6	58.0	0.218	*6.3%*
ADpp	2	90.9	266.6	56.8	62.4	0.213	*-28.7%*
ADgy	1	85.8	216.4	52.5	61.2	0.243	*66.4%*
ADgy	2	86.7	222.5	54.5	62.9	0.245	*60.6%*
ADgm	1	94.4	232.4	65.5	69.4	0.279	*101.4% / 62.2%*
ADgm	2	87.4	242.5	68.1	77.9	0.284	*96.7% / 39.1%*
ADgm-hi	1	140.2	365.4	98.7	70.4	0.270	*98.0% / 45.1%*
ADgm-hi	2	148.8	356.3	99.3	66.7	0.279	*100.0% / 69.5%*

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 4. Metabolic costs[a] and cofactor balancing of bioproducts biosynthesis from glucose in yeasts.

Bioproduct[b]	Category	Precursor	NH_3	ATP	NADH	NADPH	Y_p	Y_p^R
Stearic acid	***Fatty acids derivatives***	Acetyl-CoA	*n/a*	1	18	-16	0.271	**0.311**
Eicosapentaenoic acid (EPA)		Acetyl-CoA	*n/a*	1	20	-23	0.243	**0.336**
Docosahexaenoic acid (DHA)		Acetyl-CoA	*n/a*	1	22	-26	0.238	**0.332**
Isoprene	***Terpenoids***	IPP	*n/a*	0	6	-2	0.227	**0.252**
Limonene		IPP x 2	*n/a*	0	12	-4	0.227	**0.252**
Amorpha-4,11-diene		IPP x 3	*n/a*	0	18	-6	0.227	**0.252**
Taxadiene		IPP x 4	*n/a*	0	24	-8	0.227	**0.252**
Squalene		IPP x 6	*n/a*	0	36	-13	0.226	**0.254**
Cholesterol		IPP x 6	*n/a*	-12	36	-26	0.192	**0.239**
Lycopene		IPP x 8	*n/a*	0	48	-16	0.224	**0.249**
Astaxanthin		IPP x 8	*n/a*	0	48	-18	0.246	**0.276**
Lovastatin	***Polyketides***	Acety-CoA	*n/a*	-4	22	-13	0.341	**0.409**
6-methylsalicylic acid (6-MSA)		Acetyl-CoA	*n/a*	-2	4	-1	0.405	**0.422**
Polyhydroxylbutryrate (PHB, n=1)[c]	***Polyesters***	Acetyl-CoA	*n/a*	2*	4*	-1*	0.533	**0.578**
Glycolate	***Organic Acids***	Acetyl-CoA	*n/a*	3	2.5	-1	0.734	**0.844**
3-hyroxypropionate		Acetyl-CoA	*n/a*	0	2	-2	0.712	**0.988**
Penicillin N	***Non-ribosomal peptides***	Cysteine Valine Aminoadipate	-3	-8	5	-7	0.486	**0.559**
Hydrocodone	***Alkaloids***	Tyrosine, Acetyl-CoA	-2	-13.6	8	-7	0.344	**0.391**
Tyrosine	***Amino Acids***	PEP, E4P	-1	-2.3	3	-2	0.524	**0.574**
Cysteine		3-PGA	-1	-4	2	-5	0.734	**0.780**
Valine		PYR	-1	2	2	-2	0.558	**0.651**

[a]. The metabolic cost of each end product is calculated from glucose which was firstly metabolized through glycolysis (EMP pathway). [b]. The list of bioproducts are not a exhaust list. Representative molecules of each category was summarized in the table. C. The metabolic cost of the biopolymer was calculated to chain length of the biopolymer is 1.

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 5. Plasmids and strains used in the study.

Strains (host strain)	Genotype or plasmid characteristics	Origin
E. coli		
DH5α	*fhuA2 Δ(argF-lacZ) U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	*Invitrogen*
Plasmid		
YLEX	pINA1269-*LEU*	*Yeastern*
pMT15	YLEX derivative replacing hp4d promoter with TEF intron promoter	[5]
pMT91	pACYC derivative with URA3 marker targeting ylLip2 region	[11]
pQK7	pMT15 derivative replacing LEU2 with hygromycin resistance marker	[3]
pQkj1	pUC19 derivative containing markerless URA3 knock-out cassette	*This work*
pQkj2	pMT91 derivative with TEF intron promoter and XPR2 terminator inserted upstream of URA3 marker	*This work*
pQkj3	pQkj2 derivative expressing *E. coli* soluble pyridine nucleotide transhydrogenase UdhA	*This work*
pKJ1	pUC19 derivative with *Y. lipolytica* GPD promoter and LIP1 terminator	*This work*
pKJ2	pKJ1 derivative expressing *E. coli* membrane-bound pyridine nucleotide transhydrogenase PntA	*This work*
pQkj4	pQkj2 derivative expressing *E. coli* pyridine nucleotide PntB and PntA (cloned from pKJ2)	*This work*
pQkj5	pQkj2 derviative expressing *C. acetobutylicum* GAP dehydrogenase GapC	*This work*
pQkj6	pQkj2 derivative expressing *K. lactis* GAP dehydrogenase GPD1	*This work*
pQkj7	pQkj2 derivative expressing *M. circinelloides* malic enzyme MCE2	*This work*
pKJ3	pKJ1 derivative expressing synthetic *C. kluyveri* phosphotransacetylase PTA	*This work*
pQkj8	pQkj2 derivative expressing synthetic *L. mesenteroides* phosphoketolase PK and PTA (cloned from pKJ3)	*This work*
pQkj9	pQK7 derivative expressing *C. acetobutylicum* GAP dehydrogenase GapC	*This work*
pQkj10	pQkj2 derivative expressing *Y. lipolytica* NAD^+ kinase YEF	*This work*
pQkj11	pQkj2 derivative expressing *Y. lipolytica* NAD^+/NADH kinase POS5	*This work*
pQkj12	pQkj2 derivative expressing *Y. lipolytica* NAD^+ kinase UTR1	*This work*
pKJ4	pKJ1 derivative expressing *C. acetobutylicum* GAP dehydrogenase GapC	*This work*
pQkj13	pQkj2 derivative expressing *Y. lipolytica* NAD^+ kinase YEF and *C. acetobutylicum* GAP dehydrogenase GapC (cloned from pKJ4)	*This work*
pQkj14	pQkj2 derivative expressing *M. circinelloides* malic enzyme MCE2 and *C. acetobutylicum* GAP dehydrogenase GapC (cloned from pKJ4)	*This work*
pQkj15	pMT15 derivative expressing 6 x Histidine tagged *Y. lipolytica* glyceraldehyde dehydrogenase ylGPD	*This work*
pQkj16	pMT15 derivative expressing 6 x Histidine tagged *C. acetobutylicum* glyceraldehyde dehydrogenase ylGPD	*This work*
Y. lipolytica		
po1g	*MATa, leu2-270, ura3-302::URA3, xpr2-3*	*Yeastern*

Nature Biotechnology: doi:10.1038/nbt.3763

YL-wt	Po1g, *ylex (LEU2)*	[3]
MTYL065	Po1g, *pMT065 (LEU2, h4pd-ACC1, TEFin-DGA1)*	[5]
ADΔura3	MTYL065, Δ*URA3*	*This work*
AD[a]	ADΔura3, *pQkj2 (URA3, TEFin-XPR2t)*	*This work*
ADudhA	ADΔura3, *pQkj3 (URA3, TEFin-UdhA-XPR2t)*	*This work*
ADpntAB	ADΔura3, *pQkj4 (URA3, TEFin-PntA-XPR2t, GPDp-PntB-LIP1t)*	*This work*
ADgapc	ADΔura3, *pQkj5 (URA3, TEFin-GapC-XPR2t)*	*This work*
ADgpd	ADΔura3, *pQkj6 (URA3, TEFin-GPD1-XPR2t)*	*This work*
ADme	ADΔura3, *pQkj7 (URA3, TEFin-MCE2-XPR2t)*	*This work*
ADpp	ADΔura3, *pQkj8 (URA3, TEFin-PK-XPR2t, GPDp-PTA-LIP1t)*	*This work*
ADgapc2	ADgapc, *pQkj9 (Hygromycin, TEFin-GapC-XPR2t)*	*This work*
ADgg	ADgpd, *pQkj9 (Hygromycin, TEFin-GapC-XPR2t)*	*This work*
ADyef	ADΔura3, *pQkj10 (URA3, TEFin-YEF-XPR2t)*	*This work*
ADpos5	ADΔura3, *pQkj11 (URA3, TEFin-POS5-XPR2t)*	*This work*
ADutr	ADΔura3, *pQkj12 (URA3, TEFin-UTR-XPR2t)*	*This work*
ADgy	ADΔura3, *pQkj13 (URA3, TEFin-YEF-XPRt, GPDp-GapC-LIP1t)*	*This work*
ADgm	ADΔura3, *pQkj14 (URA3, TEFin-MCE-XPRt, GPDp-GapC-LIP1t)*	*This work*
YL-ylGPD	Po1g, *pQkj15 (LEU2, TEFin-ylGPD-6xhistag)*	*This work*
YL-GapC	Po1g, *pQkj16 (LEU2, TEFin-GapC-6xhistag)*	*This work*

a: The plasmids pQkj series are all linearized and transform into Y. lipolytica, the expression cassettes were all integrated into the genome.

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 6. Sequences of primers used in the study.

No.	Primers	Sequences
P1	URA3-KO1	GCCCCAGATAAGGTTCCGA
P2	URA3-KO2	AGTGAATTCGAGCTCGGTACCCATGCCCTCCTACGAAGCTCG
P3	URA3-KO3	CTGCGAACTTTCTGTCCTCGAA
P4	URA3-KO4	TTCGAGGACAGAAAGTTCGCAGTACTCCAAGCAGACCATTGAGCT
P5	URA3-KO5	GGTCGACTCTAGAGGATCCCCCTAACAGTTAATCTTCTGGTAAGCCTC
P6	URA3-KO6	AAGCTGAACAAGCGCTCCATA
P7	pUC19-v-r	GGGTACCGAGCTCGAATTCACT
P8	pUC19-v-f	GGGGATCCTCTAGAGTCGACC
P9	pMT91-15-f	GAGATATACATATGGCAGATCTCAATTGAGAGACCGGGTTGGCGG
P10	pMT91-15-r	TCGCGTGGCCGGCCGATATCGGACACGGGCATCTCACTTG
P11	pMT15-91-f	CAAGTGAGATGCCCGTGTCCGATATCGGCCGGCCACGCGA
P12	pMT15-91-r	CCGCCAACCCGGTCTCTCAATTGAGATCTGCCATATGTATATCTC
P13	TEF-UdhA-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGCCACATTCCTACGATTAC GATG
P14	TEF-UdhA-r	CAAGACCGGCAACGTGGGGTTAAAACAGGCGGTTTAAACC
P15	TEF-f	CCCCACGTTGCCGGTCTTG
P16	TEF-r	CTGCGGTTAGTACTGCAAAAAGTGCTGGTCG
P17	pKJ1-1f	AGTGAATTCGAGCTCGGTACCCCGCAGTAGGATGTCCTGCAC
P18	pKJ1-2f	AGATGCATAGCACGCGTGTAGATACTGTTGATGTGTGTTTAATTCAAGA ATGAAT
P19	pKJ1-3f	TACACGCGTGCTATGCATCTGGTTCATGAGAAGATAAATATATAAATAC ATTGAGA
P20	pKJ1-4f	GGTCGACTCTAGAGGATCCCCCTACCTTGCTCGAATGACTTATTG
P21	pKJ1-v-f	GTATCTACACGCGTGCTATGCA
P22	pKJ1-v-r	TGTTGATGTGTGTTTAATTCAAGAATGAATATAGAG
P23	pKJ1-PntA-f	CATTCTTGAATTAAACACACATCAACAATGCGAATTGGCATACCAAGAG AACG
P24	pKJ1-PntA-r	TGCATAGCACGCGTGTAGATACTTAATTTTTGCGGAACATTTTCAGCAT GCGCTG
P25	TEF-PntB-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGTCTGGAGGATTAGTTACA GCTGCATAC
P26	TEF-PntB-r	CAAGACCGGCAACGTGGGGTTACAGAGCTTTCAGGATTGCATCCA
P27	Qkj2-KJ1-f	CGTATTGTACACGGCCGCATAGAATTCGAGCTCGGTACCC
P28	Qkj2-KJ1-r	CAATTGAGATCTGCCATATGTATATCTCGTCGACTCTAGAGGATCCCC
P29	Qkj2-v-f	GAGATATACATATGGCAGATCTCAATTG
P30	Qkj2-v-r	TATGCGGCCGTGTACAATACG
P31	TEF-GapC-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGGCAAAGATAGCTATTAAT GGTTTTGG
P32	TEF-GapC-r	CAAGACCGGCAACGTGGGGCTATTTTGCTATTTTTGCAAAGTAAGCT
P33	TEF-Gpd1-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGCCCGATATGACCAACGAG TCC
P34	TEF-Gpd1-r	CAAGACCGGCAACGTGGGGTTACACGCCGGCCTCGAA
P35	TEF-Mce2-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGTCGCCTATTATTGATTTTG TTCGTCG
P36	TEF-Mce2-r	CAAGACCGGCAACGTGGGGCTACAATTTACCAGCTTGCTGATTGCT
P37	pKJ1-PTA-f	CATTCTTGAATTAAACACACATCAACAATGAAGTTGATGGAAAACATCT

Nature Biotechnology: doi:10.1038/nbt.3763

		TCGGT
P38	pKJ1-PTA-r	TGCATAGCACGCGTGTAGATACTTAACCTTGTGCTTGAGCTTGAAC
P39	TEF-PK-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGGCCGATTTCGACTCTAAA GAATAC
P40	TEF-PK-r	CAAGACCGGCAACGTGGGGTTATTTCAATGGGGACCAAGTCCAATC
P41	pQK7-TEF-f	CCGGAGCTTGCAGGATCGCCAGAGACCGGGTTGGCGGCGC
P42	pQK7-XPR-r	GAGCGAGTGTTACACATGGAATTGGACACGGGCATCTCACTTG
P43	pQK7-v-f	AATTCCATGTGTAACACTCGCTC
P44	pQK7-v-r	GGCGATCCTGCAAGCTCCGG
P45	TEF-YEF-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGGCCCGCAACACAACGGA
P46	TEF-YEF-r	CAAGACCGGCAACGTGGGGTTAAATGTTGGTGAACCGCTTCTGT
P47	TEF-POS5-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGCGACTACTCATCCGCCGA AC
P48	TEF-POS5-r	CAAGACCGGCAACGTGGGGCTAAGCAACATCGCCTGACG
P49	TEF-UTR1-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGAGCACTCCGGTCAGCGAG TC
P50	TEF-UTR1-r	CAAGACCGGCAACGTGGGGTTACTTCGTGTCCTCCTGCTCG
P51	pKJ1-GapC-f	CATTCTTGAATTAAACACACATCAACAATGGCAAAGATAGCTATTAATG GTTTTGG
P52	pKJ1-GapC-r	TGCATAGCACGCGTGTAGATACCTATTTTGCTATTTTTGCAAAGTAAGCT
P53	TEF-GapC-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGGCAAAGATAGCTATTAAT GGTTTTGGAA
P54	TEF-GapC-r	CAAGACCGGCAACGTGGGGTTAGTGGTGGTGGTGGTGGTGTTTTGCTAT TTTTGCAAAGTAAG
P55	TEF-GPD-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGCCCGATATGACCAACGA
P56	TEF-GPD-r	CAAGACCGGCAACGTGGGGTTAGTGGTGGTGGTGGTGGTGCACGCCGG CCTCGAAGTC
P57	TEF-ylGPD-f	CGACCAGCACTTTTTGCAGTACTAACCGCAGGCCATCAAAGTCGGTATT AACGGA
P58	TEF-ylGPD-r	CAAGACCGGCAACGTGGGGTTAGTGGTGGTGGTGGTGGTGAGCGGAAG CATCCTTCTTG

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Table 7. Sequences of synthetic genes used in this study.

Synthetic gene	Sequences
GPD, originated from *Kluyveromyces lactis*, codon optimized toward *Y. lipolytica*	atgcccgatatgaccaacgagtcctcttcgaagcccgcccagatcaacatcggcatcaacggcttcggccgaatcggacga ctggtgctgcgagccgccctgacccaccccgaggtgaaggtgcgactgatcaacaacccctctaccacccccgagtacgc cgcctacctgttcaagtacgactctacccacggcaagtaccgaggcgaggtcgagttcgacgacgagcgaatcatcatcca gaacgaccacgtgtctgcccacatccccctgtctcacttccgagagcccgagcgaatcccctgggcctcttacaacgtggac tacgtgatcgactctaccggcgtgttcaaggaagtggacaccgcctctcgacacaagggcgtgaagaaggtgatcatcacc gccccctctaagaccgcccccatgtacgtgtacggcgtgaaccacgtgaagtacaacccccctgaccgaccacgtggtgtct aacgcctcttgcaccaccaactgcctggcccccctggtgaaggccctggacgacgagttcggcatcgaagaggccctgat gaccaccatccacgccaccaccgcctctcagaagactgtcgacggcacctcttctggcggcaaggactggcgaggcggc cgatcttgccagggcaacatcatcccctcttctaccggcgctgccaaggccgtgggcaagatcctgcccgagctgaacggc aagatcaccggcatgtctatccgagtgcccaccatcaacatctccctggtggacctgaccttccgaaccgccaagaagacct cttacgacgacatcatgaaggccctcgagcagcgatctcgatctgacatgaagggcgtcctgggcgtgaccaaggacgcc gtggtgtcctctgacttcacctctgactctcgatcttctatcgtggacgccaaggccggcatcgagctgaacgaccacttcttc aaggtgctgtcttggtacgacaacgagtacggctactcttctcgagtggtcgacctgtctatcttcatggcccagaaggacttc gaggccggcgtgtaa
PK, originated from *Leuconostoc mesenteroides*, codon optimized toward *Y. lipolytica*	atggccgatttcgactctaaagaatacttggaattggttgacaaatggtggagagctaccaattatttgtctgccggt atgatcttcttgaagtctaatcctttgttctccgttaccaacactccaatcaaagctgaagatgttaaggttaagccaat tggtcattggggtactatttctggtcaaactttcttgtacgctcatgccaacagattgattaacaagtacggtttgaata tgttctacgttggtggtccaggtcatggtggtcaagttatggttactaatgcttatttggatggtgcctacactgaagat tacccagaaattacccaagacatcgaaggtatgtctcacttgtttaagagattctcattcccaggtggtatcggttctc atatgactgctcaaactccaggttctttacatgaaggtggtgaattgggttactctttgtctcatgcttttggtgctgtttt ggataacccagatcaagttgcttttgctgttgttggtgatggtgaagctgaaactggtccatctatggcttcatggcat tctattaagttcttgaacgctaagaatgatggtgccgttttgccagttttggatttgaatggtttcaagatctccaaccc aaccatcttctctagaatgtccgatgaagaaatcaccaagttctttgaaggtttgggttacagtccaagattcatcga aaacgatgatatccatgattacgccacctatcatcaattggctgctaacattttggatcaagccatcgaagatatcca agccattcaaaatgatgccagagaaaacggtaaataccaagatggtgaaattccagcttggccagttattattgcta gattgccaaaaggttggggtggtccaactcatgatgcttctaacaatccaatcgaaaactctttcagagcccatcaa gttccattgccattggaacaacatgatttggctactttgccagaattcgaagattggatgaattcctacaagcctgaa gaattattcaacgccgatggttccttgaaggatgaattgaaagctattgctccaaagggtgacaaaagaatgtctgc taatccaattactaatggtggtgccgatagatccgatttgaaattgccaaattggagagaattcgccaacgatattaa cgatgacaccagaggtaaagaattcgctgattctaagagaaacatggatatggctaccttgtctaattacttgggtg cagtttctcaattgaaccctactagattcagatttttcggtccagacgaaaccatgtctaatagattgtggggtttgttc aacgttactccaagacaatggatggaagaaatcaaagaaccacaagatcaattattgtccccaaccggtagaatc attgactctcaattgtctgaacatcaagctgaaggttggttggaaggttatactttgactggtagagttggtattttcgc ctcttacgaatctttcttgagagttgttgataccatggttacccaacatttcaagtggttgagacatgcttcagaacaa gcttggagaaatgattacccatccttgaacttgattgctacttctactgctttccaacaagatcataacggttacactc atcaagatccaggtatgttgactcatttggctgaaaagaagtccaacttcatcagagaatatttgccagctgatggta actctttgttggctgtccaagaaagagctttttccgaaagacataaggtcaacttgttgatcgcttctaagcaacctag acaacaatggttcactgttgaagaagctgaagttttggctaacgaaggtttgaagattattgattgggcttctacagc tccatcctccgatgttgatattacttttgcttctgctggtactgaacctaccattgaaactttggctgctttgtggttgatc aatcaagcttttccagatgtcaagttcagatacgttaatgtcgtcgaattattgagattgcaaaaaaagtccgaacct aacatgaacgacgaaagagaattgtctgcagaagaattcaacaagtacttccaagctgataccccagttatttttgg tttccatgcttacgaaaacttgatcgaatcattcttcttcgaacgtaaattcactggtgatgtttacgttcacggttacag agaagatggtgatattaccactacctacgatatgagagtttactcccatttggatagattccaccaagctaaagaag ctgccgaaattttgtctgcaaacggtaagatagatcaagctgctgctgatactttcattgccaagatggatgatacct tggctaagcactttcaagttactagaaacgaaggtagagatatcgaagaattcacagattggacttggtccccattg

Nature Biotechnology: doi:10.1038/nbt.3763

	aaataa
PTA, originated from *Clostridium kluyveri*, codon optimized toward *Y. lipolytica*	atgaagttgatggaaaacatcttcggtttggctaaggctgataagaagaaaatcgttttggctgaaggtgaagaag aaagaaacattagagcctccgaagaaatcatcagagatggtattgctgatatcatcttggtcggttctgaatccgtta tcaaagaaaatgctgctaagttcggtgttaacttggctggtgttgaaatagttgatccagaaacttcttctaagactgc tggttacgctaatgccttctacgaaattagaaagaacaagggtgttaccttggaaaaggcagataagatagttaga gatccaatctacttcgctaccatgatggttaagttgggtgatgctgatggtttggtttctggtgctattcatacaaccg gtgatttgttaagaccaggtttacaaatcgttaagactgttccaggtgcttccgttgtttcttctgtttttttgatgtctgttc cagactgcgaatatggtgaagatggtttttgttgttcgctgattgtgctgttaacgtttgtccaactgctgaagaattg tcctctattgctattactactgctgaaaccgctaagaacttgtgcaaaattgaacctagagttgccatgttgtctttctct actatgggttctgcttcccatgaattggttgataaggttactaaggctaccaagttggctaaagaagctagaccaga tttggatatcgatggtgaattacaattggatgcctccttggttaagaaggttgctgatttgaaagctccaggttctaaa gttgctggtaaggctaatgttttgatcttcccagatattcaagccggtaacattggttacaagttggttcaaagatttgc taaggcagaagccattggtccaatttgtcaaggttttgctaagccaatcaacgacttgtctagaggttgttctgttgat gatatcgttaaggttgttgccgttactgctgttcaagctcaagcacaaggttaa

Nature Biotechnology: doi:10.1038/nbt.3763

Supplementary Note 1.

Stoichiometric equations for de novo fatty acids from glucose in *Y. lipolytica*

The conversion of glucose to SA is depicted in equation *S1* with all the biochemical steps generating or consuming cofactors numbered accordingly. Simplification of equation *S1* led to the equation (1) as shown in **Supplementary Figure 2**.

4.5 Glucose + **8**ATP + **16**NADPH.c = **1**Stearic Acid + **9**NADH.c + **9**NADH.m (1)

Due to the higher energy density of SA in comparison to that of glucose, additional glucose has to be anabolized to generate energy cofactors, in particular, ATP and NADPH to support the fatty acid synthesis. ATPs are generated through complete oxidation of glucose to carbon dioxide with theoretical maximal depicted in equation *S2* assuming maximum efficiency of ATP production shown in equation *S4* and *S5*.

1Glucose → **32**ATP (*S4*)

1NADH ~ **2.5**ATP (*S5*)

Moreover, pentose phosphate pathway has been demonstrated to be and sufficiently supply the NADPH for fatty acid biosynthesis in *Y. lipolytica*[12] with the stoichiometry shown in equation *S6*.

1Glucose + **1**ATP → **12**NADPH + **6**CO_2 (*S6*)

Combining equations *S4-S6* allows simplification of S1 to equation *S7*, on which based the theoretical lipids yield (Y_L) of wild type *Y. lipolytica* is 0.271 g-SA/g-Glucose.

5.83 Glucose → **1**Stearic Acid + **28**ATP (*S7*)

When one assumes that NADH can be directly converted to NADPH without any energy penalty, the equation (1), sufficient NADPH is provided from *EMP* and TCA other than pentose phosphate pathway and equation (1) becomes equation *S7*. Combination of equations *S7* with *S3* and *S4* affords equation *S8*, and therefore the Y_L is elevated to 0.344 g-SA/g-glucose.

4.5 Glucose + **8**ATP = **1**Stearic Acid + **2**NADH.m (*S8*)

4.59 Glucose = **1**Stearic Acid (*S9*)

Recycling the NADH.c to NADPH.c Ideally, if each synthetic pathway operates at its own maximum efficiency, all 9 cytosolic NADH can be replaced with NADPH:

4.5 Glucose + **8**ATP + **7**NADPH.c → **1**Stearic Acid + **9**NADH.m (*S10*)

Combining equations *S10 with S3-S5* gives equation *S11*, on which based the theoretical lipids yield (Y_L) of wild type *Y. lipolytica* is 0.311 g-SA/g-Glucose.

5.08 Glucose → **1**Stearic Acid + **13.9**ATP (*S11*)

Installation of engineered POM cycle that regenerate NADPH from NADHs from both cytosol and mitochondria, leading to elevation of Y_L to 0.329 g-SA/g-glucose (Equation *S12*).

4.81 Glucose = **1**Stearic Acid (*S12*)

Functional reconstitution of NOG pathway would allow the complete carbon conservation, leading to the formation of 3 mol acetyl-CoA from 1 mol glucose.

1Glucose + **1**ATP → **3**Acetyl-CoA (*S13*)

If one assume acetyl-CoA in Eq. *S1* was exclusively supplied by NOG pathway, while ATP and NADPH are derived from oxidative respiration and pentose phosphate pathway. The assumption results in the simplification of Eq. *S1* to Eq. *S14*:

4.72 Glucose → **1**Stearic Acid (*S14*)

Therefore, the theoretical lipid yield reached 0.335 g-SA/g-glucose when NOG pathway is working at its 100% capacity.

Recycling NADH.m to NADPH.c. In contrast to NADH.c, NADH.m needs to be shuttled to the cytosol before it can be converted to NADPH. The inner membrane of the mitochondrion is impermeable to NADH but permeable to metabolites, such as malate or isocitrate, which can be oxidized by $NADP^+$-dependent malic enzyme and isocitrate dehydrogenase, respectively, to yield NADPH.c. Conversion of NADH.m to NADPH.c can therefore be achieved by constructing POM cycle (**Supplementary Fig. 4d**) or citrate/isocitrate/2-ketoglutarate (CIK) cycle[13] across the mitochondrial membrane. Though still inconclusive, the *E* of ADgm (99.1 ± 2.5%) indicates that POM cycle by itself might be sufficient. Undoubtedly, incorporating efficient malate transporters shuttling malate out of mitochondrion will help fully realize POM cycle's potential[14]. Furthermore, generation of NADH.m by pyruvate decarboxylation can be by-passed via the introduction of *E. coli* pyruvate formate lyase (PFL), which cleaves pyruvate into acetyl-CoA and formate in the cytosol. The formate can be subsequently oxidized by $NADP^+$-dependent formate dehydrogenases (FDH) to give NADPH. However, *E. coli* PFL had been demonstrated to be extremely sensitive to oxygen[15]. To make it compatible with *Y. lipolytica*, the PFL has to be engineered to improve oxygen tolerance. Alternatively, *Y. lipolytica* could be engineered to adapt to anaerobic conditions.

The synthetic pathway was proposed to be enabled by expressing *E. coli* pyruvate-formate lyase PFLB, its cognate activating enzyme PFLA and the $NADP^+$-dependent formate dehydrogenase. Combination of the synthetic pathway with $NADP^+$-dependent GPDs allows substitution of glycolytic NADH with NADPH

1Glucose → **2**Acetyl-CoA + **2**ATP + **4**NADPH (*S15*)

Since the acetyl-CoA is generated from pyruvate directly instead of through activity of ACL, the ATP needed for fatty acid synthesis is less, leading to Eq. *S16*:

4.5 Glucose + **8**ATP + **16**NADPH.c → **1**Stearic Acid + **9**ATP + **18**NADPH (*S16*)

Therefore, the theoretical maximum yield is 0.351 g-SA/g-glucose. Furthermore, per mol SA synthesized using this stoichiometry there would be 2 mol NADPH and 1 mol ATP in excess. The net ATP gain would make this synthetic pathway favorable under strict anaerobic conditions if the excess NADPHs could be oxidized in a futile pathway or consumed in the production of a reduced molecule, such as mannitol.

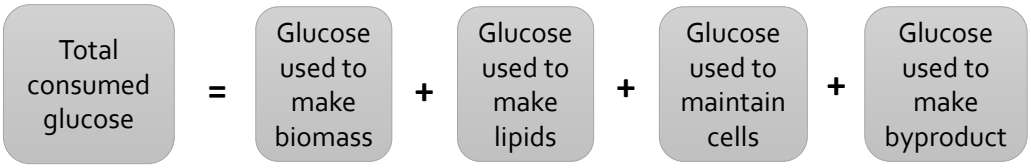
Supplementary Note 2.

Development of a mathematical model of process yield.

For a given fed-batch fermentation that produce lipids from glucose, the overall process yield Y can be defined as

$$Y = \frac{L}{G} \quad (S1)$$

Where L is the lipids titer (g/L) and G represented the totally consumed glucose (g/L). Given that glucose is the exclusive carbon source of the fermentation and it was allocated into four major products:



It can be rewritten to a mathematical expression:

$$G = \frac{B}{Y_B} + \frac{L}{Y_L} + mB + \frac{W}{Y_w} \quad (2)$$

$$C = \frac{L}{B + L} \quad (S2)$$

B: non-lipid biomass (g/L); m: cell maintenance (g/g dry cell weight); W: byproduct titer (g/L); Y_B, Y_L and Y_w respectively represent the conversion yield of glucose to the corresponding products-non-lipid biomass, lipids and byproduct(s).

To simplify the equation (2), three key assumptions are made on the basis of growth characteristic and the data obtained from fed-fermentation practice previously and in the study: 1. biomass is generated at its maximal yield Y_B and lipids are biosynthesized at its maximal yield Y_L. 2. no byproduct formation. 3. cell maintenance is negligible.

As shown in Supplementary Figure 7, a fed-batch fermentation of engineered *Y. lipolytica* (for example, baseline strain AD) features a biphasic growth characteristic. In the growth phase (from 0 h to 36 h), the *Y. lipolytica* cells grows exponentially at the specific growth rate of ~0.25 h.

After depletion of nitrogen (~40 h), the cells enters lipid production phase, in which cells stops doubling and their metabolism shifted to complete lipogenesis. Therefore, it is reasonable to assume that non-lipid biomass formation of *Y. lipolytica* occurs exclusively in the growth phase. As a result, the non-lipid biomass yield Y_B is identical to Yx/s, which can be either directly measured or obtained from literatures. Besides, the cell maintenance can be neglected in the growth phase.

$$Y_B = Y_{\frac{x}{s}} = \frac{g\ biomass}{g\ substrate} \quad (S3)$$

During lipid production phase, the increase in dry cell weight during lipid production phase almost exclusively originate from accumulation of lipids as evidenced by the unchanged cell numbers and gradual enlargement of lipid bodies. The lipids yield (defined as maximal yield) at this phase is 0.235 g-FA/g-glucose, accounting for 87% of the stoichiometric maximal (0.271 g/g). Therefore, it is fairly to assume that lipids are biosynthesized at its maximum conversion yield. Furthermore, only small portion (up to 13%) of the substrate contributes to maintain the cells during the lipid production phase. Given that the about half of the total glucose is consumed in lipid production phase, the glucose consumed for maintenance is ~6.5% of total consumption of glucose, which is neglected due to its small contribution.

Additionally, previous process engineering efforts have been centered on minimization of the chief byproduct citrate during the fermentation and resulted in limited the citrate titer around 10 g/L at the end of fermentation by fine-tuning the dissolved oxygen level at the lipid production phase. ~10 g/L of citrate was derived from ~9.4 g/L glucose (theoretical yield 1.07 g-citrate/g-glucose), accounting for <3.7 % of total consumed glucose (> 250 g/L).

Therefore, the allocation of glucose to cell maintenance and byproduct citrate only take up less than 10% of the total consumed glucose. And the cells are making biomass and lipids at high efficiencies (close to 100%) at growth phase and lipid production phase respectively.

Furthermore, we proceed to establish a stoichiometric model based on the metabolic pathways for the synthesis of lipids from glucose. We chose stearic acid as the end product to simplify the model and analysis since living microbial cells contain a variety of lipids including triacylglycerides, steryl esters, phospholipids, fatty acyl-CoAs, free fatty acids and other fatty acid derivatives. After transesterification reactions, these lipid molecules were converted to five FAMEs – methyl palmitate, methyl palmitoleate, methyl stearate, methyl oleate and methyl linoleate. The composition of the total fatty acids varies from strain to strain. Out of the three major fatty acids, stearic acid showed the best fit with the previously published results on three different microbial hosts- *E. coli*, *Y. lipolytica* and *R. toruloides*. Therefore, we selected stearic acid as our example to develop the lipid model. Undoubtedly, the model can be improved by incorporating additional considerations such as the composition of the fatty acids if they could be easily measured and remained largely unchanged across all the strains.

Supplementary Note 3.

Discussion of the established yield model

Physiological range of each parameter. Y_L is solely determined by the metabolic pathway stoichiometry involved in *de novo* biosynthesis (Supplementary Fig. 1) and it ranges from the baseline (0.271 g/g according to metabolism in the native *Y. lipolytica*)[13] to the thermodynamics maximal (0.362 g/g) which is calculated by the complete energy conservation[16]. On the other hand, Y_B can be readily measured by experiments (0.55 g-dry cell weight/g-glucose) for *Y. lipolytica* or consult literature for the empirical different microorganisms' biomass yield if one wanted extend the model to another organism. Of course, a whole genome scale metabolic model can also be used to determine Y_B but its accuracy has to be analyzed before used to predict Y. Finally, lipid content C simply ranges from 0% to 100%. It is important to notice that the definition of C is not restricted to intracellular products like lipids in the study but also can be extended to all biochemical products, such as ethanol, butanol or other natural products. Furthermore, the three parameters, in spite of possess the physical meanings and units, yet are dimensionless and allow the full transformations of equations related to the model.

Dependence of process yield Y on parameters. The above derived mathematical model (Equation 3 and Supplementary Note 2) demonstrates that Y is determined by Y_B, Y_L and C. However, it is not obvious to identify how each parameter affect the optimization target Y. To address the issue, we perform full optimization for a series of values within the physiological ranges of each parameter.

In **Supplementary Fig. 1a**, we showed how Y as functions of the lipid content C response to the increase of the other two yield parameters- Y_B and Y_L. As expected, increasing Y_B and Y_L both increases Y, though with different dynamics depending on C. Interestingly, Y is the most responsive to Y_B when C is around the mid-point-50%, with only minute increments when C is close to either 0% or 100%. On the other hand, Y_L and C seem to have a synergistic effects on optimization of Y. Most significant increase of Y is found to happen by elevating Y_L with C approaching 100%. Considering the practical situation that *Y. lipolytica* has been engineered in our hand to afford C 50%-60%, optimizing Y_B and Y_L could be both good strategies for optimizing Y for production of lipids.

In **Supplementary Fig. 1b**, we plotted Y as functions of Y_B (red) and Y_L (blue) with three different values in C. We could reach similar conclusions as in the previous analysis on **Supplementary Fig. 1a**. Firstly, the slopes of blue curves are significantly higher than those of red curves, indicating Y_L is more effective than Y_B at the range of C (50% - 70%) for optimizing Y. Secondly, Y is responsive to C regardless of the values of Y_L or Y_B.

To direct our engineering practice in elevating the process yield of lipids, it is necessary and key to identify which parameter is the most effective in optimizing Y. Toward this end, we performed single-point sensitivity analyses at two different C values-50% and 70%, which are respectively the lipid contents achieved respectively previously in our own experiments[3, 5] and are most relevant to previously reported lipid contents of oleaginous yeasts[7, 17, 18]. The results are shown in **Supplementary Figure. 1c and 1d**. Both C and Y_L were identified to be very influential on Y at both Cs, while YB is considerably less sensitive. Despite C is slightly more sensitive than YL at 50%, the Y_L becomes the most dominate parameters at 70%. The sensitivity analyses demonstrated that although all three parameters positively contribute to the optimization

of Y, yet we should probably focus on optimizations of C and Y_L in the engineering practice toward a high lipid yield.

References

1. Dellomonaco, C., Clomburg, J.M., Miller, E.N. & Gonzalez, R. Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**, 355-359 (2011).
2. Xu, P. et al. Modular optimization of multi-gene pathways for fatty acids production in E. coli. *Nature Commun.* **4**, 1409 (2013).
3. Qiao, K. et al. Engineering lipid overproduction in the oleaginous yeast Yarrowia lipolytica. *Metab. Eng.* **29**, 56-65 (2015).
4. Blazeck, J. et al. Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production. *Nature Commun.* **5**, 3131 (2014).
5. Tai, M. & Stephanopoulos, G. Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metab. Eng.* **15**, 1-9 (2013).
6. Liu, L.Q. et al. Surveying the lipogenesis landscape in Yarrowia lipolytica through understanding the function of a Mga2p regulatory protein mutant. *Metabolic engineering* **31**, 102-111 (2015).
7. Li, Y.H., Zhao, Z.B. & Bai, F.W. High-density cultivation of oleaginous yeast Rhodosporidium toruloides Y4 in fed-batch culture. *Enzyme Microb Tech* **41**, 312-317 (2007).
8. Strand, M.K. et al. POS5 gene of Saccharomyces cerevisiae encodes a mitochondrial NADH kinase required for stability of mitochondrial DNA. *Eukaryotic cell* **2**, 809-820 (2003).
9. Kawai, S., Suzuki, S., Mori, S. & Murata, K. Molecular cloning and identification of UTR1 of a yeast Saccharomyces cerevisiae as a gene encoding an NAD kinase. *Fems Microbiol Lett* **200**, 181-184 (2001).
10. Shi, F., Kawai, S., Mori, S., Kono, E. & Murata, K. Identification of ATP-NADH kinase isozymes and their contribution to supply of NADP(H) in Saccharomyces cerevisiae. *Febs J* **272**, 3337-3349 (2005).
11. Tai, M. (Massachusetts Institute of Technology, 2012).
12. Wasylenko, T.M., Ahn, W.S. & Stephanopoulos, G. The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in Yarrowia lipolytica. *Metab. Eng.* **30**, 27-39 (2015).
13. Ratledge, C. The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: a reappraisal and unsolved problems. *Biotechnology letters* **36**, 1557-1568 (2014).
14. Casal, M., Paiva, S., Queiros, O. & Soares-Silva, I. Transport of carboxylic acids in yeasts. *FEMS microbiology reviews* **32**, 974-994 (2008).
15. Wagner, A.F.V., Frey, M., Neugebauer, F.A., Schafer, W. & Knappe, J. The Free-Radical in Pyruvate Formate-Lyase Is Located on Glycine-734. *P Natl Acad Sci USA* **89**, 996-1000 (1992).
16. Dugar, D. & Stephanopoulos, G. Relative potential of biosynthetic pathways for biofuels and bio-based products. *Nature Biotechnol.* **29**, 1074-1078 (2011).
17. Liu, L., Pan, A., Spofford, C., Zhou, N. & Alper, H.S. An evolutionary metabolic engineering approach for enhancing lipogenesis in Yarrowia lipolytica. *Metabolic engineering* **29**, 36-45 (2015).
18. Abghari, A. & Chen, S. Yarrowia lipolytica as an oleaginous cell factory platform for the production of fatty acid-based biofuel and bioproducts. *Frontiers in Energy Research* **2** (2014).

Nature Biotechnology: doi:10.1038/nbt.3763