nature biotechnology

LETTERS

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

Kangjian Qiao, Thomas M Wasylenko, Kang Zhou, Peng Xu & Gregory Stephanopoulos

Microbial factories have been engineered to produce lipids from carbohydrate feedstocks for production of biofuels and oleochemicals. However, even the best yields obtained to date are insufficient for commercial lipid production. To maximize the capture of electrons generated from substrate catabolism and thus increase substrate-to-product yields, we engineered 13 strains of *Yarrowia lipolytica* with synthetic pathways converting glycolytic NADH into the lipid biosynthetic precursors NADPH or acetyl-CoA. A quantitative model was established and identified the yield of the lipid pathway as a crucial determinant of overall process yield. The best engineered strain achieved a productivity of 1.2 g/L/h and a process yield of 0.27 g–fatty acid methyl esters/g-glucose, which constitutes a 25% improvement over previously engineered yeast strains. Oxygen requirements of our highest producer were reduced owing to decreased NADH oxidization by aerobic respiration. We show that redox engineering could enable commercialization of microbial carbohydrate-based lipid production.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

Lipids, specifically fatty-acid-derived lipids, are important feedstocks for the fuel and oleo-chemical industries. Owing to the depletion of fossil fuels, plant-oil- and animal-fat-derived lipids are being developed as renewable feedstocks for biodiesel production[1]. However, availability of these edible oils is insufficient for production of biofuels at scale, and their use for fuel production conflicts with food supply[2]. Carbohydrates are the most abundant renewable feedstocks, and advances in lignocellulosic biomass hydrolysis technologies mean that microbial conversion of non-food carbohydrates to lipids is a promising option for sustainable lipid production[3].

Oleaginous yeasts are characterized by complex internal membranes that enable high storage capacity of neutral lipids (mainly triacylglycerides). These yeasts grow fast and produce lipids at high rates[2]. The model oleaginous yeast *Yarrowia lipolytica* has been engineered to improve lipid production[4–7]. Engineering strategies include increasing the availability of biosynthetic precursors[8] and lipogenic pathway flux[9], shutting down degradation pathways including lipolysis[10] and β-oxidations[11–13], and removing inhibitory intermediates[14]. Despite much progress, commercialization of microbial oils is limited to the production of high-value commodity chemicals[2,13]. For low-priced biofuels, commercialization depends on lowering the feedstock cost, which is the single most expensive component of total production cost. This requires improving the overall yield of carbohydrate to lipid conversion and productivity.

To assess the extent to which the yield of lipid production can be further improved, we had to first define the theoretical limits of process yield and delineation of the main parameters affecting it. We developed a mathematical model that establishes a quantitative relationship between process yield **Y** and three main parameters: the yield of non-lipid biomass $\mathbf{Y_B}$, cellular lipid content **C** and the yield of the lipid synthesis pathway $\mathbf{Y_L}$. *In silico* sensitivity analysis using the model identified lipid content as the critical parameter at low **C**, whereas $\mathbf{Y_L}$ was the most influential driver of process yield maximization at higher lipid contents (**Supplementary Fig. 1**). To experimentally improve $\mathbf{Y_L}$, we designed four synthetic pathways converting cytosolic NADH to NADPH, and thereby increased $\mathbf{Y_L}$ and **Y** through efficient NADH recycling.

As suggested by the stoichiometry of lipid synthesis, excess NADH is formed in a typical lipid biosynthetic pathway (**Supplementary Fig. 2**). The overall stoichiometry of glucose conversion to SA is

$$4.5\,\text{Glucose} + 8\,\text{ATP} + 16\,\text{NADPH.c} = 1\,C_{17}H_{35}COOH + 9\,\text{NADH.c} + 9\,\text{NADH.m} \quad (1)$$

If it is assumed that reducing equivalents generated in cytosol and mitochondria in the form of NADH (NADH.c and NADH.m, respectively) can be converted to the cytosolic NADPH required for fatty acid synthesis, a maximum $\mathbf{Y_L}$ for SA can be calculated as 0.344 g-SA/g-glucose in comparison to 0.271 g/g glucose in native *Y. lipolytica* (**Supplementary Note 1**). Moreover, inspection of flux distributions in a strain of *Y. lipolytica* engineered to overproduce lipids shows that the pentose phosphate pathway is the main source of NADPH used for lipid synthesis and that lipogenesis is limited by the supply of NADPH[15]. This suggests that lipid yields could be improved by increasing the supply of NADPH, which could be achieved by engineering synthetic pathways that convert NADH to NADPH.

To establish the relationship between **Y** and $\mathbf{Y_L}$, a mass balance for total consumed glucose (**G**) producing non-lipid biomass (**B**), lipids (**L**), by-products (**W**) and cell maintenance (**mB**) yields the following equation, assuming that each of the corresponding biochemical pathways operates at its maximum efficiency (where $\mathbf{Y_B}$, $\mathbf{Y_L}$

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Correspondence should be addressed to G.S. (gregstep@mit.edu).

Received 4 April 2016; accepted 8 December 2016; published online 16 January 2017; doi:10.1038/nbt.3763

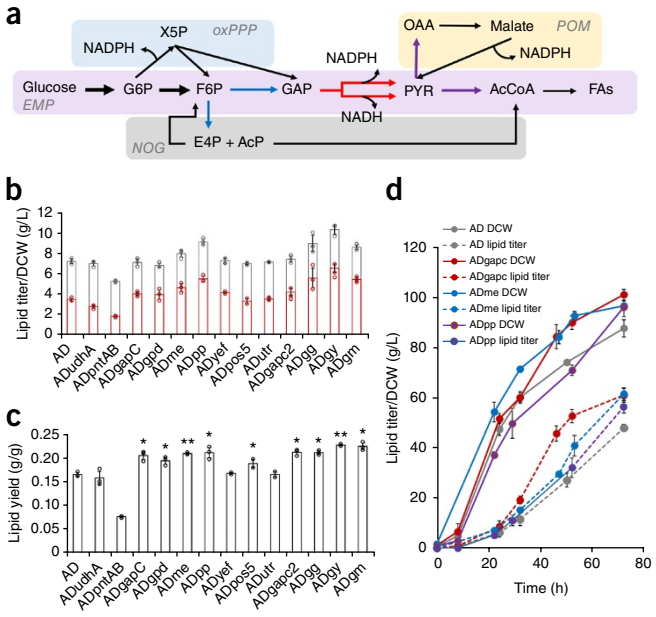

Figure 1 Improving the lipid yields of engineered *Y. lipolytica* by introducing synthetic pathways that recycles NADH.c to NADPH.c or acetyl-CoA. (**a**) Illustration of the three synthetic pathways, namely, $NADP^+$-dependent GPD (red line), POM cycle (purple line), and NOG pathway (blue line). Different color shadings indicate different pathways including EMP pathway (light purple), oxidative pentose phosphate pathway (oxiPPP, blue), POM cycle (orange) and NOG pathway (gray color). (**b**,**c**) Lipid titer and dry cell weight (**b**), and lipid yield of control and engineered *Y. lipolytica* strains (**c**) in shake flask experiments. (**d**) Time-course profiles of lipid accumulation (lipid titer) and cell growth (dry cell weight) of engineered *Y. lipolytica* strains including ADgapc, ADme and ADpp compared to those of control strain AD in fed-batch fermentations. For columns 1 –7 and 11–14 in **b**, **c**, $n = 3$; for column 8–10 in **b** and **c**, $n = 2$; for **d**, $n = 2$. Error bars, mean ± s.d. Statistically significant differences between each engineered *Y. lipolytica* strain and the baseline strain AD were denoted $^*P < 0.05$, $^{**}P < 0.01$ (two-tailed Student's *t*-test). DCW, dry cell weight.

and $\mathbf{Y_w}$ are, respectively, the actual yields of biomass, lipids and by-product formation):

$$G = \frac{B}{YB} + \frac{L}{YL} + mB + \frac{W}{Yw} \quad (2)$$

Neglecting the primary by-product citrate, which was previously minimized to less than 10 g/L in bioreactor studies[9,14], and yeast cell maintenance (contributing less than 6% of total consumed glucose) (**Supplementary Note 2**) allows the derivation of an equation for the overall process yield:

$$Y = \frac{\frac{YBYL}{YB-YL}C}{\frac{YL}{YB-YL}+C} \quad (3)$$

where C = L/(B + L) is the lipid content and Y = L/G the overall process lipid yield.

Supplementary Table 1 shows that this model predicts (with less than 5% discrepancy) experimentally measured overall lipid process yields in *Escherichia coli*[16,17], *Y. lipolytica*[9,14] and *Rhodosporidium toruloides*[18] for a range of $\mathbf{Y_B}$ and **C** values that were reported in optimized fermentations. We next examined the sensitivity of **Y** with respect to each model parameter at two different **C** (50% and 70%), both of which have been obtained in previously engineered *Y. lipolytica* strains[9,14]. A one-way sensitivity analysis showed that **C** is most important at low **C** while $\mathbf{Y_L}$ was identified as the most crucial determinant of process yield when **C** > 60%. This analysis identified maximization of $\mathbf{Y_L}$ as our engineering target (**Supplementary Fig. 1** and **Supplementary Note 3**).

$\mathbf{Y_L}$ (solely determined by the stoichiometry of the lipid pathway) could be increased by converting excess NADH to produce either of the precursors of lipid synthesis, namely, cytosolic NADPH or acetyl-CoA. If all NADH.c were converted into NADPH, $\mathbf{Y_L}$ would increase to 0.311 g-SA/g-glucose (**Fig. 1a** and **Supplementary Fig. 3**). To this end, we engineered in a *Y. lipolytica* background all variants in which acetyl-CoA carboxylase (ACC) and diacylglyceride acyltransferase (DGA) were overexpressed[9] (control strain **AD**).

It has previously been established that *E. coli* transhydrogenases UdhA and PntAB can convert excess cytosolic NADH directly to NADPH.c in *Y. lipolytica*[19]. Overexpression of *E. coli* UdhA had no significant effect on lipid yield, titer or biomass, whereas overexpression of the membrane-bound enzymes PntAB changed the cell morphology (**Supplementary Fig. 4**) and disrupted cell growth and lipid production (**Fig. 1b**).

Second, the 9 NADH.c are solely generated by the biochemical reaction catalyzed by glyceraldhyde-3-phosphate dehydrogenase (GPD), which favors NAD^+ rather than $NADP^+$ (**Supplementary Fig. 5**). To switch the cofactor preferences of this enzyme, we introduced two $NADP^+$-dependent GPDs, respectively, into the **AD**: GapC from *Clostridium acetobutylicum*[20,21] and GPD1 from *Kluyveromyces lactis*[22]. This generated strains **ADgapc** and **ADgpd**, both of which showed unchanged lipid titer and dry cell weight, but with 20.0% and 17.8% higher lipid yields, respectively, than those of **AD** (**Fig. 1c**). In bioreactor cultivations, the best performer **ADgapc** mirrored the growth of **AD** but its lipid process yield in shake flasks was 0.229 g-lipid/g-glucose (hereafter, g/g) compared with 0.184 g/g for **AD**. Notably, lipid yield during the lipid production phase (36–72 h) reached 0.279 g/g, exceeding the $\mathbf{Y_L}$ (0.271 g/g) of **AD** (**Fig. 1d** and **Supplementary Fig. 6**).

The third strategy was to activate the pyruvate/oxaloacetate/malate (POM) cycle, which can convert 1 mol NADH to 1 mol NADPH at a cost of 1 mol ATP (**Fig. 2a**). The endogenous malic enzyme (ylMAE) in *Y. lipolytica* is a mitochondria-associated NAD^+-dependent enzyme, whose overexpression and knockout has little impact on lipid production[23]. Therefore, we opted to improve the POM cycle by incorporation of a cytosolic $NADP^+$-dependent malic enzyme, a strategy known to be effective in non-*Yarrowia* oleaginous fungi[24,25]. We cloned and expressed MCE2 from *Mucor circinelloides* in strain **AD** to yield strain **ADme**[26]. The lipid yield of **ADme** in a shake flask was increased from 0.17 in **AD** to 0.21 g/g with marginal increases in biomass and lipid content (**Fig. 1c**). **ADme** exhibited a similar time course of growth as **AD** and **ADgapc** in bioreactor runs, but much higher lipid titer and yield than those of **AD** (**Fig. 1d** and **Table 1**). Strain **ADgapc** and strains with the engineered POM cycle thus showed promise of improved lipid yield. In summary, of the three tested strategies that recycle NADH.c to NADPH, overexpression of both $NADP^+$-dependent GPD and malic enzyme were demonstrated to be effective in lipid yield improvement.

An alternative strategy to conserve electrons is to divert NADH.c into acetyl-CoA. The recently reported non-oxidative glycolytic (NOG) pathway can yield 3 mol acetyl-CoA from 1 mol glucose, bypassing the Embden–Meyerhof–Parnas (EMP) pathway[27] (**Supplementary Fig. 7**). To activate the NOG pathway, we co-expressed

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

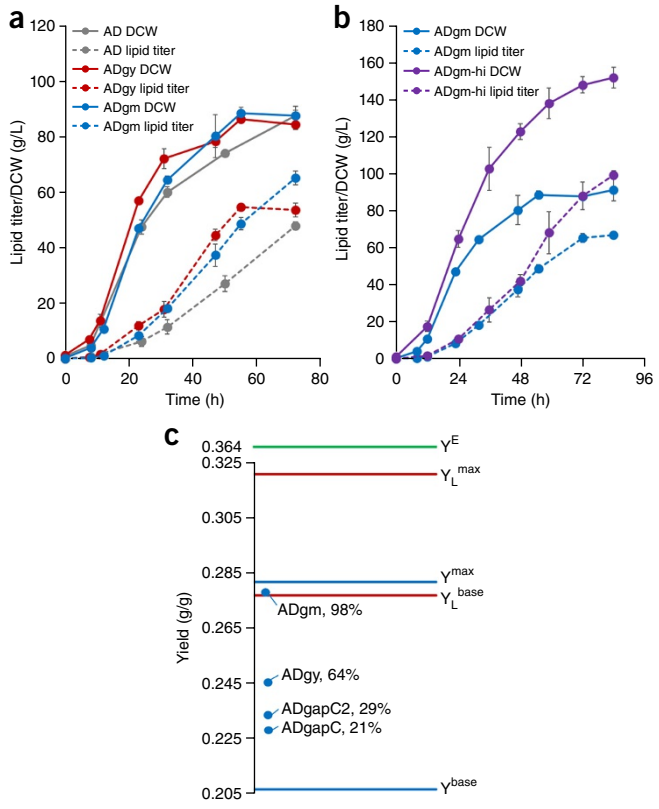

Figure 2 Yield model-directed optimization of strain engineering and the resultant fermentation profiles of engineered *Y. lipolytica* contains combination of synthetic pathways. (**a**) Time courses of cell growth and lipid production of ADgy (red line) and ADgm (blue line) in comparison to those of AD (gray line) under the same fermentation conditions. (**b**) Optimization of the fed-batch fermentation by increasing nitrogen. ADgm-hi, ADgm cultured in a high-density fed-batch fermentation in which the starting ammonium concentration doubles. (**c**) Yield model-directed optimization of process yield. The green colored line represents the maximum possible yield, while the two red lines define, respectively, the upper and lower bounds of the lipid pathway yield for a particular strain harboring a specific synthetic pathway for NADH recycling. The blue lines represent the upper and lower bounds of the overall process yield for strains harboring the same synthetic pathway. For **a**–**c**, $n = 2$. Error bars, mean ± s. d.

phosphoketolase from *Leuconostoc mesenteroides* and phosphate acetyltransferase from *Clostridium kluyveri* in **AD**[28,29]. The resulting strain **ADpp** exhibited 41% higher dry cell weight and 16.4% higher lipid content in shake flask fermentations (**Fig. 1b**). While lipid yields were moderately increased (**Fig. 1c**), the maximum yield during lipid production remained lower than 0.271 g/g (**Supplementary Fig. 8**).

Considering the positive effect of the $NADP^+$-dependent GPD pathway, we sought to increase the carbon flux through this pathway branch by introducing additional copies of GapC or GPD1 into **ADgapc**. This generated strains **ADgapc2** and **ADgg**. Despite increased transcription of the heterologous GPDs, neither strain had improved lipid yield (**Fig. 1c**). We also examined whether increasing the pool size of $NADP^+$ would strengthen the catalytic efficiency of $NADP^+$-dependent GPDs by overexpressing the NAD^+/NADH kinases, which catalyze the final step of $NADP^+$ biosynthesis. Out of the three *Y. lipolytica* NAD^+/NADH kinase isoforms (**Supplementary Table 2**), we found that only the expression of ylYEF led to improvement of lipid yield, while co-expression of functional ylYEF and GapC led to a modest process yield increase (**Fig. 1c**). Finally, co-expression of GapC and MCE2, which enables the combination of the two orthogonally functional synthetic pathways interconverting NADH to NADPH, further increased lipid yield to 0.231 g/g (**Fig. 1c**).

Shake flask tests are useful for strain comparison, however, they underestimate true strain potential due to limited culture controls. Thus, we next characterized the two best strains, **ADgy** and **ADgm**, in bioreactors and found that both strains exhibited similar growth time courses and elevated lipid production to control **AD** (**Fig. 2a**). In the strains with increased NADPH supply, both specific productivity and lipid content were improved, leading to significantly increased overall process yields ($P < 0.05$, $n = 2$) (**Fig. 2b** and **Supplementary Table 3**). Notably, although **ADgy** had the highest lipid productivity (1.57 g/L/h from 31 h to 55 h), the **Y** of strain **ADgm** was the highest (**Fig. 2c**, 0.28 g/g).

In fed-batch fermentations of *Y. lipolytica* where the substrate glucose was continuously supplemented, high dO_2 induced citrate production while *de novo* fatty acid synthesis was inhibited in microaerobic or anaerobic conditions (**Supplementary Fig. 9**). Therefore, to maximize lipid yield, bioreactor dO_2 must be carefully controlled. In a nitrogen-limited culture, the amount of nitrogen in the starting medium determines final cell concentrations, positively contributing to lipid productivity. In comparison with control strain **AD**, **ADgm** had a reduced oxygen requirement, as evidenced by the dO_2 level of **ADgm** (**Supplementary Fig. 10**). This is because reduced NADH levels in engineered strains with high pathway efficiencies require less oxygen for regeneration of NAD^+ from NADH by respiration.

We exploited this feature to further increase cell density by supplementation of ammonium sulfate to 17.6 g/L, which yielded an increase of 67.4% in biomass, while simultaneously elevating the fatty acid methyl esters titer and productivity to 99 g/L and 1.2 g/L/h, respectively. The lower specific oxygen requirements of the engineered

Table 1 Lipid production in engineered *Y. lipolytica* strains in fed-batch fermentations

Strain	Cell growth	Glucose consumed	Lipids (FAMEs) produced				
	DCW (g/L)	(g/L)	Titer (g/L)	Content (%)	Yield (g/g)	Efficiency	Rationale for genetic modulation(s)
AD	87.4	260.3	47.8	54.7	0.184	n/a	n/a
ADgapc	101.5	276.5	63.3	62.5	0.229	21.8%	NADH.c → NADPH.c
ADme	96.4	248.1	61.4	63.7	0.247	50.1%	NADH.c + ATP → NADPH.c
ADpp	93.4	260.8	56.2	52.7	0.216	−11.2%	NADH.c → Acetyl-CoA.c
ADgy	86.3	223.2	54.6	63.2	0.244	63.5%	NAD^+ → $NADP^+$; NADH.c → NADPH.c
ADgm	90.9	237.5	66.8	75.5	0.280	99.1/52.8%	NADH.c → NADPH.c; NADH.c + ATP → NADPH.c NADH.m→ NADPH.c (tbd)
ADgm-hi[a]	148.0	367.4	98.9	66.8	0.269	100.4/48.9%	

Average characteristics from duplicate fed-batch fermentations including titer, consumed glucose, lipid titer, content and yield are shown here. Detailed information regarding the individual bioreactor experiments is available in **Supplementary Table 3**.
[a]ADgm-hi represents the strain ADgm cultured in a high density fed-batch fermentation in which the starting ammonium concentration doubles. Average fermentation characteristics from triplicate fed-batch fermentations are shown here. FAMEs, fatty acid methyl esters; DCW, dry cell weight.

Table 2 Comparison of fatty acid and fatty acid methyl ester (FAMEs) production in selected engineered microorganisms

Product	Microorganism host	Fermentation time (h)	Titer (g/L)	Yield (g/g)	Productivity (g/L/h)	Lipid content (%)	Carbon source	Medium/fermentation mode	Ref.
Fatty acids	*E. coli*	60	~7	0.28	0.117	78	Glucose	MM/batch	16
Fatty acids	*E. coli*	70	8.6	0.092	0.124	21.05	Glucose	MM/fed-batch	17
Fatty acids	*E. coli*	72	5.2	0.26	0.072	n/a	Glucose	MM, M9/batch	31
Fatty acids	*S. cerevisiae*	48	2.2	0.115	0.046	~18[b]	Glucose	Rich YPD/fed-batch	32
Fatty acids	*S. cerevisiae*	72	0.4	n/a	0.006	n/a	Glucose	YNB/shake flask	33
FAMEs	*R. toruloides*	134	71.9	0.23	0.540	67.5	Glucose	Rich 15.7 g/L YE 10 g/L peptone/fed-batch	18
FAMEs	*R. toruloides*[a]	217	16.4	0.23	0.08	61.1	Glucose	Rich 0.75 g/L YE/shake flask	27
FAMEs	*Y. lipolytica*	120	28.5	0.195	0.143	61.7	Glucose	MM YNB + 1 g/L YE/batch	9
FAMEs	*Y. lipolytica*	145	15.25	0.202	0.105	75	Glucose	YNB/batch	11
FAMEs	*Y. lipolytica*	78	55	0.234	0.707	69	Glucose	YNB+YE/fed-batch	14
FAMEs	*Y. lipolytica*[b]	145	39.1	0.244	0.270	77.60	Glucose	YNB/batch	34
FAMEs	*Y. lipolytica*	192	25	0.213	0.130	79–85	Glucose	YNB/fed-batch	35
FAMEs	*Y. lipolytica*	120	2.47	0.132	0.020	55.8	Glucose	YNB/shake flask	36
FAMEs	*Y. lipolytica*	144	89.4	0.22	0.62	75.6	Glucose	YNB + 15.7 g/L YE/fed-batch	37
FAMEs	*Y. lipolytica*	84	98.9	0.269	1.300	66.8	Glucose	YNB + 3 g/LYE/fed-batch	This work

[a]The fermentation metrics from *R. toruloides* A4 strain represented the highest lipid titer and productivity among all the previously reported isolated microorganisms that have been cultured and fermented. [b]The fermentation characteristics are estimated from the plots provided in the original paper[34]. n/a, not available.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

strains yielded higher biomass density and lipid titers for the same aeration capacity, so that the fermentation results exceeded any other work reported in the literature (**Table 2**). This feature will facilitate scaling up in industrially relevant bioreactors where dissolved oxygen is the limiting factor of the process.

To quantitatively assess the efficiency of our synthetic pathways, we defined the fraction of the total carbon flux flowing through the synthetic pathway as Pathway Efficiency *E*, obtained from the following equation where $\mathbf{Y_m}$ and $\mathbf{Y_b}$, are, respectively, the maximum and baseline process yields.

$$E = \frac{Y - Yb}{Ym - Yb} \quad (4)$$

Using experimentally determined values for $\mathbf{Y_B}$ and **C**, upper and lower bounds can be calculated for $\mathbf{Y_L}$ assuming 100% and 0% capacity of the corresponding synthetic pathway. Alternatively, pathway efficiencies can be estimated for a given lipid pathway yield (**Table 1**) showing the efficiency of each synthetic pathway calculated from equation (4). Pathway efficiency is a useful concept to guide strain design toward achieving the full potential of the synthetic pathways (**Fig. 2c**). Stoichiometric analysis and flux balance calculations define the upper and lower bounds of the lipid pathway yields (red lines) and the yield model defines upper and lower bounds of the actual overall process yields (blue lines). The yields of various strain constructs fall between these bounds (circles in **Fig. 2c**). For example, **ADgapC** has an efficiency of 21%, which can be further improved by increasing the competitiveness of the synthetic pathways. At least three strategies can be applied to this end leading to overall process yield improvement: (i) overexpressing the committed enzyme GapC by doubling gene copy number (**ADgapc2**, $E = 29\%$); (ii) enlarging the pool size of the preferred cofactor of GapC–$NADP^+$ by overexpressing NAD^+ kinase (**ADgy**, $E = 64\%$); and (ii) introducing a second, functionally similar synthetic-pathway-activated POM cycle (**ADgm**, $E = 98\%$). Clearly, the synthetic pathways of GAPDH and POM cycle outcompete the native pathway and convert most of the cytosolic NADH to NADPH.

Our experiments indicate that **Y** was closely correlated with pathway efficiency (**Supplementary Fig. 11**), which was increased by improving synthetic pathway competitiveness. Even higher process yields are possible by deleting native pathways that compete with the synthetic bypasses, if this can be done without impairing cell growth and physiology. Further improvements of Y_L could be made by recycling mitochondrial NADH.m (**Supplementary Fig. 12**). By combining the genetic modifications that enable recycling of NADH.c and NADH.m, a maximum stoichiometric yield of 0.351 g/g, accounting for 96% of the maximum thermodynamic yield (0.364 g/g), was reached.

Here, we report that titer, yield and productivity of fatty acid methyl esters obtained using pure glucose are close to fulfilling the requirements (90 g/L, 1.3 g/L/h and 0.28 g-fatty acids/g-lignocellulosic sugars) determined by the National Renewable Energy Laboratory to support the US Department of Energy's 2017 cost goal of $5/gallon gasoline equivalent[30]. It will be important to identify whether other carbohydrate substrates from renewable feedstocks (sugarcane, lignocellulosic hydrolysates) can also be used to obtain equally high lipid yields with our engineered strains. One potential problem with these substrates is the presence of inhibitors derived from pretreatment. Besides sugar substrates, other substrates, such as, volatile fatty acids and products of anaerobic digestion and gas fermentation, could be used as substrates for our engineered strains. Redox engineering has previously been used in *Saccharomyces cerevisiae* but those efforts were not very successful, either because NADPH was not the limiting resource, or because *S. cerevisiae* cells resist changes in redox homeostasis. Indeed, we report here that some of our engineered pathways recycling cytosolic NADH did not improve process yields.

Redox engineering to improve productivity and yields could be applied to other bioproducts, such as, fatty-acid-derived oleochemicals (e.g., biosurfactants), and polyketides, isoprenoids and polyhydroxyalkanoates, which require ATP, NADPH and acetyl-CoA for biosynthesis (**Supplementary Fig. 13** and **Supplementary Table 4**).

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work is supported by the Department of Energy (grant DE-SC0008744). Some material from this paper has been included in a proposal submitted to the Department of Energy for a Bioenergy Research Center. The authors would like to acknowledge J. Shaw from Novogy, Inc. and H. Zhou, Woo-suk Ahn and A.Silverman from Massachusetts Institute of Technology for useful discussions and review of the manuscript.

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

AUTHOR CONTRIBUTIONS

K.Q. and G.S. conceived the project and wrote the manuscript. K.Q., T.M.W., K.Z. and P. X. designed and performed all the experiments. K.Q., T.M.W., K.Z., P.X. and G.S. analyzed the results.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc. Natl. Acad. Sci. USA* **103**, 11206–11210 (2006).
2. Sitepu, I.R. *et al.* Oleaginous yeasts for biodiesel: current and future trends in biology and production. *Biotechnol. Adv.* **32**, 1336–1360 (2014).
3. Jin, M. *et al.* Microbial lipid-based lignocellulosic biorefinery: feasibility and challenges. *Trends Biotechnol.* **33**, 43–54 (2015).
4. Beopoulos, A. *et al.* Yarrowia lipolytica as a model for bio-oil production. *Prog. Lipid Res.* **48**, 375–387 (2009).
5. Sheng, J. & Feng, X. Metabolic engineering of yeast to produce fatty acid-derived biofuels: bottlenecks and solutions. *Front. Microbiol.* **6**, 554 (2015).
6. Abghari, A. & Chen, S. *Yarrowia lipolytica* as an oleaginous cell factory platform for the production of fatty acid-based biofuel and bioproducts. *Front. Energy Res.* **2**, 21 (2014).
7. Dulermo, R., Gamboa-Meléndez, H., Ledesma-Amaro, R., Thévenieau, F. & Nicaud, J.M. Unraveling fatty acid transport and activation mechanisms in Yarrowia lipolytica. *Biochim. Biophys. Acta* **1851**, 1202–1217 (2015).
8. Dulermo, T. & Nicaud, J.M. Involvement of the G3P shuttle and β-oxidation pathway in the control of TAG synthesis and lipid accumulation in Yarrowia lipolytica. *Metab. Eng.* **13**, 482–491 (2011).
9. 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).
10. Dulermo, T. *et al.* Characterization of the two intracellular lipases of Y. lipolytica encoded by TGL3 and TGL4 genes: new insights into the role of intracellular lipases and lipid body organisation. *Biochim. Biophys. Acta.* **1831**, 1486–1495 (2013).
11. Blazeck, J. *et al.* Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production. *Nat. Commun.* **5**, 3131 (2014).
12. Beopoulos, A. *et al.* Control of lipid accumulation in the yeast Yarrowia lipolytica. *Appl. Environ. Microbiol.* **74**, 7779–7789 (2008).
13. Xue, Z. *et al.* Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nat. Biotechnol.* **31**, 734–740 (2013).
14. Qiao, K. *et al.* Engineering lipid overproduction in the oleaginous yeast Yarrowia lipolytica. *Metab. Eng.* **29**, 56–65 (2015).
15. 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).
16. Dellomonaco, C., Clomburg, J.M., Miller, E.N. & Gonzalez, R. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**, 355–359 (2011).
17. Xu, P. *et al.* Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nat. Commun.* **4**, 1409 (2013).
18. Li, Y.H., Zhao, Z.B. & Bai, F.W. High-density cultivation of oleaginous yeast Rhodosporidium toruloides Y4 in fed-batch culture. *Enzyme Microb. Technol.* **41**, 312–317 (2007).
19. Sauer, U., Canonaco, F., Heri, S., Perrenoud, A. & Fischer, E. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J. Biol. Chem.* **279**, 6613–6619 (2004).
20. Iddar, A., Valverde, F., Serrano, A. & Soukri, A. Expression, purification, and characterization of recombinant nonphosphorylating NADP-dependent glyceraldehyde-3-phosphate dehydrogenase from Clostridium acetobutylicum. *Protein Expr. Purif.* **25**, 519–526 (2002).
21. Martínez, I., Zhu, J., Lin, H., Bennett, G.N. & San, K.Y. Replacing *Escherichia coli* NAD-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH) with a NADP-dependent enzyme from Clostridium acetobutylicum facilitates NADPH dependent pathways. *Metab. Eng.* **10**, 352–359 (2008).
22. Verho, R. *et al.* Identification of the first fungal NADP-GAPDH from Kluyveromyces lactis. *Biochemistry* **41**, 13833–13838 (2002).
23. Zhang, H. *et al.* Regulatory properties of malic enzyme in the oleaginous yeast, Yarrowia lipolytica, and its non-involvement in lipid accumulation. *Biotechnol. Lett.* **35**, 2091–2098 (2013).
24. Liang, Y.J. & Jiang, J.G. Characterization of malic enzyme and the regulation of its activity and metabolic engineering on lipid production. *RSC Adv.* **5**, 45558–45570 (2015).
25. Li, Z. *et al.* Overexpression of malic enzyme (ME) of Mucor circinelloides improved lipid accumulation in engineered Rhodotorula glutinis. *Appl. Microbiol. Biotechnol.* **97**, 4927–4936 (2013).
26. Savitha, J., Wynn, J.P. & Ratledge, C. Malic enzyme: Its purification and characterization from Mucor circinelloides and occurrence in other oleaginous fungi. *World J. Microbiol. Biotechnol.* **13**, 7–9 (1997).
27. Zhang, S. *et al.* Engineering Rhodosporidium toruloides for increased lipid production. *Biotechnol. Bioeng.* **113**, 1056–1066 (2016).
28. Lee, J.M. *et al.* Cloning and characterization of the gene encoding phosphoketolase in Leuconostoc mesenteroides isolated from kimchi. *Biotechnol. Lett.* **27**, 853–858 (2005).
29. Hawkins, K.M. *et al.* Use of phosphoketolase and phosphotransacetylase for production of acetyl-coenzyme a derived compounds. US patent 20140273144 (2014).
30. Davis, R. *et al. Process Design and Economics for the Conversion of Lignocellulosic Biomass to Hydrocarbons: Dilute-Acid and Enzymatic Deconstruction of Biomass to Sugars and Biological Conversion of Sugars to Hydrocarbons* No. NREL/TP-5100–60223. (National Renewable Energy Laboratory (NREL), Golden, CO., 2013).
31. Zhang, F. *et al.* Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metab. Eng.* **14**, 653–660 (2012).
32. Leber, C., Polson, B., Fernandez-Moya, R. & Da Silva, N.A. Overproduction and secretion of free fatty acids through disrupted neutral lipid recycle in *Saccharomyces cerevisiae*. *Metab. Eng.* **28**, 54–62 (2015).
33. Runguphan, W. & Keasling, J.D. Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid-derived biofuels and chemicals. *Metab. Eng.* **21**, 103–113 (2014).
34. Liu, L., Pan, A., Spofford, C., Zhou, N. & Alper, H.S. An evolutionary metabolic engineering approach for enhancing lipogenesis in Yarrowia lipolytica. *Metab. Eng.* **29**, 36–45 (2015).
35. Liu, L. *et al.* Surveying the lipogenesis landscape in Yarrowia lipolytica through understanding the function of a Mga2p regulatory protein mutant. *Metab. Eng.* **31**, 102–111 (2015).
36. Silverman, A.M., Qiao, K., Xu, P. & Stephanopoulos, G. Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast Yarrowia lipolytica. *Appl. Microbiol. Biotechnol.* **100**, 3781–3798 (2016).
37. Zhang, S., Ito, M., Skerker, J.M., Arkin, A.P. & Rao, C.V. Metabolic engineering of the oleaginous yeast Rhodosporidium toruloides IFO0880 for lipid overproduction during high-density fermentation. *Appl. Microbiol. Biotechnol.* **100**, 9393–9405 (2016).

ONLINE METHODS

Construction of plasmids and *Y. lipolytica* strains. All the plasmids constructed in this study were validated by DNA sequencing and are summarized in **Supplementary Table 5**. All the primers used in this study were purchased from Integrated DNA Technologies and are listed in **Supplementary Table 6**. Plasmids were constructed exclusively by Gibson Assembly. Genomic DNA isolations from bacteria (*E. coli* and *Clostridium acetobutylicum*) and fungi (*Yarrowia lipolytica*) were performed using Wizard Genomic DNA purification kit according to manufacturer's protocol (Promega, USA). Total RNAs from *Mucor circinelloides* and *Y. lipolytica* were isolated using Ribopure-Yeast RNA Kit (Life Technologies) and RT-PCR was performed using ImProm-II Reverse Transcription Kit according to manufacturer's protocol (Promega, UAS). PCR was done using KAPA HiFi PCR Kit (KAPAbiosystems). Synthetic genes were codon-optimized using Optimizer[38] and assembled from the 500-bp or 1-kb DNA strings purchased from GeneArt (see **Supplementary Table 7** for detailed sequences). All the engineered *Y. lipolytica* strains were constructed by transforming the corresponding plasmids, which were linearized by restriction enzymatic digestion using either NotI or AseI, and were verified via PCR amplifications of the synthetic genes from the genomic DNA and DNA sequencing (Quintara, USA) of the corresponding amplicons. The transformation protocol has been reported previously[14].

The parent strain in this study is MTYL065, which was previously constructed by overexpressing acetyl-CoA carboxylase 1 (ACC1) and diacylglyceride acyltransferase 1 (DGA1) in the parent *Y. lipolytica* po1g (Yeastern, Taiwan). An auxotrophic marker URA3 was introduced by knocking out ~250 bp in the center of URA3 CDS region by knockout cassettes from pQkj1 (primers P2-P5 and P7-P8 used). The transformants were pre-selected using 5-fluoorotic acid followed by PCR verifications (primers P1 and P6 used). To express genes of interest in the strain ADura3, a compatible vector pQkj2 was created by combining TEF intron promoter and XPR2 terminator into pMT91 (primers P9-P12 used). The control strain AD was constructed by transforming pQkj2 into ADura3 to serve as a baseline control for all the experiments. Furthermore, another gene expression vector was created by combining *Y. lipolytica* GPD promoter and LIP1 terminator in cloning vector pUC19 to give pKJ1 (primers P18-P20 used). Plasmid pKJ1 was primarily used as a cloning vector to express a second gene of interest.

To construct *Y. lipolytica* strains with increased cytosolic NADPH availability, a series of genes, including UdhA and PntAB from *E. coli*, glyceraldehyde-3-phosphate (GAP) dehydrogenase from *C. acetobutylicum* and malic enzyme from *M. circinelloides* were cloned from the genomic DNA of *E. coli* and *C. acetobutylicum* and cDNA library prepared from mRNA extracted from *M. circinelloides*. The GAP dehydrogenase from *Kluyveromyces lactis* was codon-optimized and synthesized. All the genes were cloned into vector pQkj2 (primer pair P15-P16) to give, respectively, pQkj3, pQkj4, pQkj5, pQkj7 and pQkj6 (primers used: P13-14, P21-30, P33-P34, P35-P36 and P33-P34). The engineered *Y. lipolytica* strains ADudhA, ADpntAB, ADgapc, ADgpd and ADme were constructed by transforming linearized plasmids from pQkj3 to pQkj7, respectively. To activate the non-oxidative glycolysis pathway in yeast, phosphoketolase (PK) from *Leuconostoc mesenteroides* and phosphotransacetylase (PTA) from *Clostridium kluyveri* were synthesized and cloned into vector pQkj2 to yield pQkj8. The strain ADpp was created by integrating pQkj8 into ADura3.

To create the *Y. lipolytica* strains with enhanced GAP dehydrogenase activity, pQkj9 was constructed in pQK7 backbone (hygromycin resistance). The pQkj9 was transformed into ADgapc and ADgpd to afford ADgapc2 and ADgg, respectively. To functionally characterize endogenous NAD^+/NADH kinases from *Y. lipolytica*, three candidate genes in the *Y. lipolytica* genome were identified by BLASTX using three previously characterized *Saccharomyces cerevisiae* NAD^+/NADH kinase sequences (**Supplementary Table 2**). The corresponding genes encoding ylYEF, ylPOS5 and ylUTR1 were cloned into vector pQkj2 to yield pQkj10, pQkj11 and pQkj12 (primers used: P45-P50). Subsequently, ADyef, ADpos5 and ADutr were constructed. Finally, pQkj13 and pQkj14 (primers P27-P30, P51-P52) were individually constructed to co-express GapC and ylYEF; GapC and MCE2 in ADura3 (strains ADgy and ADgm). To characterize the GAP dehydrogenases GapC and ylGPD *in vitro*, the genes encoding GapC and ylGPD were cloned, respectively, into pMT15 vector with the C terminus 6× histidine tag to give pQkj15 and pQkj16.

Culture media, chemicals and conditions. *E. coli* was grown in Luria-Bertani (LB) medium with appropriate antibiotics at 37 °C/250 r.p.m. The antibiotics were added at the following concentrations: carbenicillin, 50 μg/mL; kanamycin, 50 μg/mL; chloramphenicol, 34 μg/mL. *Y. lipolytica* was selected in defined yeast nitrogen base (YNB) medium containing 6.7 g/L yeast nitrogen base w/o amino acids and 20 g/L glucose supplemented with appropriate concentrations of complete supplementary mixture (CSM) dropout mixtures, including CSM-Leu, CSM-Ura, CSM-Leu-Ura (Sunrise Science, USA) or 250 μL hygromycin at 30 °C/250 r.p.m. For genomic DNA extractions, bacteria were grown in LB medium, while all fungal strains were cultured in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose). 16 g/L Bacto agar was added for Petri dish cultivations (Becton, Dickinson, USA). All chemicals and substrates were purchased from Sigma-Aldrich unless otherwise indicated.

Shake flask culturing of engineered *Y. lipolytica* strains. A single colony of *Y. lipolytica* was grown in 2 mL defined medium (6.7 g/L yeast nitrogen base w/o amino acids, complete and 20 g/L glucose) at 250 r.p.m. for 30 h in 10 mL test tubes (Corning). The cells were harvested by centrifugation at 8,000 r.p.m. and washed twice with low-nitrogen defined medium (1.7 g/L yeast nitrogen base w/o amino acids and ammonium sulfate, 1.1 g/L ammonium sulfate and 50 g/L glucose). The washed cells were inoculated into 50 mL low-nitrogen defined medium (250 mL shake flask) at OD_{600} of 0.05 and grown at 30 °C/250 r.p.m. for 120 h. ~2 mL of cell suspension was sampled every 24 h (from 48 h to 120 h) for OD_{600}, dry cell weight, extracellular metabolite, and lipid measurements as described below. For enzymatic activity and cofactor quantification assays, ~10 mL cell suspension was harvested at exponential growth phase (24 h) and lipid production phase (60 h).

Bioreactor experiments for lipid production by engineered *Y. lipolytica*. The 3-liter Bioflo bioreactor (New Brunswick) was operated with 1.6-liter working volume for all the bioreactor experiments in this study. The seed cultures of *Y. lipolytica* were prepared by inoculating a single colony of *Y. lipolytica* into YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) and growing at 30 °C/250 r.p.m. for 30 h. The seed culture was harvested by centrifugation at 4,000 r.p.m., washed twice using fermentation medium (3.4 g/L yeast nitrogen base, 2.5 g/L yeast extract, 8.8 g/L ammonium sulfate, 100 g/L glucose) and inoculated into a bioreactor containing 1.6 liters fermentation medium. The starting OD_{600} of each bioreactor run was ~0.5. During the fermentation, oxygen was supplied by sterile filtered air at 3 v.v.m. (volume O_2 per volume liquid/min) and agitation speed was cascaded (200 r.p.m. to 750 r.p.m.) such that dissolved oxygen levels were maintained at 20% during growth phase (typically from 0 h to 36 h) and ~5% during lipid production phase (~36 h to the end of fermentation). The temperature was constantly controlled at 28 °C, and pH was maintained at 5.5 with the controlled addition of 6 M sodium hydroxide solution. During the course of fermentation, glucose concentration was monitored by high-performance liquid chromatograph (HPLC) analysis (see "Quantification of extracellular metabolites"), and the bioreactor was continuously supplemented with glucose. The concentration of glucose in the feed bottle was 600 g/L, and the feeding rate was controlled at 6.5 mL/h from 5 h to 55 h.

Protein purification and enzymatic activity of GAP dehydrogenases. Cells were harvested at exponential growth phase (36 h) from 50 mL of shake flask cultures by centrifugation at 8,000 r.p.m. for 2 min. The supernatants were removed and cell pellets were washed once with and resuspended in 10 mL SBE buffer (0.9 M sorbitol, 0.1 M EDTA and 0.8% β-mercaptoethanol) containing 50 mg zymolyase-20T (Sunrise Science Products). The cell suspension was incubated at 30 °C for 1 h and the *Y. lipolytica* protoplasts were harvested by centrifugation at 800 r.p.m. for 5 min followed by suspending in the phosphate buffer (50 mM NaH_2PO_4, 300 mM NaCl and 10 mM imidazole, pH = 7.5). Cells were disrupted by heavy-duty vertex with 0.7 mm-diameter acid-washed glass beads (Sigma-Aldrich) at 4 °C and the rate of 1,600 r.p.m. for 30 min. Cell debris was removed by centrifugation at 18,000*g* for 30 min at 4 °C. The clear supernatant was filtered through 0.45 um syringe filter and the 5× histidine-tagged proteins were purified using Ni-NTA Spin Kit (Qiagen) according to the manufacturer's protocol. The buffer of the final elution was exchanged with phosphate buffer (50 mM NaH_2PO_4 and 300 mM NaCl, pH = 7.5), and protein

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.

doi:10.1038/nbt.3763

samples were concentrated using Nanosep 10K omega (Pall) for enzymatic assays. The protein concentration was determined using Bio-Rad protein assay kit (Bio-Rad Laboratories). All the operations were carried out at 4 °C.

The enzymatic assays of GAP dehydrogenases-GapC and ylGPD were performed according to a previously established protocol[39]. Briefly, the enzymatic activity of GAP dehydrogenases was determined at 25 °C by monitoring the absorbance change of XTT (Sigma-Aldrich) reacted with NAD(P)H and phenazin methosulfate (Sigma-Aldrich) at 460 nm using a SpectraMax M2e Multi-Mode Microplate Reader. Initial rate measurements were carried out in assay buffer (50 mM triethanolamine hydrochloride, 0.2 mM EDTA, 50 mM Na_2HPO_4, pH = 8.5) in the presence of 2 mM glyceraldehyde 3-phosphate.

Quantification of cell density. Cell densities were monitored by measuring optical density at 600 nm wavelength and dry cell weight. For dry cell weight measurement, 800 µL cell suspension was harvested by centrifugation at 18,000*g* for 15 min in preweighed microcentrifuge tubes. The cell pellets were washed twice in water and dried at 60 °C until the mass of each tube remained constant over time (typically after ~36 h).

Extractions, derivatizations and quantifications of lipids. Depending on cell density, 50 µL to 1 mL of cell suspension (~2 mg dry cell weight) was sampled and centrifuged at 18,000*g* for 10 min and the supernatant was carefully removed. The cell pellets can be stored in a –20 °C freezer until further derivatization. For lipid analysis, 100 µL hexane containing 2 mg/mL methyl tridecanoate (internal standard for volume change) and 2 mg/mL glyceryl triheptadecanoate (Internal standard for transesterification efficiency) was added to the cell pellet. Lipid transesterifications were initiated by addition of 500 µL 0.5 N sodium methoxide (20 g/L sodium hydroxide in anhydrous methanol) followed by vortexing at 1,200 r.p.m. for 60 min at room temperature. Then the samples were neutralized with 40 µL sulfuric acid (98% purity) and the synthesized FAMEs were extracted with 500 µL hexane by vortexing for an additional 30 min at 1,200 r.p.m. The samples were centrifuged at 8,000 r.p.m. for 1 min and 1 µL of the top hexane layer was analyzed by a Bruker 450-GC Gas Chromatograph equipped with a flame ionization detector (GC-FID). The sample was injected into a HP-INNOWAX capillary column (Agilent Technologies, USA) with split ratio of 10 and injector temperature of 260 °C. The flowrate of carrier gas was 1.5 mL/min and the oven temperature was held at a constant temperature of 200 °C for the duration of 13 min to analyze all five major FAME species (C16, C16:1, C18, C18:1 and C18:2).

Quantification of extracellular metabolites. At the indicated time points, ~500 µL of cell suspension was sampled and centrifuged at 16,000*g* for 10 min. The supernatant was further filtered with a 13-mm syringe filter with 0.2 µm PTFE membrane (VWR international). 10 µL of the filtered supernatant was analyzed using high-performance liquid chromatography (Agilent 1,200 HPLC system equipped with G1362A Refractive Index Detector) to quantify the concentrations of metabolites, including glucose, citrate, mannitol, glycerol and erythritol. The mobile phase (14 mM sulfuric acid) was used to flow through a separation column (Bio-Rad HPX-87H column) at a rate of 0.7 mL/min.

Statistical analysis. Sample size was chosen based on the experience with similar experiments from other research groups in the field of metabolic engineering of microbial workhorses. No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. Differences between the performances of control and engineered strains were analyzed using Student's *t*-test as indicated in the corresponding figure captions. $P < 0.05$ was considered significant.

38. Puigbò, P., Guzmán, E., Romeu, A. & Garcia-Vallvé, S. OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res.* **35**, W126–W131 (2007).
39. Bommareddy, R.R., Chen, Z., Rappert, S. & Zeng, A.P. A de novo NADPH generation pathway for improving lysine production of Corynebacterium glutamicum by rational design of the coenzyme specificity of glyceraldehyde 3-phosphate dehydrogenase. *Metab. Eng.* **25**, 30–37 (2014).

© 2017 Nature America, Inc., part of Springer Nature. All rights reserved.