

Engineering de novo formatotrophy in the non-model yeast *Y. lipolytica*

Received: 18 February 2025

Accepted: 14 April 2026

Published online: 12 June 2026

 Check for updatesWilliam Newell¹, Piotr Hapeta¹, David Bell^{2,3,4} & Rodrigo Ledesma-Amaro^{1,2,3} 

Formate is an exciting potential microbial feedstock as it can be derived from CO₂ and electricity. Despite this, limited progress has been made in engineering formatotrophy in yeasts, and no yeasts grow using formate naturally. Here we use metabolic modelling to find two potential formatotrophy pathways in *Yarrowia lipolytica*. We then use C13 tracer analysis and computationally guided growth experiments to show that wild-type *Y. lipolytica* possesses strong formate dissimilation and a cyclical C1 pathway with similar architecture to the synthetic serine–threonine cycle, which it uses to co-assimilate formate and glycerol. Messenger RNA sequencing shows that formate exposure results in increased oxidative stress and changes in the tricarboxylic acid cycle, redox and C1 metabolism. Following this, we use model-guided adaptive laboratory evolution to produce a formatotrophic strain of *Y. lipolytica* using the eukaryotic serine–threonine cycle. We then use further messenger RNA sequencing to show that formatotrophy is supported by changes in adenosine triphosphate and reactive oxygen species metabolism. Subsequently, we engineer nicotinamide adenine dinucleotide phosphate (NADPH) and reactive oxygen species metabolism to create a strain with substantially improved growth. This strain reaches about 10% of the theoretical maximum biomass yield, highlighting its potential for additional engineering approaches. Finally, we show that beta-carotene production from formate is possible in our engineered strain, opening the door to formatotrophic eukaryote bioprocesses.

The ongoing climate crisis demands substantial changes in the ways fuel, food and building materials are produced. Biomanufacturing is well suited to become a large part of this solution, but is constrained by the limited sustainability of conventional sugar-based feedstocks^{1,2}. Formic acid can be synthesized at very high efficiency by electrically reducing CO₂, making it an ideal sustainable feedstock for bioprocesses³.

Despite this potential, progress has been limited in industrializing formate-based bioprocesses. On the one hand, microorganisms with endogenous formate assimilation pathways have been explored as production hosts. However, many are difficult to engineer and have

not yet shown the yields needed for industrial-scale production^{4–6}. Others, such as acetogens, are strict anaerobes, limiting which products can be made⁷. Several bioproducts have been made by formate conversion via the Calvin cycle, but this is less efficient than direct formate assimilation^{8–10}.

On the other hand, some headway has been made in engineering well-established industrial prokaryotes to grow on formate. *Escherichia coli*, *Pseudomonas putida* and *Cupriavidus necator* have all been engineered to grow on formate via the highly efficient reductive glycine pathway (RGP)^{11–13}, with *E. coli* also engineered to use the

¹Department of Bioengineering and Imperial College Centre for Engineering Biology, Imperial College London, London, UK. ²Bezos Centre for Sustainable Protein, Imperial College London, London, UK. ³UKRI Engineering Biology Mission Hub on Microbial Food, Imperial College London, London, UK. ⁴London Biofoundry, Imperial College London, London, UK. ✉e-mail: r.ledesma-amaro@imperial.ac.uk

serine–threonine cycle¹⁴. Production of lactic acid and polyhydroxybutyrate from formate was recently demonstrated in *E. coli* at relatively high yields, showing that non-native assimilation pathways can support effective bioproduction¹⁵. Prokaryotes, however, may suffer from scale-up issues such as susceptibility to extremes of pH¹⁶. Meanwhile, eukaryotes sometimes offer greater robustness and can be used to produce different compounds from prokaryotic hosts¹⁷.

However, eukaryotes have proved much more difficult to convert into formatotrophs, and no eukaryote is known to grow on formate as sole carbon source naturally. RuBisCo has been functionally expressed in *S. cerevisiae* and combined with a formate dehydrogenase (FDH) to supply roughly 10% of required carbon, but full formatotrophy was not observed¹⁸. *Komagataella phaffi* has been engineered to grow via the RGP¹⁹, however, growth is slow. *Saccharomyces cerevisiae* has also been evolved to grow on formate via the RGP during efforts to create a RuMP-cycle using methylotroph²⁰. However, formatotrophy via other pathways has not been observed and eukaryotic formatotrophy remains poorly understood.

Yarrowia lipolytica is a yeast that has shown great potential for biomanufacturing¹⁷. Industrial-scale culture to produce a wide spectrum of organic acids and food additives has been successfully accomplished and alkane, alkene and terpenoid production has proved possible^{17,21,22}. Unlike many microbes, *Y. lipolytica* is tolerant to extremes of pH, low temperatures, high concentrations of inhibitory compounds and high salt concentrations^{17,23}. Furthermore, a strong endogenous flux towards lipid biosynthesis and NADPH and a high tolerance to expression burden also strengthens its bioproduction ability²⁴. Therefore, engineering *Y. lipolytica* to grow on formate could provide a way to convert electricity and CO₂ to industrial products at a much larger scale than currently possible. *Y. lipolytica* has been shown to use formate as a source of NADH during growth on glucose²⁵ and formate assimilation via a cyclical assimilation pathway has been proposed on the basis of transcriptomics²⁶. In this work, we use computationally guided metabolic engineering and adaptive laboratory evolution to create a formatotrophic strain of *Y. lipolytica*.

Results and discussion

In silico analysis of cryptic formate metabolism

We initially used genome-scale metabolic modelling to find possible cryptic formate-assimilating pathways in *Y. lipolytica*. Simulating formate uptake showed that cytoplasmic methyl transfer, anapleurosis and threonine biosynthetic reactions could be patched together to assimilate formate via a cyclic pathway connecting serine and threonine (Fig. 1). Here formate is attached via 5,10-methylene-tetrahydrofolate (mTHF) to glycine to form serine. Serine is then recycled back to glycine via pyruvate and threonine. Glycine can also be synthesized using formate-derived mTHF by the RGP, via the reverse glycine cleavage system²⁷. Therefore, we added this reaction to the model, which then permitted growth via the linear RGP (Fig. 1). In both cases, the model regenerates NADH and generates CO₂ simultaneously using *Y. lipolytica*'s 11 FDHs. An analysis of the distribution of serine–threonine cycle and RGP genes using FUNGIpath showed each were widely distributed (59% and 75%) in the 166 genomes covered.

Co-assimilation of formate during growth on glycerol

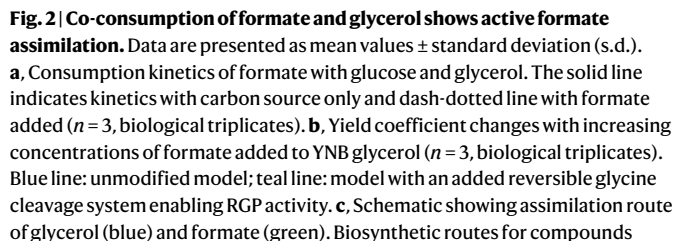
Growth on formate as a sole carbon and energy source has never arisen spontaneously in any eukaryote, including *Y. lipolytica*. However, *Y. lipolytica* has been shown to consume formate under several conditions^{25,26}. Therefore, we wondered whether *Y. lipolytica* host cryptic formate assimilation pathways that could be awoken with adaptive evolution. To explore formate metabolism in greater detail, we measured biomass and substrate consumption kinetics during growth on glucose and glycerol with and without formate supplementation (Fig. 2a). Cell dry weight (CDW) accumulation appeared roughly correlated with carbon source use in both cases, while formate use

was roughly constant throughout growth. Biomass yield on glycerol was similar to that on glucose (Supplementary Fig. 2). However, formate addition to glycerol increased CDW yield by more than glucose: around 80% for glycerol but around 50% on glucose (Fig. 2a and Supplementary Fig. 2). We therefore decided to investigate growth on glycerol and formate in more detail. Modelling suggested that formate was used as both a supplementary carbon source via ligation to THF and a source of reducing power for NADH regeneration by FDH (Supplementary Dataset 3). These simulations showed that if formate was only used as an energy source, CDW yield plateaus at around 0.8, as growth becomes limited by the availability of carbon in the media. However, simulating use as a carbon and energy source predicts that yield can increase indefinitely if formate can be incorporated as a carbon source. When we contrasted growth via the serine–threonine cycle with the RGP, we saw that the RGP permits higher yields, probably due to the lower ATP demand. This is reflected by the higher biomass yield predicted on formate (Supplementary Fig. 1). When *Y. lipolytica* is grown on glycerol, yields match curves predicted for formate incorporation into biomass (Fig. 2b). The lower yields at the highest formate ratio for the higher glycerol concentration probably represent formate toxicity, as the concentration here is 1.5 M: around which *S. cerevisiae* also experiences impaired growth²⁸. Modelling predicted that during growth via the serine–threonine cycle, formate enters metabolism via THF-mediated attachment to glycine synthesized from glycerol (Supplementary dataset 1). This forms serine, which is then converted into threonine via pyruvate carboxylation and aspartate biosynthesis. Threonine is then cleaved into glycine and acetaldehyde, which then enters the tricarboxylic acid cycle (TCA) as acetyl-CoA after conversion by acetaldehyde dehydrogenase and acetate kinase (Fig. 2c). Our analysis of C13-labelling patterns showed that serine-derived metabolites (serine, threonine, aspartate, asparagine, uracil, glycerophosphoethanolamine) were once labelled. Other mTHF-derived amino acids (histidine, methionine) were once labelled, suggesting that formate enters via the THF cycle. Adenine monophosphate (AMP) and xanthine were m + 1 and m + 2 labelled, reflecting the two molecules of mTHF used in their biosynthesis (Fig. 2c). Signs of labelled glycine were not detected: serine was only labelled once and m + 3 glutathione (GSH), AMP and xanthine, which would stem from labelled glycine, were only detected in small amounts suggesting that negligible amounts of single-labelled glycine were produced. This suggests that the RGP, which synthesizes glycine from formate, is not active (Supplementary Note 1) and instead formate enters via the serine–threonine cycle.

mRNA sequencing shows formate-responsive transcriptional response

We next used messenger RNA (mRNA) sequencing to explore the changes in gene expression during growth on glycerol and formate. Sequencing showed widespread differences in gene expression when formate was added (Supplementary Fig. 4). As a result, we decided to focus on metabolic genes specifically, as well as focusing on highly upregulated or downregulated genes (Fig. 3). Some of these changes reflect variations in C1 metabolism. NADH is regenerated by formate oxidation via FDH, and indeed four FDH isozymes were strongly upregulated (Fig. 4b) when formate was added. We also observed strong upregulation of NQO, coding for NADH-quinone oxidoreductase. This provides another route for electrons to enter the respiratory chain and is used to restore NADH/nicotinamide adenine dinucleotide ratios under strongly reducing conditions^{29–31}. TCA cycle genes coding for malic enzyme, isocitrate dehydrogenase, pyruvate dehydrogenase subunits PDC1 (E1 subunit) and Lat1 (E2 subunit) also all fell in expression level. Both TCA cycle flux and several TCA cycle genes are downregulated when *Y. lipolytica* is grown on more reducing substrates due to higher levels of NADH generated by a more reducing cellular redox status³⁰. TCA downregulation could, therefore, result from NADH





C13 tracer analysis confirmed our initial modelling prediction: that growth on formate follows the serine–threonine cycle. In this case, the single-labelled serine that we observed can only arise if glycine is recycled via the serine–threonine cycle; likewise, twice labelled xanthine suggests formate assimilation via THF does not from mTHF and CO₂ in the RGP, which would create double labelled serine and triple labelled xanthine. Similar to growth on glycerol and formate, we observe carry-through of labelled carbon to aspartate, threonine, branched chain amino acids and TCA metabolites (Fig. 4c), all of which indicate formate-derived carbon being converted to acetyl-CoA via threonine biosynthesis. Modest amounts of m + 2 labelled citrate show that 2 molecules of acetyl-CoA are assimilated via the glyoxylate cycle. This contradicts the cyclic assimilation pathway in *Y. lipolytica* proposed by one research group, where glyoxylate is predicted to regenerate glycine as opposed to being used to generate malate²⁶. Single-labelled phenylalanine and glucose also demonstrate that the synthesis of gluconeogenesis metabolites also occurs from formate. *Y. lipolytica* accumulates very large amounts of lipids and carbohydrates⁴¹, which could support growth without formate assimilation and may dilute our C13 labelling signal. Indeed, the non-evolved control was able to sustain roughly two doublings on formate in a single passage (Supplementary Fig. 5). Therefore, we grew NYF1 on formate-only media over multiple passages. We observed consistent formatotrophic growth, and our strain failed to grow without a carbon source, making growth on intracellular storage metabolites alone unlikely (Supplementary Figs. 6 and 10).

We also saw that aspartate aminotransferase 1, coding for cytoplasmic aspartate aminotransferase, was upregulated (Fig. 5a). Aspartate aminotransferase interconverts oxaloacetate and aspartate, allowing electrons to be shuttled into the mitochondria via the aspartate–malate

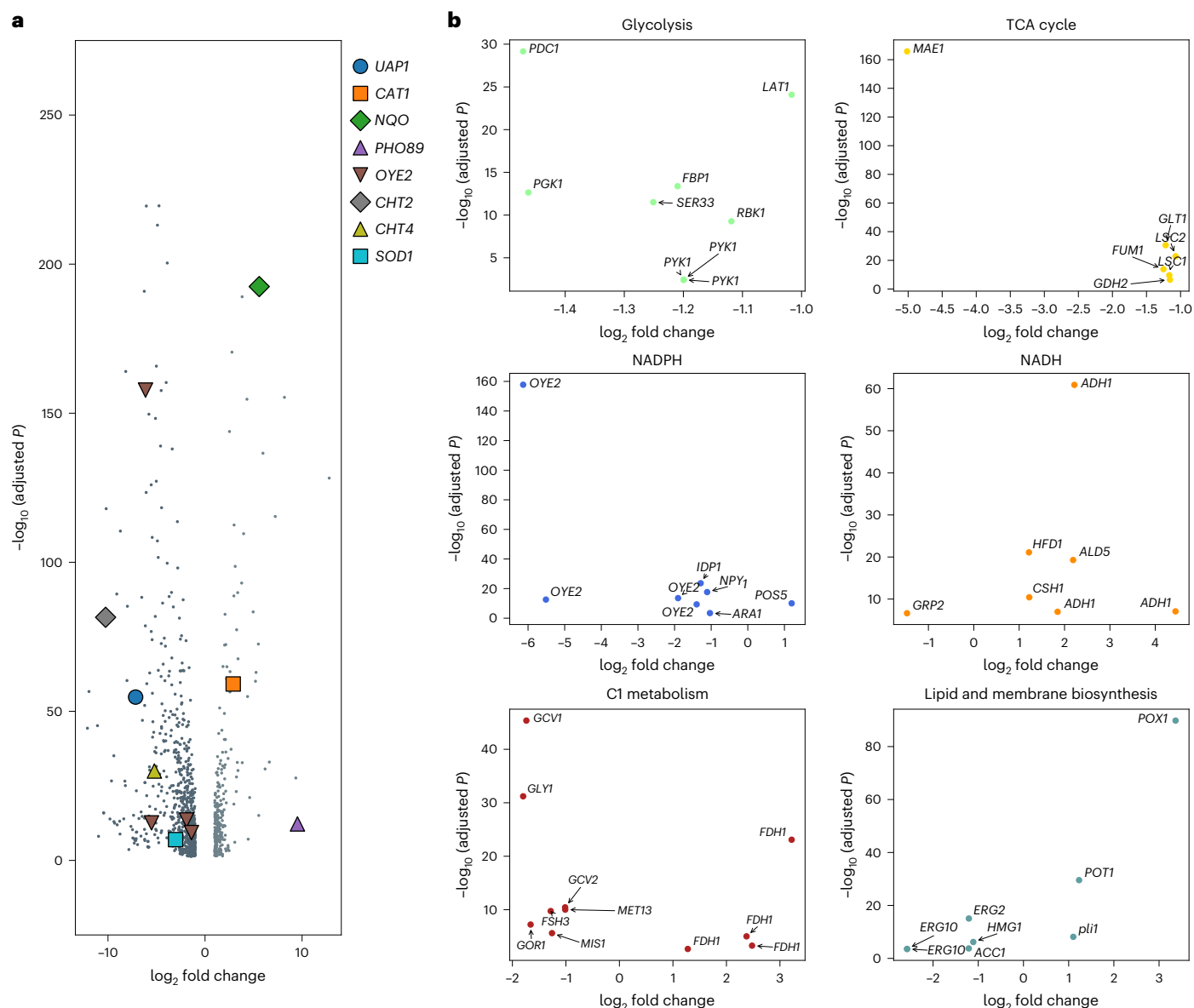


Fig. 3 | Transcriptomic changes in *Y. lipolytica* during formate exposure.

a, Selected genes with added formate. **b**, Transcriptomic analysis of differentially expressed metabolic genes involved in major metabolic pathways: glycolysis—conversion of upper metabolism to pyruvate; TCA cycle—conversion of

pyruvate to TCA intermediates and assimilation of NH_3 via glutamate; NADPH/NADH—NAD(P)H-interconverting oxidoreductase reactions; C1—single carbon metabolism genes and lipid; and membrane biosynthesis—lipid anabolism or catabolism and mevalonate pathway genes.

shuttle⁴³. It is also required to generate aspartate in the serine–threonine cycle. Both processes should require high flux, requiring a greater amount of aminotransferase.

We also observed a clear shift towards ROS resistance, showing that, once again, formate-driven ROS production was a substantial challenge to cells. NYF1 appeared to have evolved higher resistance to ROS. For instance, expression levels of genes IML2 (encoding a DNA oxidative damage-repair protein) and Cat1 (catalase) fell suggesting lower oxidative stress (Fig. 5a). Expression levels of differentially expressed genes between Po1D and NYF1 growing in formate were also consistent with evolution towards lower levels of oxidative stress. For one, AMO1 (encoding a broad specificity amine oxidase) and various homologues of DUR3 (urea transporters) were downregulated. These participate in the allantoin degradation pathway: a product of oxidative damage of purines and a biomarker for oxidative stress⁴⁴ (Fig. 5d). Amine oxidase may also show xanthine oxidase activity, a known source of ROS that may need to be downregulated to allow lower ROS generation^{45,46}.

A cytosolic superoxide dismutase (SOD1) was also increased in expression compared with Po1D. SOD1 catalyses the conversion of superoxide radicals to water⁴⁷ (Fig. 4d), which are produced when electrons ‘leak’ from the electron transport chain during inhibition of respiratory complexes⁴⁸ (Fig. 5a). Upregulation of SOD1 should therefore counteract radical production caused by formate inhibition of complex IV.

We then measured relative ROS production using a 2',7'-dichlorofluorescein diacetate (DCFDA) assay and indeed found that NYF1 showed lower ROS production, both during growth on glycerol and when formate was added (Fig. 5b), suggesting that NYF1 had evolved better ROS degradation. A ROS-resistant phenotype could also improve mitochondrial function during formate exposure. Radical oxygen species promote mitochondrial depolarization and loss of function, decreasing ATP production⁴⁹. We used Mitotracker Deep Red, a dye that binds to active mitochondria⁵⁰, to assess whether NYF1 had more functional mitochondria. NYF1 showed a higher fluorescence at 535 nm than Po1D during growth on glycerol with formate, suggesting

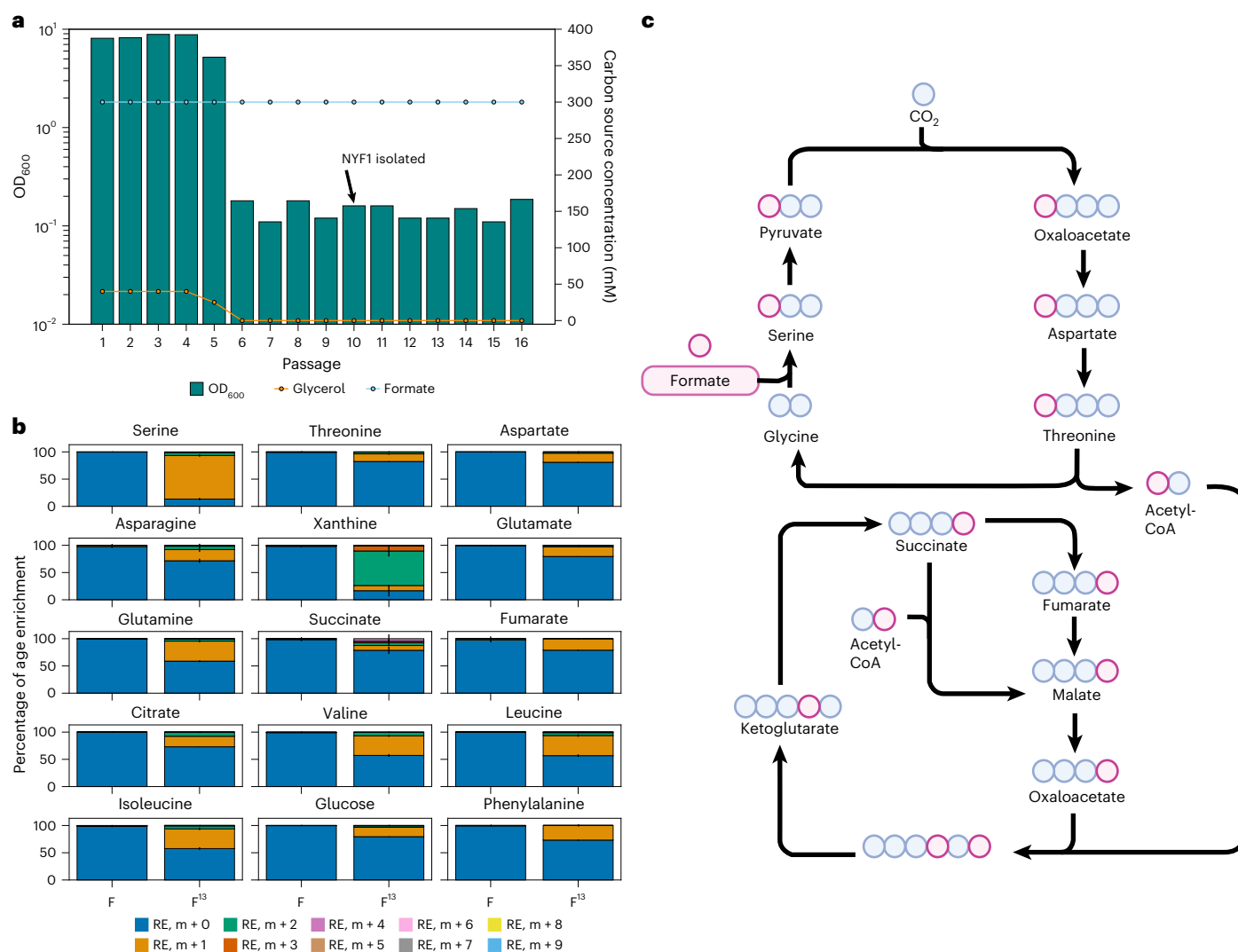


Fig. 4 | Evolution of a formatotrophic *Y. lipolytica* strain. a, Summarized adaptive laboratory evolution of NYF1. Teal bars indicate OD, blue and orange markers indicate formate and glycerol concentrations. **b**, C13 tracer and amino

acid labelling in strain NYF1. Data are presented as mean values $\pm$ s.d. ($n = 3$, biological triplicates). **c**, Schematic showing incorporation of formate (purple) and CO₂ (blue) into semi-labelled biomass during single passages on formate.

that mitochondria were indeed more active (Fig. 5c). Furthermore, NYF1 shows higher expression of a homologue of PET9, which encodes for a cytosolic ADP–mitochondrial ATP exchanger, which implies greater ATP synthesis⁵¹ (Fig. 5a). Together, these findings suggest that NYF1 has evolved greater degradation of ROS, leading to lower oxidative stress and preserving mitochondrial function when formate is added.

Balancing NADPH and ROS metabolism improves formatotrophic growth

We then attempted to improve NYF1's growth on formate. Extensive experiments to optimize culture conditions showed that NYF1 grew best in shake-flasks, whereas altering formate concentration, phosphate concentration, initial OD or CO₂ availability did not change growth on formate (Supplementary Figs. 5 and 11).

We then decided to explore other strategies to improve growth. Substantial amounts of NADPH are required during growth on formate: modelling showed that almost half of all formate consumed is used to regenerate NADPH during simulated growth on formate (Supplementary Fig. 12). While several endogenous routes are available to convert electrons into NADPH, all require ATP to be spent⁵², potentially increasing the demand on oxidative phosphorylation

systems already-taxed by formate poisoning. We therefore decided to express the NADPH-dependent FDH from *Burkholderia stabilis*⁵³ (FDHbs), which converts formate into NADPH without spending ATP. This approach should also improve growth: our modelling showed that avoiding this unnecessary energy sink increases yields by around 20% (Supplementary Fig. 12).

Despite this, expressing FDHbs in NYF1 failed to improve growth on formate (Supplementary Fig. 13). A highly reduced NADPH pool can lead to 'reductive stress', where electrons leak from reduced NADPH-dependent oxidoreductases and form radical oxygen species⁵⁴. Therefore, FDHbs expression could further increase oxidative stress, inhibiting growth. When we measured ROS production and mitochondrial activity in NYF1 + FDHbs, we observed the same trend of higher oxidative stress accompanying lower mitochondrial function as previously (Fig. 6b,c). FDHbs expression seemed to have a mild pro-oxidant effect even without formate addition, potentially the result of NADPH production from intracellular formate synthesis from the THF cycle (Supplementary Fig. 15 and Supplementary Note 4).

We then decided to test the effects of adding butylated hydroxytoluene (BHT). BHT has been used to reduce oxidation in *Y. lipolytica* cultures and is a known ROS scavenger^{55,56}. BHT addition showed reduced oxidative stress in both NYF1 and FDHbs (Supplementary Fig. 15).

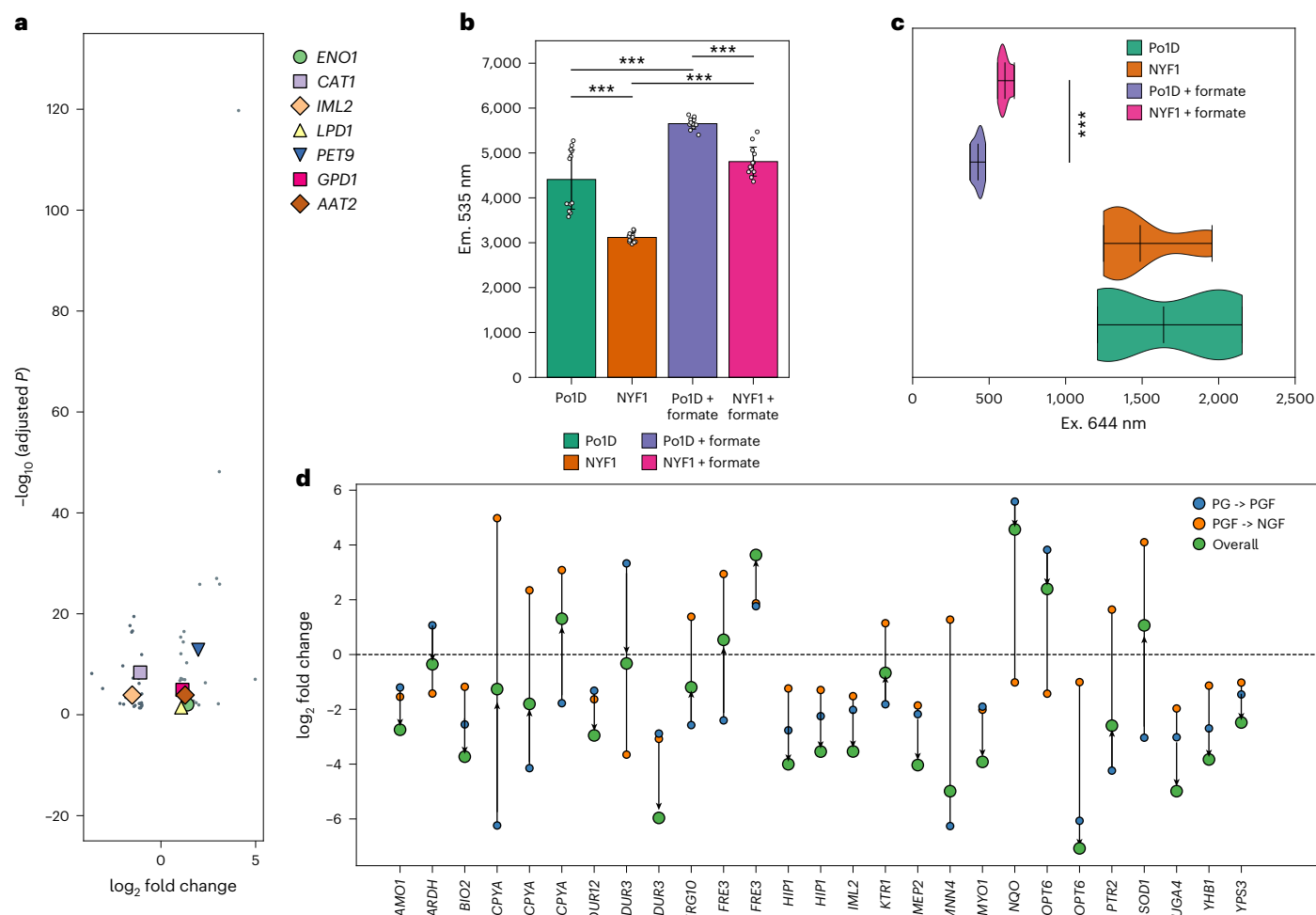


Fig. 5 | Transcriptomic and physiological changes in NYF1. a, Selected differentially expressed genes between NYF1 and Po1D in glycerol and formate. **b**, DCFDA assay of ROS in Po1D and NYF1 grown in glycerol with added formate (results of a two-sample paired t -test: $P = 6.78 \times 10^{-5}$, 4.55×10^{-5} , 4.40×10^{-10} , 6.93×10^{-7}). Data are presented as mean values $\pm$ s.d. ($n = 9$, biological triplicates or technical triplicates). **c**, Violin plot of fluorescence at 644 nm of mitotracker-stained cells. Left and right bounds of violin shows maximum and minimum, central line indicates mean ($n = 30$, biological triplicates; $P = 7.72 \times 10^{-11}$).

d, Transcript changes in differentially expressed genes shared between Po1D and NYF1. Blue marker: fold change in Po1D when formate is added. Orange marker: fold change between gene in NYF1 and Po1D. Green marker: total change from Po1D grown in glycerol to NYF1 grown in formate. Black arrow: direction of overall change; upwards: positive change; downwards: negative. Horizontal line indicates expression level of genes in Po1D grown in glycerol. All P values are summarized as * $P > 0.05$, ** $P > 0.01$, *** $P > 0.005$. Ex., excitation; Em., emission.

It also resulted in better growth for NYF1 + FDHs, with a final OD roughly double that of NYF1 (Fig. 6a). This did not occur with NYF1, suggesting that lowering ROS was required to exploit increased cellular reducing power.

Menadione stimulates superoxide formation and is known to induce ROS-tolerance in *Y. lipolytica* when added to precultures by increasing the expression of oxidative stress resistance genes⁵⁷. Therefore, menadione addition in precultures of NYF1 + FDHs should improve growth. Indeed, doing so showed the same pattern of better growth as seen with other antioxidant strategies (Supplementary Fig. 14).

On this basis, we experimented with overexpressing superoxide-degrading enzymes in *Y. lipolytica*. Several of these are available: superoxide dismutase 2 (SOD2) (mitochondrial) and superoxide dismutase 3 (SOD3) (cytoplasmic) catalyse the conversion of superoxide to water, while reduced GSH catalyses the degradation of a wide range of radical species, including oxygen radicals, lipid peroxyls produced by oxidative damage and hydrogen peroxide⁵⁷ (Fig. 6a). We then expressed SOD2, 3, and the rate-limiting step for GSH production⁵⁸ (glutathione reductase, GSRI) in NYF1. Once more, we observed

better growth in NYF1 + FDHs for all three enzymes (Fig. 6d). We then selected the two best growing strains, FDHs + SOD3/GSRI, and measured mitochondrial function and ROS production. This showed that both strategies greatly reduced ROS production (Fig. 6b) with FDH expression. Mitochondrial function was higher when both antioxidant enzymes were expressed, with SOD3 maximizing mitochondrial output (Fig. 6c).

Finally, we experimented with producing beta-carotene from formate. Beta-carotene is of considerable nutraceutical interest as an antioxidant supplement, as well as being a useful benchmark for bioproduction from acetyl-CoA^{59–61}. Expressing carotenogenesis genes in NYF1 allowed B-carotene production from formate (Fig. 7d and Supplementary Fig. 16). Expression of carotenogenesis genes alongside FDHs led to a similar increase in final OD to that seen with BHT addition (Fig. 6b). Beta-carotene is known to act as a radical scavenger, which probably suppresses ROS-formation in a similar manner to BHT and so allows better growth. We then used brightfield and fluorescence microscopy to image cells during growth on a relaxing carbon source (glycerol) and during formatotrophic growth (Fig. 6c). This showed that beta-carotene was localized to lipid bodies commonly

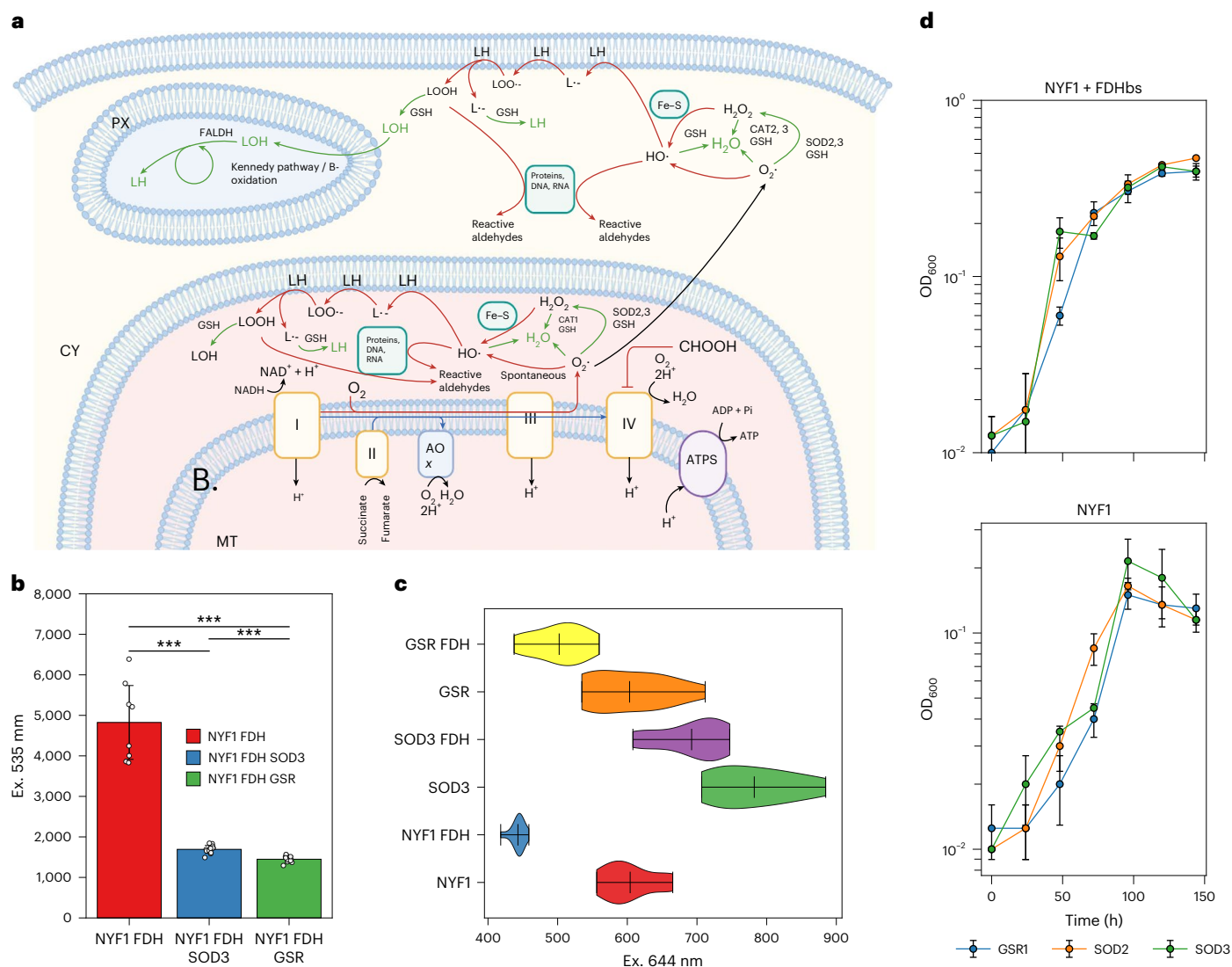


Fig. 6 | Engineering the ROS metabolism of *Y. lipolytica* improves growth on formate. Data presented as mean $\pm$ s.d. **a**, Schematic showing mechanisms of ROS-related damage in *Y. lipolytica*. MT, mitochondria; CY, cytoplasm. Blue arrow along inner mitochondrial membrane: normal flow of electrons. Red inhibitor line at complex IV: inhibitory effect of formate binding. Red arrows: flow of electrons to O_2 to form superoxide and superoxide reactions to form hydroxyl radicals, then further reactions with lipids (LH) to form lipyl radicals and lipid peroxides (L- and LOOH). Green arrows: detoxification pathways to form water, lipid alcohols (LOH), short lipids (LH) or short

aldehydes. **b**, DCFDA assay of ROS in Po1D and NYF1 grown in glycerol with added formate ($n = 9$, biological triplicates or technical triplicates; *** $P > 0.001$; $P = 7.76 \times 10^{-11}$, 1.67×10^{-11} , 1.29×10^{-6}). **c**, Violin plot of fluorescence at 644 nm of mitotracker-stained cells. Left and right bounds of violin shows maximum and minimum, central line indicates mean ($n = 30$, biological triplicates; $P = 8.23 \times 10^{-13}$, 1.62×10^{-14} , 2.58×10^{-5} , 4.61×10^{-4} , 3.91×10^{-5}). **d**, Growth curves of NYF1 or NYF1 + FDHbs expressing antioxidant enzymes ($n = 2$, biological triplicates).

seen with beta-carotene in *Y. lipolytica*⁶². Production was high overall, with a high proportion of cell mass made up of accumulated carotenes (Supplementary Fig. 16). Carotene production per cell was lower in the better growing FDHbs strain, suggesting that there may be a trade-off between carotene and biomass production.

While our modelling shows that a biomass yield of around 6 g l^{-1} is possible with the serine–threonine cycle, our strain achieved lower CDW concentrations, around 0.6 g l^{-1} (Supplementary Fig. 1). This reflects the difficulty encountered when engineering C1-trophic metabolism. For example, when one research group replaced the Calvin cycle with the RGP, their initial formatotrophic strain only achieved around 50% of the theoretical maximum due to non-optimized gene expression and unexpected metabolic side reactions¹². Therefore, a long road lies ahead, with further engineering required to balance the complex network supporting formatotrophic metabolism

Conclusions and perspective

Y. lipolytica is a yeast with considerable industrial applications, and here we have evolved it to grow on formate. We observed growth via the eukaryotic serine–threonine cycle and noted that formatotrophy arose via evolution alone. In two other examples of eukaryotic formatotrophy, poor energy generation was suggested to be a limiting factor of growth on formate^{19,20}. In this case, *Y. lipolytica*'s ability to increase the expression of several FDH genes probably allows it to generate large amounts of NADH to sustain energy-hungry formate assimilation, making it a strong prospective platform for C1 bioprocesses.

Moreover, *Y. lipolytica* appears to be a cryptic formatotroph; no genes involved in formate assimilation were upregulated after formatotrophy evolved and growth on formate and glycerol was highly efficient, implying that existing metabolism could support formate incorporation. ROS degradation is consistently involved with

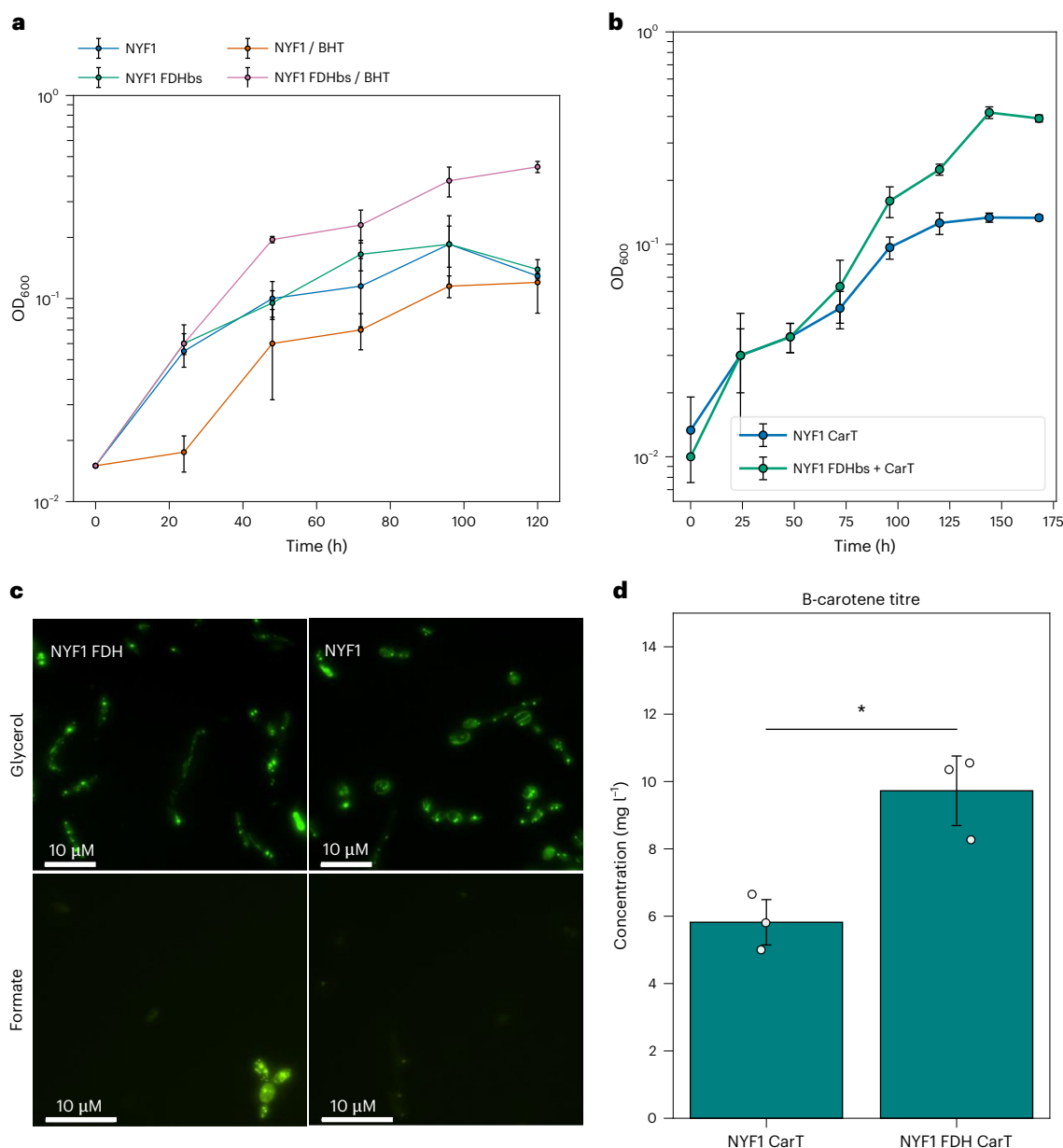


Fig. 7 | Carotene production from formate. All data are presented as mean values $\pm$ s.d. **a**, Growth curve of NYF1 and NYF1 expressing FDH from *B. stabilis* with added BHT. **b**, Growth curve of NYF1 and NYF1 FDH expressing carotenogenesis cassette CarT. **c**, Fluorescence micrograph of

NYF1 + carotenogenesis + FDH from *B. stabilis* (left) and NYF1 + carotenogenesis (right), grown on glycerol and formate (upper) or formate alone (lower). **d**, Titre and yield of carotenes produced by strain NYF1 and NYF1 with FDHbs expressing carotenogenesis genes (CarT; $P = 0.012$, $^*P > 0.05$).

supporting growth on formate. Given that formate is common in the environment⁶³ and a potent respiratory chain inhibitor³⁸, this cryptic metabolism could represent a detoxification mechanism that allows *Y. lipolytica* to avoid poisoning by assimilating environmental formate to biomass. Finally, we show that metabolic engineering of cofactor metabolism and oxidative stress resulted in a twofold higher final OD, as well as allowing strong carotene biosynthesis. Efforts to increase cellular reducing power—important during growth on oxidized carbon sources—appeared to generate further oxidative stress. This seemed to counteract the benefits of better cofactor regeneration unless ROS production was countered by improved antioxidant metabolism. With these findings in place, future research efforts should focus on improving ATP generation, ROS resistance and maximizing flux via the serine–threonine pathway. Experiments further optimizing physico-chemical culture parameters, such as pH, oxygenation and carbon source feeding

could also be explored as an avenue to improve growth. This work demonstrates the capacity for complex cryptic metabolism to emerge with surprising speed. The eukaryotic serine–threonine pathway is widespread in fungi and could therefore support formatotrophy in a wide range of organisms, furthering the development of a fungal C1-trophic bioeconomy.

Methods

General strain engineering

Y. lipolytica engineering was carried out according to protocols developed in refs. 64 and 65. Overexpression constructs were constructed with Golden Gate assembly of Tef, Tef4UAS or Tef8UAS synthetic promoters, the gene of interest, a LipT1 terminator and an integration vector carrying a nourseothricin (pZNS-1.3/pJMP62-NAT) or hygromycin resistance marker (pJMP62-Hyg), and Zeta transposon-mediated

random integration sites (sequences in Supplementary Tables 2 and 3). Endogenous genes of interest were amplified by PCR from purified *Y. lipolytica* chromosomal DNA with primers bearing appropriate overhangs with Q5 polymerase and inserted by restriction ligation into expression vector pZNS1.3. Codon-optimized FDH (GenScript) from *B. stambis* was expressed under promoter pTEF with a Nourseothricin marker from the JMP series of zeta-mediated integration vectors (sequences in Supplementary Table 2)⁶⁶. All *Y. lipolytica* transformations were carried out via a standard chemical method using 2 M dithiothreitol and lithium acetate⁶⁷. Following transformation, cells were plated on either rich media (yeast extract + 2% glucose and 1% sodium formate) with 250 mg ml⁻¹ nourseothricin or defined minimal media (yeast nitrogen base (YNB) + 2% glucose, 2% citrate, 1% sodium formate and 100% synthetic complete uracil dropout supplement). Glycerol stocks were made up to 25% glycerol by mass and the appropriate volume of overnight culture made in selective media. Colony PCR was carried out using primers binding to the gene of interest and the Tef promoter of synthetic constructs (sequences in Supplementary Table 4). Laboratory *Y. lipolytica* strain Po1D with restored prototrophy was used as the base strain for high-performance liquid chromatography (HPLC) analysis and initial evolution experiments, as well as controls for strain engineering studies. The formatotrophic strain, NYF1, was isolated by streaking out the mixed culture on rich media followed by growth characterization of a single colony.

Genome-scale metabolic modelling of yield coefficients

Genome-scale metabolic modelling was carried out using the COBRA toolbox for Python and the *Y. lipolytica* genome-scale metabolic model, iYali4.2, developed in ref. 68. Theoretical yield calculations were accomplished by setting formic acid uptake rates to 1 mmol gCDW⁻¹ h⁻¹ and running Flux Balance Analysis as included with the base COBRA toolbox. New reactions corresponding to heterologous pathways were added by manually inserting reactions with appropriate reaction stoichiometry to the model. Expected yields for co-use modelling were calculated by setting carbon source uptake rates to 1 g per gram of biomass. Formic acid was either fed to the model as a supplementary carbon source using the formate metabolite provided with iYALI, or as a pseudometabolite only able to regenerate one molecule of NADH. Resultant yields were divided by the molar mass of each carbon source to derive yield coefficients. Any OD–CDW conversions required were assumed to follow the correlation reported in ref. 69 (CDW = 1.26 × OD – 0.35).

Microorganism cultivation for growth assays

Y. lipolytica strain Po1D (uracil and leucine auxotroph bearing restored prototrophies) was used as the base, ‘wild-type’ strain. Growth experiments were carried out in minimal media (YNB) with the stated carbon source concentration, with 50 mM sodium diphosphate buffer added to maintain neutral pH. C/N ratios for growth characterization experiments were maintained at 20/1 by varying ammonia concentration in the media. Formatotrophic strains were grown at ‘high’ ammonia concentrations (60 mM, C/N = 5). Strains were grown on formate as a sole carbon source at 300 mM sodium formate (Sigma). Culture conditions were varied between narrow, wide-necked and baffled flasks purchased from VWR, Corning 98-well or 24-well plates placed in an Infors HT shaking incubator set at 250 rpm. All experiments were carried out in duplicate (Figs. 5d and 6a and Supplementary Figs. 5–7, 13 and 14) or triplicate (Fig. 6d) and statistical significance was assessed using a two-tailed Student’s *t*-test. Timepoint OD₆₀₀ readings were taken with a Synergy HTX plate reader from Agilent Technologies. Experiments at high-CO₂ concentrations were carried out in 50-ml flasks shaken at 250 rpm in an Algaetron AG230 incubator with CO₂ maintained at 10,000 ppm (1%). All OD₆₀₀ readings were normalized to spectrophotometer readings by multiplying by the difference between spectrophotometer and plate reader path length (10:1). Blank measurements

were made using non-inoculated media for each growth experiment. Precultures were made 24 h before experiments in yeast extract peptone dextrose 0.5% glycerol + carbon source of interest, followed by a day culture of 6 h before inoculation of each growth experiment.

C13-labelling experiments

For C13-labelled intracellular analysis experiments, strains were cultured in 20 ml of media in 50-ml flasks. For labelling experiments, YNB + 30 mM glycerol + 150 mM formate was used. For C13 labelling of the formatotrophic strain, cultures were grown in YNB + 300 mM formate in triplicate.

For tracer experiments, metabolic reactions were quenched by adding 20 ml of precooled (–20 °C) methanol to 15 ml of culture in an ethanol-dry ice bath, followed by vortexing. Cells were washed by centrifugation at –10 °C and 4,000g and resuspension in methanol, and pellets were then dissolved in 2.5 ml of 75% ethanol at 85 °C. After incubation for 5 min, suspensions were vortexed to completely lyse cells and frozen by immersion in dry ice. Debris was pelleted by centrifugation and 1 ml of supernatant was recovered for liquid chromatography–mass spectrometry (LC–MS). For amino acid labelling in the formatotrophic strain, cell pellets left over from tracer experiments were resuspended in 4 ml of 6 M HCl and incubated at 75 °C overnight to lyse proteins. Lysates were dried by incubation at 75 °C for a further 24 h in open Falcon tubes. Amino acids were resuspended in 500 µl of H₂O, spun to remove insoluble cell debris and used as analyte. Labelled metabolites were detected using an Agilent Infinity ultra-performance liquid chromatography system and 6550 quadrupole time of flight mass spectrometers. Metabolites were separated with an Agilent Zorbax Eclipse Plus C18 column (2.1 × 50 mm, 1.8 µm) at 0.3 ml min⁻¹ flow rate. Solutions of 0.1% aqueous formic acid (inorganic) and 0.1% formic acid in acetonitrile (organic) were used as mobile phase. Mobile phase gradient was moved from 2% to 98% organic over 2.5 min, followed by 1 min of constant concentration. An The Agilent JST electrospray interface in negative ion mode was used to generate spectra.

C13 relative enrichment values were calculated by first calculating relative percentages from corrected ion percentages. These were found by dividing *m* + 1 corrected ion percentages by *m* + 0 percentages. Relative enrichment values were then calculated by subtracting average control corrected ion percentages from labelled replicates and dividing by the total relative percentage. In the case of formate-only growth, corrected ion percentages were corrected for unlabelled biomass contamination by dividing each *m* + 1, 2 (and so on) value by the OD reached by Po1D on formate over the average final OD of NYF1 grown on formate. Statistical significance was assessed using a two-tailed Student’s *t*-test.

HPLC and biomass yield analysis

HPLC experiments were carried out using a Thermo Scientific U3000 Standard HPLC machine. Mobile phase for metabolite analysis was composed of 5 mM H₂SO₄ at 0.6 ml min⁻¹ and an Aminex HPX-87H column for sugar detection set at 30 °C was used for chromatography of hydrophilic metabolites. Flow rate was set at 0.6 ml min⁻¹, with ultraviolet detection at 210 nm. Standards of each metabolite of interest were prepared by making up volumetric flasks at 20 g l⁻¹ and creating dilution series from 0.1 g l⁻¹ to 20 g l⁻¹. Calibration curves resulting from this were then used to detect metabolite concentrations using Chromeleon v.6.8 software. Samples were prepared by taking 250 µl of culture, centrifuging at 14,000g for 5 min and transferring supernatant to HPLC vials. Volumes were then made up to 500 µl with 250 µl of Millipore-purified water. For initial consumption experiments, 250-ml flasks were used. Evaporation from these flasks was accounted for by assuming 25% volume evaporation per day. In initial yield experiments, glycerol and glucose concentrations were normalized to 2% glucose by dividing enthalpies of formation: formate was supplemented at 20 g l⁻¹.

Dry biomass readings were made as follows: 2 ml of washed culture was transferred to weighed and predried 2-ml microcentrifuge tubes, and then dried for 72 h. Masses were then recorded using an Ohaus Pioneer PX142M analytical balance. Final biomass was calculated by subtracting the residual carbon in the media from the raw mass data. All experiments were carried out in triplicate.

ROS-response characterization experiments

Strains were cultivated as stated previously for experiments with formate. Where stated, 0.01 mM menadione was added to precultures for 6 h before growth experiments started. Alternatively, 0.01 g l⁻¹ BHT was added to growth media by making a 10 g l⁻¹ BHT stock solution in 100% acetone and adding as needed. BHT flasks were then vortexed to prevent precipitation of the hydrophobic antioxidant. ROS quantification was carried out using a DCFDA assay. Briefly, cells were grown at 30 °C on glycerol or glycerol + formate, then collected by centrifugation and washing in PBS. Cells were stained with 100 µl of 20 µM DCFDA solution, incubated for 30 min at 30 °C, then measured for fluorescence at excitation and emission of 485 nm and 535 nm, respectively, in an Agilent Synergy H1 plate reader. OD₆₀₀ was also assessed to normalize data. Measurements were carried out three times, in biological triplicates. Statistical significance was assessed using a two-tailed Student's *t*-test.

Carotenoid bioproduction

Carotenoid producing strains were constructed using a hygromycin-resistant 3 gene overexpression cassette based on JMP vectors with GGS1, CarB and CarRP under strong synthetic Tef8UAS promoters⁷⁰. Carotenoid producing strains were streaked to single colonies, with the strongest orange colony picked to ensure strong carotenoid production. Before culture in formate, carotenoid producing strains were cultivated in yeast extract peptone dextrose glycerol + 300 mM formate for 96 h to allow carotenoid accumulation, followed by the standard 6-h day culture before transfer to growth experiment media. Carotenoid expressing strains were cultured in 40 ml in 250-ml flasks of YNB formate 300 mM. Carotenoids were extracted for HPLC by spinning 40 ml of culture volume down, resuspending in 1 ml and then bead-beating in 100% acetone at 6,037g for 60 s. Tubes were then centrifuged at 4,193g for 3 min and supernatant removed for analysis. Carotenoids were detected by HPLC using a YMC Carotenoid C30 column at 455 nm. All growth and production experiments were carried out in triplicate. Micrograph images of carotenoid producing cells were taken in fluorescence and brightfield mode by placing 10 µl of culture media directly on slides. Culture media was either concentrated 100× by centrifuging (formate-only media) or taken directly from flasks. Statistical significance was assessed using a two-tailed Student's *t*-test.

mRNA and DNA sequencing

Samples for mRNA sequencing were prepared by growing cells in flasks with 20 ml of YNB + 30 mM glycerol (Po1D) or glycerol + 150 mM formate (Po1D, NYF1). Experiments were carried out in triplicate. Cells were collected at the end of exponential phase, pelleted by centrifugation at 4,193g at -10 °C and immediately frozen on dry ice. Whole cell pellets were shipped for commercial mRNA sequencing by Azenta Life Sciences Ltd. Briefly, mRNA fragmentation and random priming were used to append poly A tails to RNA, followed by complementary DNA amplification. dA tailing and 5' phosphorylation was then used to prepare fragments for adaptor ligation, after which PCR enrichment was used to amplify strands for sequencing. Reads trimmed using Trimmomatic v.0.36 and mapped to *Y. lipolytica* genome assembly ASM3722351v1 using a splice alignment algorithm (STAR aligner v.2.5.2b). Unique mRNA reads were calculated with Subread v.1.5.2, with reads outside intron regions discarded. DESeq2 was then used to generate *P* values and log₂ fold values using the Wald test for pairwise

differences. Adjusted *P* values >0.05 and log₂ fold change >1 were called as differentially expressed genes for analysis and repeated reads were removed. Upregulated genes were functionally annotated by homology to known sequences in the EggNOG database using eMapper with gene names reported as the 'preferred names' output for ease of interpretation. Functional clustering of genes for both strain Po1D and NYF1 was carried out using PantherDB to assess clustering in gene ontology terms for molecular functions, biological processes or cellular components. Clustering was carried out with Fisher's exact test with Bonferroni correction for multiple testing, with a *P* < 0.05 cut off. All mRNA sequencing experiments were carried out in triplicate.

For assessment of metabolic genes, entries were sorted by annotation (above) into TCA cycle, glycolysis, NADH, NADPH, cofactor, C1, lipid and membrane, electron transport chain or secondary metabolism-related genes on the basis of their annotated reactions. Here C1 metabolic genes were genes involved in C1-related reactions or assimilation pathways; NADH/NADPH were genes involved in NAD(P)H dependent redox reactions; lipid and membrane biosynthesis genes were biosynthetic genes for lipids or cell membrane components; electron transport chains were specific electron transport chain components; cofactor metabolic genes as genes involved in cofactor biosynthesis; TCA or glycolysis genes as nodes in the TCA cycle or glycolysis or gluconeogenesis (including serine biosynthesis via 3PG) and secondary metabolism genes taken to be all other genes involved in metabolite biosynthesis (amino acids, sugars and so on).

NGS DNA sequencing of strain Po1D and NYF1 was carried out by Biomarker Technologies GmbH. Whole biomass samples of each were prepared as above with strains grown in rich media for 48 h. Reference genome GCA_000002525 (strain CLIB 122) was used, with 90% of NYF1 genes and 88% of Po1D genes correctly mapped. S1000 Analysis Pipeline was used to assemble genomes using in-house methods. Single-nucleotide polymorphisms and small insertion-deletion (InDel) mutations were detected using the HaplotypeCaller algorithm from GATK software to detect variations between clean reads and the reference genome. Twenty-four single-nucleotide polymorphisms and three small InDels not present in Po1D were found in NYF1. Of these, three single-nucleotide polymorphisms and two InDels appeared in coding regions of genes. DIAMOND was used to functionally annotate genes of interest using NCBI's Non Redundant database (NR), SwissProt, Gene Ontology terms, Clusters of Orthologous Gene (COG) terms and Kyoto Encyclopedia of Genes and Genomes terms. Assessment of distribution of serine-threonine cycle or RGP genes was carried out using FUNGIpath's pathway explorer tool⁷¹. A custom BioPax2 script was used to search for EC numbers belonging to genes from each pathway across 166 fungal genomes (Supplementary Dataset 1). Genomes lacking one or more genes were assumed to be incapable of supporting the given pathway.

Mitochondrial mass quantification and mitochondrial imaging

Comparative mitochondrial mass was measured using Mitotracker Red FM. Here 1 mM stock solutions were made up by dissolving 50 µg in 92 µl of dimethyl sulfoxide, followed by dilution to 200 nM by adding 0.4 µl to 2 ml of PBS. Strains Po1D and NYF1 were grown in 2% glycerol or 2% glycerol + 300 mM formate in duplicate, then added at an OD₆₀₀ of 1.0 into 100 µl of 200 nM Mitotracker Red FM. Cells were stained by incubation at 30 °C for 30 min, then spun at 10,733g for 10 min to pellet cells. Cell pellets were resuspended in fresh PBS and measured for fluorescence and OD₆₀₀ in a Synergy H1 plate reader in triplicate (excitation 581, emission 644 nm) 10 times. Mitochondria were imaged using fluorescence (excitation 581, emission 644 nm) and brightfield microscopy after 30 min of staining with 200 nM Mitotracker Red FM in PBS. Statistical significance was assessed using a two-tailed Student's *t*-test.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data for sequencing, genome-scale metabolic models and analysis of pathway distribution is included in Supplementary Datasets 1–3. Sequencing data can be found at the NCBI BioSample database under accession numbers [SAMN50287922](#), [SAMN50287923](#), [SAMN50287924](#), [SAMN50287925](#), [SAMN50287926](#), [SAMN50287927](#), [SAMN50287928](#), [SAMN50287929](#), [SAMN50686001](#) and [SAMN50287930](#). Data supporting the findings of this study are available via Figshare at <https://doi.org/10.6084/m9.figshare.31249870> (ref. 72).

Code availability

Code for all modelling is available via GitHub at <https://github.com/WNewell21/LatentCIYLip25>.

References

- Scown, C. D. Prospects for carbon-negative biomanufacturing. *Spec. Issue Biomass Path Sustain.* **40**, 1415–1424 (2022).
- Cotton, C. A., Claassens, N. J., Benito-Vaquerizo, S. & Bar-Even, A. Renewable methanol and formate as microbial feedstocks. *Curr. Opin. Biotechnol.* **62**, 168–180 (2020).
- Yishai, O., Lindner, S. N., Gonzalez de la Cruz, J., Tenenboim, H. & Bar-Even, A. The formate bio-economy. *Curr. Opin. Chem. Biol.* **35**, 1–9 (2016).
- Hu, B. & Lidstrom, M. E. Metabolic engineering of *Methylobacterium extorquens* AM1 for 1-butanol production. *Biotechnol. Biofuels* **7**, 156 (2014).
- Yang, Y.-M. et al. Production of 3-hydroxypropionic acid in engineered *Methylobacterium extorquens* AM1 and its reassimilation through a reductive route. *Microb. Cell Factories* **16**, 179 (2017).
- Lim, C. K. et al. Designing and engineering *Methylobacterium extorquens* AM1 for itaconic acid production. *Front. Microbiol.* **10**, 1027 (2019).
- Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood–Ljungdahl pathway of CO₂ fixation. *Biochim. Biophys. Acta* **1784**, 1873–1898 (2008).
- Collas, F. et al. Engineering the biological conversion of formate into crotonate in *Cupriavidus necator*. *Metab. Eng.* **79**, 49–65 (2023).
- Gassler, T. et al. The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO₂. *Nat. Biotechnol.* **38**, 210–216 (2020).
- Panich, J., Fong, B. & Singer, S. W. Metabolic engineering of *Cupriavidus necator* H16 for sustainable biofuels from CO₂. *Trends Biotechnol.* **39**, 412–424 (2021).
- Bang, J., Hwang, C., Ahn, J. H., Lee, J. A. & Lee, S. Y. *Escherichia coli* is engineered to grow on CO₂ and formic acid. *Nat. Microbiol.* **5**, 1459–1463 (2020).
- Claassens, N. J. et al. Replacing the Calvin cycle with the reductive glycine pathway in *Cupriavidus necator*. *Metab. Eng.* **62**, 30–41 (2020).
- Bruinsma, L., Wenk, S., Claassens, N. J. & Martins dos Santos, V. A. Paving the way for synthetic C1 - Metabolism in *Pseudomonas putida* through the reductive glycine pathway. *Metab. Eng.* **76**, 215–224 (2023).
- Wenk, S. et al. Evolution-assisted engineering of *E. coli* enables growth on formic acid at ambient CO₂ via the Serine Threonine Cycle. *Metab. Eng.* **88**, 14–24 (2025).
- Kim, S. et al. Optimizing *E. coli* as a formatotrophic platform for bioproduction via the reductive glycine pathway. *Front. Bioeng. Biotechnol.* **11**, 1091899 (2023).
- Blombach, B., Grünberger, A., Centler, F., Wierckx, N. & Schmid, J. Exploiting unconventional prokaryotic hosts for industrial biotechnology. *Trends Biotechnol.* **40**, 385–397 (2022).
- Park, Y.-K. & Ledesma-Amaro, R. What makes *Yarrowia lipolytica* well suited for industry? *Trends Biotechnol.* **41**, 242–254 (2022).
- Wang, K. et al. A one-carbon chemicals conversion strategy to produce precursor of biofuels with *Saccharomyces cerevisiae*. *Renew. Energy* **208**, 331–340 (2023).
- Mitic, B. M. et al. The oxygen-tolerant reductive glycine pathway assimilates methanol, formate and CO₂ in the yeast *Komagataella phaffii*. *Nat. Commun.* **14**, 7754 (2023).
- Guo, F. et al. Evolutionary engineering of *Saccharomyces cerevisiae*: crafting a synthetic methylotroph via self-reprogramming. *Sci. Adv.* **10**, eadq3484 (2024).
- Xu, P., Qiao, K., Ahn, W. S. & Stephanopoulos, G. Engineering *Yarrowia lipolytica* as a platform for synthesis of drop-in transportation fuels and oleochemicals. *Proc. Natl Acad. Sci. USA* **113**, 10848–10853 (2016).
- Vandermies, M. & Fickers, P. Bioreactor-scale strategies for the production of recombinant protein in the yeast *Yarrowia lipolytica*. *Microorganisms* **7**, 40 (2019).
- Konzock, O., Zaghen, S. & Norbeck, J. Tolerance of *Yarrowia lipolytica* to inhibitors commonly found in lignocellulosic hydrolysates. *BMC Microbiol.* **21**, 77 (2021).
- Robles-Rodríguez, C. E. et al. Modeling and optimization of lipid accumulation by *Yarrowia lipolytica* from glucose under nitrogen depletion conditions. *Biotechnol. Bioeng.* **115**, 1137–1151 (2018).
- van Winden, W. A. et al. Towards closed carbon loop fermentations: cofeeding of *Yarrowia lipolytica* with glucose and formic acid. *Biotechnol. Bioeng.* **119**, 2142–2151 (2022).
- Chen, Q. et al. Unlocking the formate utilization of wild-type *Yarrowia lipolytica* through adaptive laboratory evolution. *Biotechnol. J.* **19**, 2400290 (2024).
- Gonzalez de la Cruz, J., Machens, F., Messerschmidt, K. & Bar-Even, A. Core catalysis of the reductive glycine pathway demonstrated in yeast. *ACS Synth. Biol.* **8**, 911–917 (2019).
- Du, C., Li, Y., Xiang, R. & Yuan, W. Formate dehydrogenase improves the resistance to formic acid and acetic acid simultaneously in *Saccharomyces cerevisiae*. *Int. J. Mol. Sci.* **23**, 3406 (2022).
- Yamashita, T., Nakamaru-Ogiso, E., Miyoshi, H., Matsuno-Yagi, A. & Yagi, T. Roles of bound quinone in the single subunit NADH-quinone oxidoreductase (Ndi1) from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **282**, 6012–6020 (2007).
- Worland, A. M. et al. Elucidation of triacylglycerol catabolism in *Yarrowia lipolytica*: how cells balance acetyl-CoA and excess reducing equivalents. *Metab. Eng.* **85**, 1–13 (2024).
- Kerscher, S. J., Okun, J. G. & Brandt, U. A single external enzyme confers alternative NADH:ubiquinone oxidoreductase activity in *Yarrowia lipolytica*. *J. Cell Sci.* **112**, 2347–2354 (1999).
- Du, L. et al. Formic acid induces Yca1p-independent apoptosis-like cell death in the yeast *Saccharomyces cerevisiae*: formic acid induces yeast apoptosis-like cell death. *FEMS Yeast Res.* **8**, 531–539 (2008).
- Young, A., Gardiner, D., Brosnan, M. E., Brosnan, J. T. & Mailloux, R. J. Physiological levels of formate activate mitochondrial superoxide/hydrogen peroxide release from mouse liver mitochondria. *FEBS Lett.* **591**, 2426–2438 (2017).
- Zeng, L. et al. Metabolome analysis of the response and tolerance mechanisms of *Saccharomyces cerevisiae* to formic acid stress. *Int. J. Biochem. Cell Biol.* **148**, 106236 (2022).
- Odat, O. et al. Old yellow enzymes, highly homologous FMN oxidoreductases with modulating roles in oxidative stress and programmed cell death in yeast. *J. Biol. Chem.* **282**, 36010–36023 (2007).

36. Zvyagil'skaya, R. A. et al. Characterization of the Pho89 phosphate transporter by functional hyperexpression in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* **8**, 685–696 (2008).
37. Chen, M. et al. Yeast increases glycolytic flux to support higher growth rates accompanied by decreased metabolite regulation and lower protein phosphorylation. *Proc. Natl Acad. Sci. USA* **120**, e2302779120 (2023).
38. Zeng, L. et al. Transcriptomic analysis of formic acid stress response in *Saccharomyces cerevisiae*. *World J. Microbiol. Biotechnol.* **38**, 34 (2022).
39. Knoblach, B. & Rachubinski, R. A. Reconstitution of human peroxisomal β -oxidation in yeast. *FEMS Yeast Res.* **18**, foy092 (2018).
40. Jin, X., Chen, M., Coldea, T. E., Yang, H. & Zhao, H. Protective effects of peptides on the cell wall structure of yeast under osmotic stress. *Appl. Microbiol. Biotechnol.* **106**, 7051–7061 (2022).
41. Dourou, M., Mizerakis, P., Papanikolaou, S. & Aggelis, G. Storage lipid and polysaccharide metabolism in *Yarrowia lipolytica* and *Umbelopsis isabellina*. *Appl. Microbiol. Biotechnol.* **101**, 7213–7226 (2017).
42. Wenk, S. et al. An 'energy-auxotroph' *Escherichia coli* provides an in vivo platform for assessing NADH regeneration systems. *Biotechnol. Bioeng.* **117**, 3422–3434 (2020).
43. Cavero, S. et al. Identification and metabolic role of the mitochondrial aspartate-glutamate transporter in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **50**, 1257–1269 (2003).
44. Martínez-Moral, M. Allantoin as a marker of oxidative stress: inter- and intra-individual variability in urinary concentrations in healthy individuals. *Environ. Sci. Technol. Lett.* **6**, 283–288 (2019).
45. Bortolotti, M., Polito, L., Battelli, M. G. & Bolognesi, A. Xanthine oxidoreductase: one enzyme for multiple physiological tasks. *Redox Biol.* **41**, 101882 (2021).
46. Linder, T. A genomic survey of nitrogen assimilation pathways in budding yeasts (sub-phylum Saccharomycotina). *Yeast* **36**, 259–273 (2019).
47. Xu, P., Qiao, K. & Stephanopoulos, G. Engineering oxidative stress defense pathways to build a robust lipid production platform in *Yarrowia lipolytica*. *Biotechnol. Bioeng.* **114**, 1521–1530 (2017).
48. Li, Y., Zhu, H. & Trush, M. A. Detection of mitochondria-derived reactive oxygen species production by the chemiluminescent probes lucigenin and luminol. *Biochim. Biophys. Acta* **1428**, 1–12 (1999).
49. Tajeddine, N. How do reactive oxygen species and calcium trigger mitochondrial membrane permeabilisation?. *Biochim. Biophys. Acta* **1860**, 1079–1088 (2016).
50. Neikirk, K. et al. MitoTracker: a useful tool in need of better alternatives. *Eur. J. Cell Biol.* **102**, 151371 (2023).
51. Cléménçon, B., Rey, M., Trézéguet, V., Forest, E. & Pelosi, L. Yeast ADP/ATP carrier isoform 2: conformational dynamics and role of the RRRMMM signature sequence methionines. *J. Biol. Chem.* **286**, 36119–36131 (2011).
52. Minard, K. I. & McAlister-Henn, L. Sources of NADPH in yeast vary with carbon source. *J. Biol. Chem.* **280**, 39890–39896 (2005).
53. Hatrongjit, R. & Packdibamrung, K. A novel NADP⁺-dependent formate dehydrogenase from *Burkholderia stabilis* 15516: screening, purification and characterization. *Enzyme Microb. Technol.* **46**, 557–561 (2010).
54. Xiao, W. & Loscalzo, J. Metabolic responses to reductive stress. *Antioxid. Redox Signal.* **32**, 1330 (2020).
55. Festjens, N. et al. Butylated hydroxyanisole is more than a reactive oxygen species scavenger. *Cell Death Differ.* **13**, 166–169 (2006).
56. Park, H. et al. Efficient production of retinol in *Yarrowia lipolytica* by increasing stability using antioxidant and detergent extraction. *Metab. Eng.* **73**, 26–37 (2022).
57. Biriukova, E. N., Medentsev, A. G., Arinbasarova, A. I. & Akimenko, V. K. Respiratory activity of yeast *Yarrowia lipolytica* under oxidative stress and heat shock. *Mikrobiologiya* **77**, 448–452 (2008).
58. Jobbagy, S. et al. Electrophiles modulate glutathione reductase activity via alkylation and upregulation of glutathione biosynthesis. *Redox Biol.* **21**, 101050 (2018).
59. Larroude, M. et al. A synthetic biology approach to transform *Yarrowia lipolytica* into a competitive biotechnological producer of β -carotene. *Biotechnol. Bioeng.* **115**, 464–472 (2018).
60. Gao, S. et al. Iterative integration of multiple-copy pathway genes in *Yarrowia lipolytica* for heterologous β -carotene production. *Metab. Eng.* **41**, 192–201 (2017).
61. Kawata, A., Murakami, Y., Suzuki, S. & Fujisawa, S. Anti-inflammatory activity of β -carotene, lycopene and tri-*n*-butylborane, a scavenger of reactive oxygen species. *In Vivo* **32**, 255–264 (2018).
62. Ma, Y. et al. Removal of lycopene substrate inhibition enables high carotenoid productivity in *Yarrowia lipolytica*. *Nat. Commun.* **13**, 572 (2022).
63. Adeleke, R., Nwangburuka, C. & Oboirien, B. Origins, roles and fate of organic acids in soils: a review. *South Afr. J. Bot.* **108**, 393–406 (2017).
64. Yuzbashev, T. V. et al. A DNA assembly toolkit to unlock the CRISPR/Cas9 potential for metabolic engineering. *Commun. Biol.* **6**, 858 (2023).
65. Larroude, M. et al. A modular Golden Gate toolkit for *Yarrowia lipolytica* synthetic biology. *Microb. Biotechnol.* **12**, 1249–1259 (2019).
66. Trassaert, M. et al. New inducible promoter for gene expression and synthetic biology in *Yarrowia lipolytica*. *Microb. Cell Factories* **16**, 141 (2017).
67. Tsakraklides, V., Brevnova, E., Stephanopoulos, G. & Shaw, A. J. Improved gene targeting through cell cycle synchronization. *PLoS ONE* **10**, e0133434 (2015).
68. Kerkhoven, E. J., Pomraning, K. R., Baker, S. E. & Nielsen, J. Regulation of amino-acid metabolism controls flux to lipid accumulation in *Yarrowia lipolytica*. *Npj Syst. Biol. Appl.* **2**, 16005 (2016).
69. Worland, A. M. et al. Analysis of *Yarrowia lipolytica* growth, catabolism, and terpenoid biosynthesis during utilization of lipid-derived feedstock. *Metab. Eng. Commun.* **11**, e00130 (2020).
70. Pesantes-Munoz, M. & Ledesma-Amaro, R. in *Yarrowia lipolytica: Methods and Protocols* (eds Wheeldon, I. & Blenner, M.) 191–204 (Springer, 2021).
71. Grossetête, S., Labedan, B. & Lespinet, O. FUNGIpath: a tool to assess fungal metabolic pathways predicted by orthology. *BMC Genomics* **11**, 81 (2010).
72. Newell, W., Ledesma-Amaro, R. & Hapeta, P. Dataset—'Engineering de novo formatotrophy in the non-model yeast *Y. lipolytica*'. *figshare* <https://doi.org/10.6084/m9.figshare.31249870> (2026).

Acknowledgements

This work was funded by an Engineering and Physical Sciences Research Council (EPSRC) Biodesign Bioengineering CDT studentship and a Carbon Recycling Network (NIBB) proof of concept grant. R.L.-A. received funding from the BBSRC (grant nos. BB/R01602X/1, BB/T013176/1, BB/T011408/1-19-ERACoBioTech-33 SyCoLim, BB/X01911X/1, BB/Y008510/1: Engineering Biology Hub for Microbial Foods), EPSRC (grant no. AI-4-EB BB/W013770/1 and EEBio Programme grant no. EP/Y014073/1), Yeast4Bio Cost Action 18229, European Research Council (ERC) (grant no. DEUSBIO-949080), the Bio-based Industries Joint (grant no. PERFECOAT-101022370) under the European

Union's Horizon 2020 research and innovation programme and the European Innovation Council under grant agreement no. 101098826 (SKINDEV). Imperial College London UKRI Impact Acceleration Account (grant nos. EPSRC-EP/X52556X/1 and BBSRC-BB/X511055/1). We thank the Bezos Earth Fund through the Bezos Centre for Sustainable Protein (grant no. BCSP/IC/001). We thank K. Roy (Biome Bioplastics, Southampton, UK), J. Fu (Bezons Centre for Sustainable Protein, Imperial College London, London, UK), J. Lopez (Bezons Centre for Sustainable Protein, Imperial College London, London, UK), T. Yuzbashev (Rothamsted Research, Harpenden, UK) and A. Graham (Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London, London, UK) for their scholarly conversations and advice concerning experimental procedures and article formatting.

Author contributions

Study conception and design were developed by W.N., P.H. and R.L.-A. Data were collected by W.N., P.H. and D.B. Analysis and interpretation of results were carried out by W.N., P.H. and R.L.-A. The draft paper was prepared by W.N. All authors reviewed the results and approved the final version of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44160-026-01076-7>.

Correspondence and requests for materials should be addressed to Rodrigo Ledesma-Amaro.

Peer review information *Nature Synthesis* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Thomas West, in collaboration with the *Nature Synthesis* team.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- | | | |
|-------------------------------------|-------------------------------------|--|
| ☐ | ☒ | The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☒ | ☐ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection All genome scale modelling was performed using CobraPy with IYali4.2. Code for all modelling is available at <https://github.com/WNewell21/LatentC1YLip25>.

Data analysis All data analysis was performed in Microsoft Excel

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Source data for sequencing, GSMMs and analysis of pathway distribution is included in supplementary datasets. Sequencing data can be found at the NCBI BioSample database under accession numbers SAMN50287922, SAMN50287923, SAMN50287924, SAMN50287925, SAMN50287926, SAMN50287927,

SAMN50287928, SAMN50287929, SAMN50686001 and SAMN50287930. Data supporting the findings of this study are available at Fishare, DOI: 10.6084/m9.figshare.31249870

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Use the terms *sex* (biological attribute) and *gender* (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.

Reporting on race, ethnicity, or other socially relevant groupings

Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status). Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.) Please provide details about how you controlled for confounding variables in your analyses.

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size decisions were based on available equipment and variation in data observed,
Data exclusions	No data were excluded
Replication	For larger (n>3) datasets, technical replicates were carried out in as large a number as feasible with available equipment. For smaller sample sizes, pilot experiments were carried out to establish trends reported in data.
Randomization	Randomisation was not applicable as control experiments showed no effects of measurement
Blinding	N/A as trends were clearly observable in all cases

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.