

Elucidation of triacylglycerol catabolism in *Yarrowia lipolytica*: How cells balance acetyl-CoA and excess reducing equivalents



Alyssa M. Worland^a, Zhenlin Han^b, Jessica Maruwan^b, Yu Wang^b, Zhi-Yan Du^b, Yinjie J. Tang^a, Wei Wen Su^{b,*}, Garrett W. Roell^{b, **}

^a Department of Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO, 63130, United States

^b Department of Molecular Biosciences and Bioengineering, University of Hawai'i at Mānoa, Honolulu, HI, 96822, United States

ARTICLE INFO

Keywords:

Yarrowia lipolytica
¹³C-metabolic flux analysis
Genome-scale model
NADPH
Lipid catabolism
Mannitol cycle
Glyoxylate shunt

ABSTRACT

Yarrowia lipolytica is an industrial yeast that can convert waste oil to value-added products. However, it is unclear how this yeast metabolizes lipid feedstocks, specifically triacylglycerol (TAG) substrates. This study used ¹³C-metabolic flux analysis (¹³C-MFA), genome-scale modeling, and transcriptomics analyses to investigate *Y. lipolytica* W29 growth with oleic acid, glycerol, and glucose. Transcriptomics data were used to guide ¹³C-MFA model construction and to validate the ¹³C-MFA results. The ¹³C-MFA data were then used to constrain a genome-scale model (GSM), which predicted *Y. lipolytica* fluxes, cofactor balance, and theoretical yields of terpene products. The three data sources provided new insights into cellular regulation during catabolism of glycerol and fatty acid components of TAG substrates, and how their consumption routes differ from glucose catabolism. We found that (1) over 80% of acetyl-CoA from oleic acid is processed through the glyoxylate shunt, a pathway that generates less CO₂ compared to the TCA cycle, (2) the carnitine shuttle is a key regulator of the cytosolic acetyl-CoA pool in oleic acid and glycerol cultures, (3) the oxidative pentose phosphate pathway and mannitol cycle are key routes for NADPH generation, (4) the mannitol cycle and alternative oxidase activity help balance excess NADH generated from β-oxidation of oleic acid, and (5) asymmetrical gene expressions and GSM simulations of enzyme usage suggest an increased metabolic burden for oleic acid catabolism.

1. Introduction

Modern sustainability initiatives emphasize resource conservation and waste reduction. While the majority of waste valorization efforts have focused on carbohydrates, waste oil feedstocks generated from food and agriculture remain underutilized, on the order of billions of pounds per year in the US alone (Awogbemi and Kallon, 2022; Nelson, 2016; Sawangkeaw and Ngamprasertsith, 2013). Fermentations with the oleaginous yeast *Yarrowia lipolytica* offer a route to convert these substrates to value-added acetyl-CoA-derived products, such as terpenes (Czajka et al., 2018; Papanikolaou and Aggelis, 2010; Worland et al., 2020a; Zhang et al., 2017). However, the mechanisms underlying *Y. lipolytica*'s ability to consume and utilize triacylglycerols (TAGs) are not adequately understood, specifically the major routes in which NADPH is synthesized and how excess NADH and acetyl-CoA generated from fatty acid β-oxidation are balanced.

When *Y. lipolytica* is grown with glucose or acetate, the oxidative pentose phosphate pathway is the primary source of NADPH generation (Liu et al., 2016; Wasylenko et al., 2015). However, the contribution of this pathway to the NADPH pool remains unclear when *Y. lipolytica* is cultured with TAG substrates. Alternative NADPH sources include NADP⁺-dependent isocitrate dehydrogenase (IDH), the mannitol cycle via mannitol-2-dehydrogenase (MDH) through the conversion of NADH and ATP, or the folate cycle via methylenetetrahydrofolate dehydrogenase (Dulermo et al., 2015; Kavšek et al., 2015; Wang et al., 2020). Alternative oxidase (AOX) offers a potential route to balance excess NADH, and may be activated to generate NAD⁺ needed to support the TCA cycle. However, energy is not conserved through this reaction and is instead released as heat. Understanding the implications of AOX activity and activation can offer insights into cellular thermogenesis, and has been previously shown to impact citrate production and lipid accumulation in this yeast (da Veiga Moreira et al., 2021; Rogov and

* Corresponding author.

** Corresponding author.

E-mail addresses: wsu@hawaii.edu (W.W. Su), groell@hawaii.edu (G.W. Roell).

<https://doi.org/10.1016/j.ymben.2024.06.010>

Received 4 March 2024; Received in revised form 6 June 2024; Accepted 24 June 2024

Available online 26 June 2024

1096-7176/© 2024 International Metabolic Engineering Society. Published by Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Zvyagilskaya, 2015).

Moreover, it is unclear which pathways most actively consume the acetyl-CoA generated through β -oxidation. In addition to contributing to biomass synthesis, cytosolic acetyl-CoA can be transported into the mitochondria via the carnitine shuttle (Messina et al., 2023), where it enters the TCA cycle via citrate synthase. Alternatively, cytosolic acetyl-CoA can be directed to the glyoxylate shunt to form malate that can then be either transported into the mitochondria (Luévanos-Martínez et al., 2010) or converted to oxaloacetate via cytosolic malate dehydrogenase. Cytosolic oxaloacetate can be transformed into phosphoenolpyruvate (PEP) via PEP carboxykinase for gluconeogenesis.

Y. lipolytica secretes extracellular lipases and surfactants to break down TAG substrates into glycerol and fatty acids (Fig. 1) (Gonçalves et al., 2014; Spagnuolo et al., 2018). Glycerol enters central metabolism via phosphorylation to glycerol-3-phosphate followed by oxidation to dihydroxyacetone phosphate (DHAP), interchangeably converted to glyceraldehyde-3-phosphate (GAP). Concurrently, the free fatty acids undergo ATP-mediated activation into fatty acyl-CoAs before entering the peroxisome for β -oxidation. The complete oxidation of oleic acid generates 9 acetyl-CoA, 8 NADH, and 7 FADH₂ (Supplementary File 2).

Glycerol is a preferred carbon substrate for *Y. lipolytica*, but glycerol makes up less than 10% of vegetable oils by weight (Erian et al., 2022; Worland et al., 2020b). An analysis of spent medium after 8 days of *Y. lipolytica* growth in 20% plant oil found no residual glycerol, despite finding that 65% of the residual oil had been hydrolyzed to free fatty acids (Li et al., 2020). While our previous report showed that glycerol and oleic acid can be consumed concurrently in the short term (Worland et al., 2020b), this observation suggests that, during longer fermentations, cells could only have access to free fatty acids after glycerol is fully depleted. Further, waste cooking oil can contain up to 40% free fatty acids (Kulkarni and Dalai, 2006). Thus, our study focuses on evaluating glycerol and long-chain fatty acid catabolism individually.

¹³C-metabolic flux analysis (¹³C-MFA) is a tool that can help delineate the reaction rates of central metabolic pathways. We conducted stationary ¹³C-MFA on *Y. lipolytica* W29 grown with glycerol and oleic

acid, separately, and as a control, ¹³C-MFA was also conducted with glucose. Previous ¹³C-MFA studies on *Y. lipolytica* have focused on glucose (Wasylewko et al., 2015; Zhang et al., 2016), glycerol (Sabra et al., 2017), or acetate (Liu et al., 2016), leaving a gap in understanding how this yeasts' metabolism functions when consuming TAG substrates. To our knowledge, ¹³C-MFA has not been conducted with *Y. lipolytica* when grown with a long-chain fatty acid. Oleic acid was chosen as the long-chain fatty acid in this study it is one of the most abundant fatty acids in many plant oils (Orsavova et al., 2015). The ¹³C-MFA models for the three carbon sources were curated using transcriptomics data to exclude 1–2 reactions with low transcript counts. Additionally, transcript levels were used to validate reaction rate differences observed between the carbon sources for the corresponding enzymes. Beyond the capabilities of ¹³C-MFA and transcriptomics, genome-scale models (GSMs) can explore genome wide reaction networks (Edwards and Palsson, 2000). Here, a *Y. lipolytica* GSM and flux variability analysis (FVA) were used to examine metabolic reaction rates and measured biomass yields (Mahadevan and Schilling, 2003), followed by using the ¹³C-MFA confidence intervals to further constrain the GSM. This iterative process narrowed the distance between minimum and maximum reaction rate values, resulting in improved genome-scale flux predictions and understanding of lipid catabolism in *Y. lipolytica*.

2. Materials and methods

2.1. Strains, media, culture conditions, and strain engineering

Y. lipolytica strain W29 (ATCC, 20460) is investigated in this work. All cultures were grown at 30 °C shaking at 250 rpm. For ¹³C-MFA experiments, two sequential seed cultures were used to reduce non-labeled carbon carryover. Specifically, the first seed culture in YPD medium (10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose) was grown to stationary phase and used to inoculate the second seed culture at 0.3% culture volume. The second seed culture contained minimal YNB medium (1.7 g/L yeast nitrogen base without amino acids or carbohydrates, with 5 g/L ammonium sulfate) and 4 g/L positionally labeled carbon substrate for the respective ¹³C-MFA experiment. 1,2-¹³C glucose, 1,3-¹³C glycerol (Cambridge Isotopes, Tewksbury, MA, USA), and 1,2,3,7,8-¹³C oleic acid (C18:1) (MilliporeSigma, Burlington, MA, USA) were the labeled carbon tracers. The oleic acid tracer has a ¹²C/¹³C isotope ratio of 0.28, suitable for achieving sufficient labeling in proteinogenic amino acids. ¹³C-MFA cultures (minimal YNB medium and 2 g/L of the respective carbon tracer) were inoculated at 2% volume ratio from the second seed and were harvested at mid-exponential phase for proteinogenic amino acid analysis. Cultures for growth and RNA-sequencing analysis followed this procedure with unlabeled carbon substrates. A mannitol-2-dehydrogenase knockout strain was developed to evaluate the role of the mannitol cycle in this yeast. For details on this strain engineering, see Supplementary Materials and Methods.

2.2. Calculations of growth parameters

Biomass concentration was measured using a spectrophotometer and the conversion factor 1 OD₆₀₀ = 0.2955 gDCW/L. This conversion factor was derived from a standard curve of OD₆₀₀ vs. non-ash-free biomass dry cell weight measurements ($R^2 = 0.99$, data not shown). The growth rate (h^{-1}) for the three substrates was determined as the slope of a linear regression fit to the natural logarithm of biomass concentration plotted against time (Supplementary Fig. 2). Substrate concentration curves for glucose and glycerol were determined using enzyme kits purchased from R-Biopharm (Darmstadt, Hessen, Germany). The yield coefficients (gDCW mmol⁻¹) for glucose and glycerol were calculated as the change in biomass concentration (ΔX) divided by the change in substrate concentration (ΔS) over the duration of the exponential phase. The yield coefficient for oleic acid was determined by log-normal fitting of colony forming units (CFU) vs. time and converting CFU to cell dry weight with

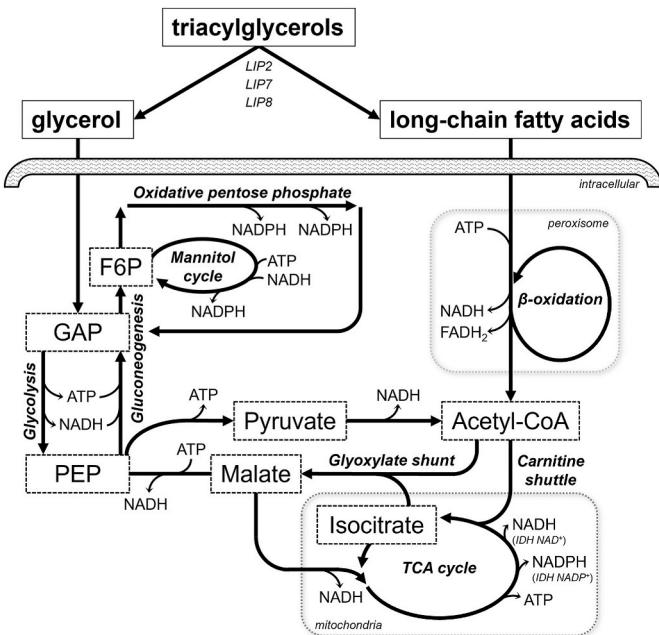


Fig. 1. Metabolic pathways for the utilization of triacylglycerols (TAGs). The diagram illustrates the metabolic precursors and intermediates involved in the consumption of TAGs, highlighting the separate pathways for glycerol and long-chain fatty acid integration into key metabolites such as GAP (glyceraldehyde-3-phosphate), PEP (phosphoenolpyruvate), F6P (fructose-6-phosphate), pyruvate, malate, isocitrate, and acetyl-CoA.

a conversion factor of 1.96×10^{-8} gDCW per CFU (Myers et al., 2013). Oleic acid consumption was determined via GC-FID (flame ionization detection) following a hexane extraction of the growth medium and FAME (fatty acid methyl ester) analysis. The specific substrate uptake rate (mmol gDCW $^{-1}$ h $^{-1}$) was determined by dividing the growth rate by the yield coefficient (Supplementary Notebook A).

2.3. RNA sequencing analysis

Y. lipolytica cultures grown with the three substrates were centrifugally harvested and washed with nanopure water. Total RNA was extracted using the Quick-RNATM Fungal/Bacterial MiniPrep kit from Zymo Research Corp (Irvine, California, USA). The high-quality RNA, as verified by an Agilent 2100 bioanalyzer at the University of Hawai'i Cancer Center, was sequenced at the DOE Joint Genome Institute (Berkeley, CA, USA). Gene expression levels were quantified using fragments per kilobase of transcript per million mapped reads (FPKM) counts, which normalizes the number of fragments to the length of the gene and the total number of mapped reads, providing a standardized measure of expression (Conesa et al., 2016). Additionally, the DESeq2 package was used to identify differentially expressed genes (Love et al., 2014).

2.4. ^{13}C -MFA

^{13}C -cultures were harvested in mid-log phase and hydrolyzed for proteinogenic amino acid analysis. A GC-MS based TBDMS (N-(tert-butyldimethylsilyl)-N-methyltrifluoroacetamide) method was used to analyze the resulting amino acid labeling patterns (Worland et al., 2020b; You et al., 2012). Each experiment was performed with 4 biological replicates ($n = 4$) collected at two time points in the exponential phase. Glucose and glycerol cultures were harvested at OD $_{600}$ 2.0 and 3.5, and oleic acid cultures were harvested at OD $_{600}$ 1.7 and 2.7. Mass isotopomer distribution (MID) variations were <1% between time-course samples, confirming a steady-state metabolism suitable for MFA calculations (Supplementary File 2). A compartmentalized metabolic network was constructed from a previously published model (Liu et al., 2016) (Supplementary File 2). Due to low carbon substrate concentrations (2 g/L), extracellular flux of overflow organic acids (citrate, acetate, isocitrate, and malate) was not detected. Therefore, the MFA model inputs were amino acid MIDs and biomass yield for each substrate. Custom biomass equations were made for these substrates. The coefficients for amino acids were determined based on measurements from our previous work (Worland et al., 2020b). The coefficients for acetyl-CoA and DHAP were determined by reported lipid levels, G6P's coefficient by carbohydrate levels, and coefficients for nucleic acid precursors by a similar method (Luttermann et al., 2021; Niehus et al., 2018). The growth-associated ATP maintenance value was based on a previous report (Liu et al., 2016). For simplicity, our models assume FADH $_2$ is the equivalent of 1.5 ATP. The stationary flux calculations were done using the software INCA (Isotopomer Network Compartmental Analysis) (Young, 2014). Fluxes reported are based on the specific substrate uptake rate for each carbon source. 95% and 68% confidence intervals (CIs) were determined using the parameter continuation function in INCA (Supplementary File 2) (Antoniewicz et al., 2006).

2.5. Integrative GSM

The iYLI647 GSM (Mishra et al., 2018) was modified to improve the accuracy of its predictions. First, the custom biomass equations previously described were added to the GSM for compatibility with our ^{13}C -MFA models. The coefficients of the biomass reactions in the GSM were scaled to ensure that the biomass reactions had the same carbon requirements in both the MFA and GSM models. (Supplementary Notebook B). Additionally, a reaction for cytosolic malate synthase, the

reversibility for two reactions in the carnitine cycle, and a reaction producing β -carotene from farnesyl pyrophosphate (FPP) were added to the GSM. Finally, several corrections were made to the gene annotations for reactions in iYLI647 based on the RNA-seq annotations. For diphosphoglyceromutase, a duplicate gene was removed from the model. For the cytochrome c oxidase reactions, gene names that were not present in the RNA-seq dataset were removed. For hexadecanoate transport, decanoate transport, and acetyl-CoA:acetoacetyl-CoA transferase, typographical gene annotation errors were fixed. The corrected model containing these modifications is denoted as iYLI647_corr_3 and is available online (see Code Availability, Supplementary File 1). To integrate ^{13}C -MFA with the GSM, a total of 46 reactions were mapped between the ^{13}C -MFA metabolic network and reactions in the GSM. Some ^{13}C -MFA reactions were linked to multiple GSM reactions. For example, the ^{13}C -MFA reaction F6P + ATP \leftrightarrow FBP was mapped to the GSM reactions phosphofructokinase and fructose bisphosphatase. The StrainDesign Toolbox (Schneider et al., 2022) was used to determine feasible ^{13}C -MFA bounds given the measured yield coefficient for each carbon source (Supplementary Notebook B). In brief, this was done using the 'fba' function of the toolbox with the condition that the biomass flux is greater than or equal to the biomass cutoff value, and the objective function was set to either maximize or minimize the rate of each reaction. In the case of F6P + ATP \leftrightarrow FBP, the objective function was to maximize or minimize the difference between phosphofructokinase and fructose bisphosphatase, since the enzymes catalyze the same reaction in opposite directions. The measured yield coefficients and their standard deviations for each carbon source were used to calculate a biomass cutoff flux set at 2 standard deviations below the measured yield coefficient. This was done to roughly estimate the 95% confidence interval lower bound of the yield coefficient. The resulting values from these optimizations were used as the upper and lower bounds for each reaction in the ^{13}C -MFA. Then, the resulting ^{13}C -MFA flux confidence intervals were used to constrain the GSM. Importantly, no feasible solutions were possible in the GSM if all the ^{13}C -MFA reaction bounds were used as constraints. Therefore, for each carbon source, a subset of at least ten ^{13}C -MFA reactions were used to constrain the GSM (Supplementary Notebook C). Each subset was chosen to include key reactions from glycolysis, the pentose phosphate pathway, and the TCA cycle. The selection process involved iterative refinement where subsets of reactions were tested for their ability to allow the GSM to yield feasible solutions. This was done to ensure the inclusion of reaction bounds from many metabolic pathways while maintaining the computational tractability of the model.

3. Results

3.1. Fermentation profiles and growth parameter calculations

The metabolism of *Y. lipolytica* was investigated through growth profile analysis when cultivated with glucose, glycerol, and oleic acid (Supplementary Table 1). Glycerol-fed cells exhibited the fastest specific growth rate of 0.36 h $^{-1}$, whereas cultures fed with either glucose or oleic acid had similar growth rates of around 0.26 h $^{-1}$. These data are in line with previous research (Papanikolaou et al., 2001; Worland et al., 2020b). The growth rate differences between glycerol and glucose and between glycerol and oleic acid were statistically significant ($P < 0.001$, two-tailed Student's t-test). When normalized to carbon per mol of the substrate, the yield coefficients of the three carbon sources were similar, indicating that the accelerated growth rate of glycerol-fed cells was achieved through a faster substrate uptake rate rather than increased yield. When the specific substrate uptake rates were normalized to carbon per mol of each carbon source, we found that *Y. lipolytica* was able to uptake 18.8 mmol C gDCW $^{-1}$ h $^{-1}$ when consuming glycerol, 14.5 mmol C gDCW $^{-1}$ h $^{-1}$ when consuming glucose, and 13.7 mmol C gDCW $^{-1}$ h $^{-1}$ when consuming oleic acid. This finding indicates that *Y. lipolytica* can catabolize oleic acid efficiently, similar to glucose. Plots

comparing measured and fitted growth and consumption curves can be found in [Supplementary Fig. 2](#).

In this work, we conducted stationary ^{13}C -MFA which relies on the assumption that the isotopic labeling patterns of the proteinogenic amino acids and the central metabolic pathway reaction rates are at steady state when the cultures are harvested. It's critical to determine when these assumptions are satisfied during cultivation with each carbon source. Thus, growth profiles for glucose and glycerol cultures showed the mid-exponential phase to be between OD₆₀₀ 1.5 and 3.5, and in oleic acid cultures this phase was determined between OD₆₀₀ 1.0 and 2.8 ([Supplementary Fig. 1A](#)).

3.2. RNA-sequencing analysis

RNA-sequencing analysis examined *Y. lipolytica* transcriptional regulation in each substrate. Of the three carbon sources, differential gene expression (DGE) analysis showed the largest difference in gene regulation was between cells grown on glucose and oleic acid ([Fig. 2A](#)). There were 182 unique genes that were differentially expressed with an absolute \log_2 fold change greater than 2 and $p_{\text{adj}} < 0.05$. These results are expected since glucose and oleic acid represent two fundamentally different types of carbon sources and their utilization involves distinct entry points into the cell's metabolic network. Glucose enters through glycolysis, while oleic acid is funneled into the TCA cycle or the glyoxylate shunt following β -oxidation ([Fig. 1](#)). Glycerol enters midway through glycolysis, and we found that gene expression was more similar to glucose than to oleic acid was it consumed.

Interestingly, many of the differentially expressed genes in glucose and glycerol cultures were from central pathways ([Fig. 2B](#)). For

example, pyruvate carboxylase and formate dehydrogenase were significantly upregulated when *Y. lipolytica* was grown in glycerol than when it was grown with glucose. Conversely, several TCA cycle genes showed higher expression in glucose cultures than in glycerol. A closer analysis of the differentially expressed genes between glucose and oleic acid cultures shows an asymmetrical distribution in *Y. lipolytica* gene expression when consuming the two carbon sources ([Fig. 2C](#)). We found that, of the differentially expressed genes with an absolute \log_2 fold change greater than 1, 80.5% showed higher expression in oleic acid compared to glucose. Compared with glycerol, 73.8% of differentially expressed genes are upregulated in oleic acid. Between glucose and glycerol, DGE showed 68.8% of genes upregulated in glycerol cultures. The increased gene expression in oleic acid cultures may indicate that a higher enzyme abundance is required to direct flux and sustain the metabolic network during oleic acid catabolism. Higher enzyme expression can increase metabolic burden, leading to reduced cell growth. Additionally, during glucose catabolism, cells may choose to suppress genes needed for oleic acid consumption, while still expressing glucose-specific enzymes for gluconeogenesis during oleic acid catabolism. This regulation would allow for a rapid response in case glucose becomes available. A similar trend was found in the comparison of glycerol and oleic acid cultures.

3.3. ^{13}C -MFA model refinement guided by RNA-sequencing

The ^{13}C -MFA reaction network contains compartments with separate metabolite pools and reactions occurring in both the cytosol and the mitochondria, and reactions transporting metabolites across the mitochondrial membrane. Here, we used transcriptomics data to guide the

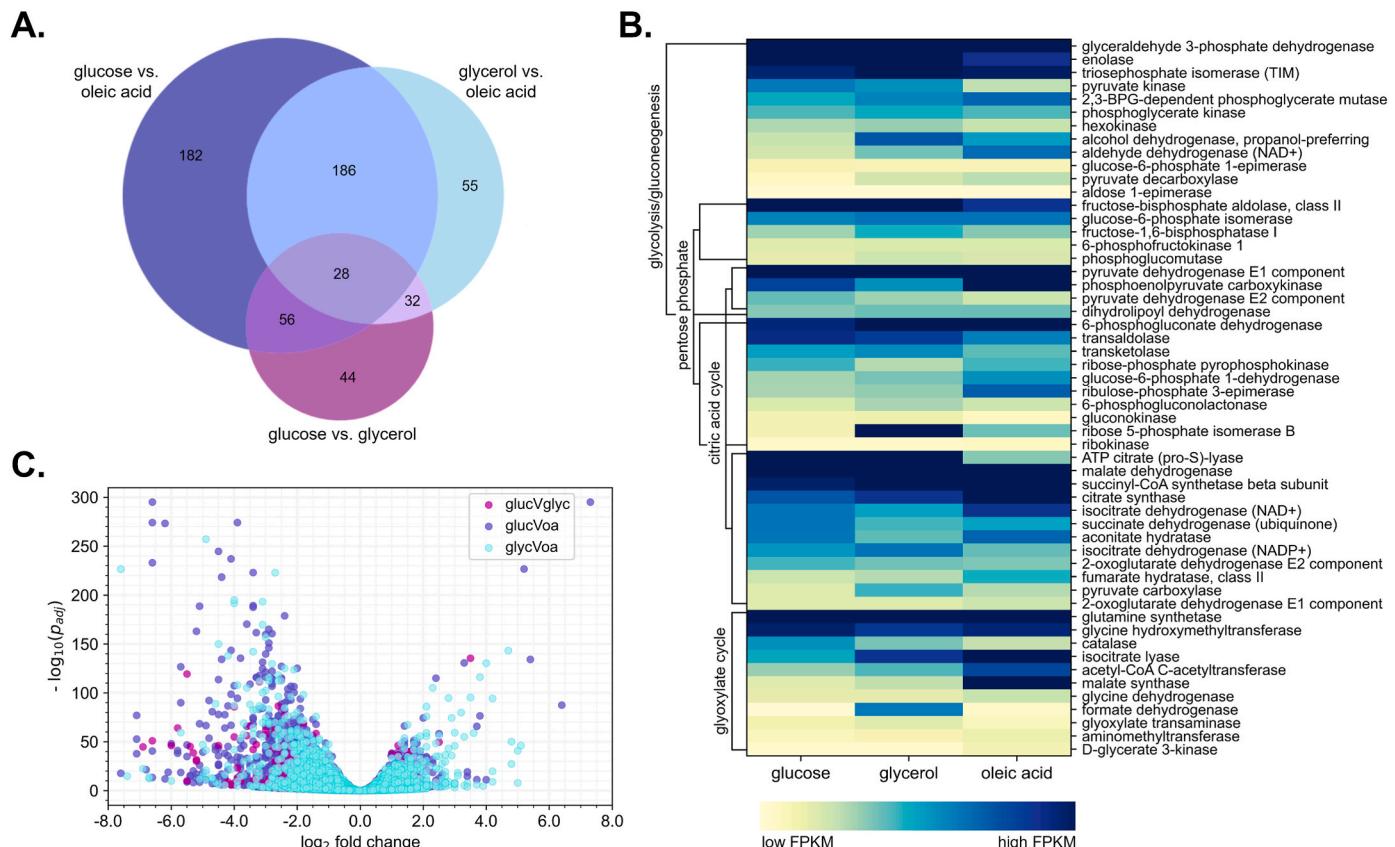


Fig. 2. Transcription data analysis. (A) Venn diagram showing differentially expressed gene count ($p_{\text{adj}} < 0.05$) with an absolute \log_2 fold change greater than 2. (B) Heat map showing transcript values for the major central carbon pathways. Expression levels are shown in fragments per kilobase of transcript per million mapped reads (FPKM). (C) Volcano plot showing the differential gene expression in *Y. lipolytica* compared across the three substrates. Data is plotted as \log_2 fold change versus $-\log_{10}(p_{\text{adj}})$. Abbreviations used to denote the carbon substrate: gluc = glucose, glyc = glycerol, oa = oleic acid, where glucVoa is the differential gene expression of glucose versus oleic acid, etc.

refinement of the ^{13}C -MFA models and reduce solution spaces. Ultimately, one reaction was omitted from the glucose model and two reactions were omitted from the oleic acid model.

Specifically, carnitine acetyltransferase is the enzyme responsible for transporting acetyl-CoA between the cytosol and mitochondria, also referred to as the carnitine shuttle. It catalyzes the reversible transfer of acetyl groups from acetyl-CoA to carnitine, forming an acetyl-carnitine complex that can then be transported across the mitochondrial membrane. On the other side, the acetyl-CoA can be separated by another carnitine acetyltransferase. Transcript expression analysis showed genes involved in the carnitine shuttle (*YICAT2*: YALI0B10340g, *YIYAT1*: YALI0F21197g, and *YICRC1*: YALI0C02431g) (Messina et al., 2023) were downregulated in glucose cultures compared to glycerol and oleic acid ($p_{\text{adj}} < 0.005$, Fig. 3A). This finding confirms a previous report of carnitine acetyltransferase upregulation in cells grown on oleic acid compared to glucose (Messina et al., 2023). We used this evidence to remove acetyl-CoA compartmental transport from the glucose ^{13}C -MFA model. By removing this reaction, the average range of 95% confidence intervals (measured by the difference between the lower and upper bounds) was reduced by 84%. In glycerol-grown cells, while one

carnitine acetyltransferase gene (YALIB10340g) was downregulated by a \log_2 fold change of 2.6 compared to oleic acid, the other (YALI0F21197g) was only downregulated by a \log_2 fold change of 0.7, and in fact upregulated by a \log_2 fold change of 1.8 in glycerol cultures compared to glucose ($p_{\text{adj}} < 0.005$, Fig. 3A), a factor supporting the inclusion of this reaction in the glycerol ^{13}C -MFA model. Importantly, when the acetyl-CoA transport reaction was removed from the glycerol model, no acceptable flux solutions were found, as determined by the sum of squared residuals between measured and simulated amino acid data. These results indicate the carnitine shuttle is an important regulator in the flux network of glycerol-grown cells. Therefore, this reaction was not removed from the glycerol or oleic acid ^{13}C -MFA models.

The oleic acid ^{13}C -MFA confidence intervals were impacted by two reactions. First, transcription data of oleic acid cultures showed downregulation of ATP citrate lyase (ACL: YALI0D24431g and YALI0E34793g), which catalyzes the cytosolic reaction of citrate to acetyl-CoA and oxaloacetate, compared to both glucose and glycerol ($p_{\text{adj}} < 0.005$, Fig. 3B). We used this evidence to remove the ACL reaction from the ^{13}C -MFA metabolic network for oleic acid. Second, a previous report showed that pyruvate carboxylase (YALI0C24101g),

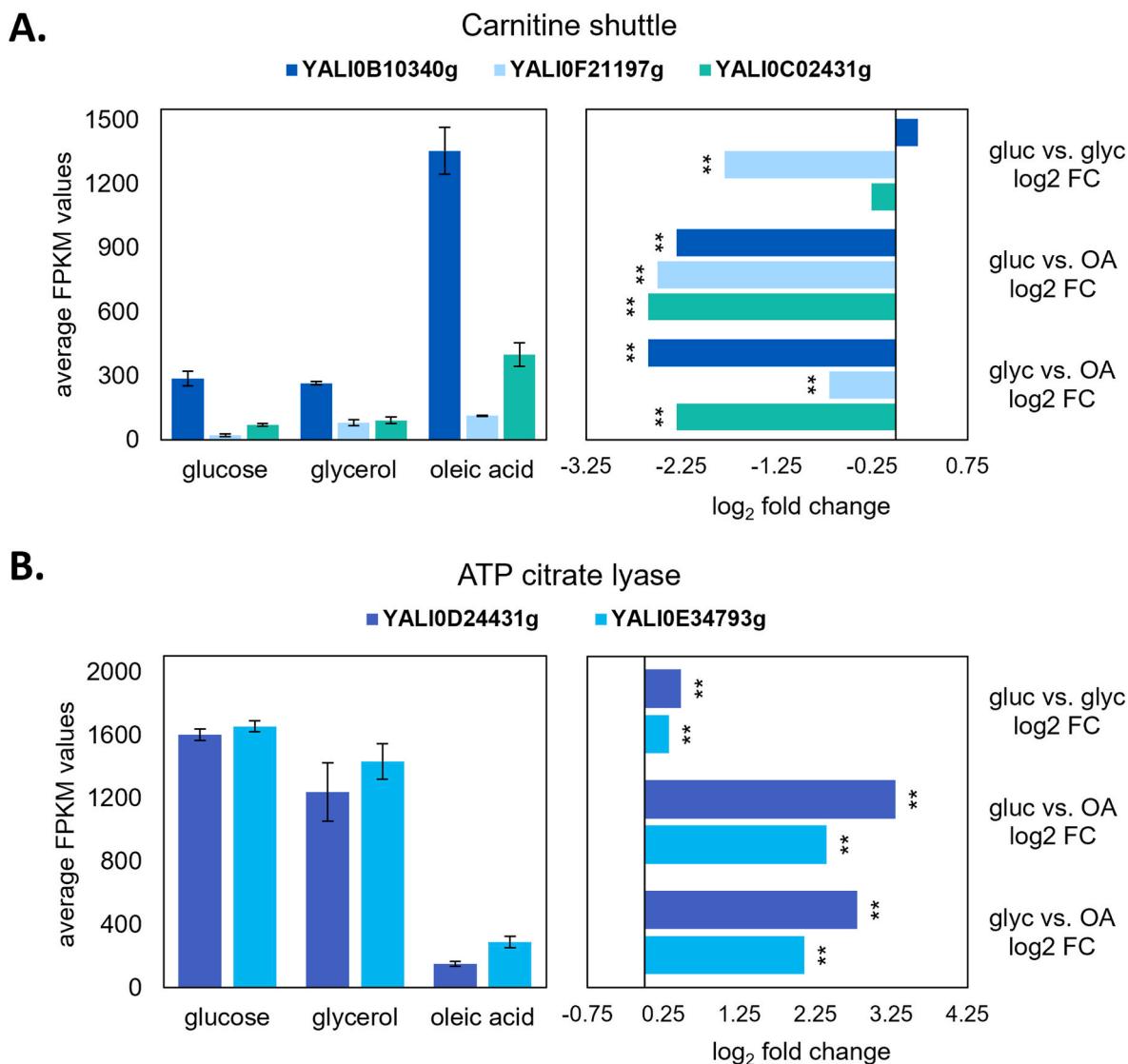


Fig. 3. Comparison of gene expression levels for (A) the carnitine shuttle and (B) ATP citrate lyase. Data is shown in average FPKM (left) and DGE analysis (right). Error bars on FPKM plots represent the standard deviation ($n = 3$). On DGE graphs, two asterisks (**) indicate $p_{\text{adj}} < 0.005$ statistical significance difference between the reported values. Abbreviations used to denote the carbon substrates: gluc = glucose, glyc = glycerol, OA = oleic acid, where gluc vs. OA is the differential gene expression comparing glucose to oleic acid, etc.

which catalyzes the carboxylation of pyruvate to form oxaloacetate, is not essential when the glyoxylate shunt is active (Flores and Gancedo, 2005). In the case of oleic acid cultures, the glyoxylate shunt has very high expression and flux (Figs. 5A and 4, respectively), thus, pyruvate carboxylase was removed from the model for oleic acid flux calculations. A previous ^{13}C -MFA report on *Y. lipolytica* acetate metabolism also omitted both ATP citrate lyase and pyruvate carboxylase reactions from their model (Liu et al., 2016). By excluding these reactions from the oleic acid ^{13}C -MFA calculations, the model confidence improved by 74%, measured by the reduction in the average range of 95% confidence intervals.

3.4. ^{13}C -MFA results

^{13}C -MFA results are shown in Fig. 4. For the glucose model, the sum of squared residuals (SSR) between measured and simulated amino acid data was 95.8, within the expected range of [80.3, 137.5]. For glycerol, the SSR was 69.9, within the range of [41.3, 84.5]. Finally, the oleic acid SSR was 78.8, within the range of [72.5, 127.3]. During oleic acid growth, the glyoxylate shunt is highly active, generating oxaloacetate for anaplerotic support to PEP and flux through gluconeogenesis. Of the acetyl-CoA generated through the β -oxidation of oleic acid, 42% is routed directly to the glyoxylate shunt through malate synthase, and 39% is transported into the mitochondria via the carnitine shuttle. The remaining 19% is used for biomass synthesis. Of the 39% of acetyl-CoA transported into the mitochondria, 76% is nearly immediately transported back into the cytosol in the form of citrate, which then enters the glyoxylate shunt following conversion to isocitrate and cleavage to succinate and glyoxylate via isocitrate lyase. Ultimately, over 80% of the acetyl-CoA generated from oleic acid is directed toward the glyoxylate shunt. The cytosolic succinate generated from isocitrate lyase is returned to the mitochondria to power the TCA cycle. Low TCA cycle fluxes are observed through isocitrate and α -ketoglutarate during oleic acid catabolism due to 91% of the mitochondrial citrate pool being transported out to the cytosol. The citrate remaining in the mitochondria generates flux through this part of the TCA cycle to support α -ketoglutarate's role in amino acid production. The low flux here does accomplish minimizing carbon loss as both isocitrate dehydrogenase (IDH) and α -ketoglutarate dehydrogenase (α -KGDH) generate CO_2 . By reducing the flux through these two reactions, only 1 in 4 carbons from oleic acid are lost in route to PEP. This low flux could also be due in part to changes in cell regulation attempting to balance the abundance of NADH generated through β -oxidation. In response, cells reduce flux through α -KGDH and NAD^+ -dependent IDH to prevent additional NADH generation. However, low flux through isocitrate dehydrogenase also suggests there's also low activity of the NADP^+ -dependent IDH. The compartmental pools of NADH and NADPH are regulated indirectly through the transport of TCA cycle metabolites (Xiao et al., 2017), thus requiring cells to adjust flux networks in response to balance these cofactors. Comparison of our long-chain fatty acid ^{13}C -MFA model with a previously published acetate model, wherein both substrates are converted to acetyl-CoA upon catabolism, we see certain similarities (Liu et al., 2016). Specifically, flux through the glyoxylate shunt is high in both models, but oleic acid gives higher glyoxylate shunt fluxes and lower TCA cycle fluxes through IDH and α -KGDH compared to acetate metabolism. Both models also determined the oxidative pentose phosphate pathway to be a major contributor to the NADPH pool.

In comparison to oleic acid, ^{13}C -MFA with glycerol cultures indicates that all TCA cycle enzymes are active and only 22% of citrate is transported out of the mitochondria, similar to a previous report (Sabre et al., 2017). Flux through the glyoxylate shunt is considerably lower in glycerol cultures than oleic acid. Following glycerol uptake, approximately 62% of the flux is driven towards PEP and the TCA cycle, while the other 38% enters gluconeogenesis towards the pentose phosphate pathway for NADPH generation. During glycerol catabolism, the cytosolic acetyl-CoA pool is supported by the carnitine shuttle, transporting

36% of mitochondrial acetyl-CoA to the cytosol following generation via pyruvate dehydrogenase. The carnitine shuttle is the primary contributor of the cytosolic acetyl-CoA pool, with a small portion generated via acetyl-CoA synthase, and less than 1% generated via ACL, as determined by ^{13}C -MFA (Supplementary File 2).

For glucose metabolism, our findings are similar to previous reports (Sabre et al., 2017; Wasylewko et al., 2015). Specifically, 75% of the citrate transported from the mitochondria to the cytosol is used to generate the cytosolic acetyl-CoA pool via ACL. This reaction is the primary source of cytosolic acetyl-CoA, with only 8% coming from the anaplerotic reaction acetyl-CoA synthase. It is important to note that even when the acetyl-CoA transport reaction was included in the glucose model, though the confidence intervals were large, acetyl-CoA flux from the mitochondria to the cytosol was not evidenced by ^{13}C -MFA. 30% of the mitochondrial citrate pool remains to participate in the TCA cycle. Supplementary Figs. 9–11 show comparisons of our ^{13}C -MFA results with previously published models.

3.5. GSMS integrated with ^{13}C -MFA

GSMS include peripheral reaction pathways and can provide insights beyond the resolution of ^{13}C -MFA. However, given GSMS' high degree of freedom, they require extensive data to constrain with precision. In this work, an iterative approach narrowed the feasible range of flux values for each reaction in the GSMS (Fig. 5A). The initial constraint came from reaction reversibility. For example, if a reaction was irreversible, it would have a range of [0, 1000], and if it was reversible, it would have a range of [-1000, 1000]. Important to note, the default constraints for oleic acid were expanded to [0, 10,000] and [-10000, 10,000], respectively, to account for the larger flux values from the 18 carbon atoms in oleic acid. This expansion prevented reaction rate limitations by arbitrary flux bounds. Next, flux variability analysis (FVA) was conducted with a biomass cutoff value determined by yield coefficients for each substrate (see section 2.5) to find the feasible flux range of each reaction. The flux bounds determined by this step ensured that the ^{13}C -MFA results would be compatible with the GSMS. Then, the ^{13}C -MFA 95% confidence intervals were used to constrain the central carbon reactions of the GSMS (Supplementary Figs. 3–5). This approach significantly reduced the reaction rates bound ranges for GSMS. For example, the malate synthase flux for oleic acid initially spanned [0, 10,000]. FVA with the biomass constraint reduced the upper limit to 1241, while ^{13}C -MFA further refined the flux range to [355, 422]. Finally, running FVA with ^{13}C -MFA constraints narrowed the reactions' flux bounds to [355, 356] (Fig. 5B). This approach of combining ^{13}C -MFA and FVA consistently tightened the GSMS reaction flux bounds in each carbon source, halving the average range of all GSMS reaction fluxes compared to when only biomass constraints were used (Fig. 5C). This strategy demonstrates that integrating ^{13}C -MFA constraints with FVA enhances the feasible flux resolution within GSMS predictions.

4. Discussion

4.1. Correlation between gene expressions and fluxomes

This work evaluated the correlation between ^{13}C -MFA and expression data to validate the flux results. For example, the glyoxylate shunt during oleic acid catabolism was measured to have both high flux and high gene expression. Isocitrate lyase (YALI0C16885g and YALI0F31999g) and malate synthase (YALI0D19140g and YALI0E15708g) were significantly upregulated in oleic acid cultures compared to glucose and glycerol ($P_{\text{adj}} < 0.005$, Fig. 6A). The glyoxylate shunt is highly active during oleic acid metabolism to drive flux towards gluconeogenesis for NADPH generation, and as a result allows for less carbon loss from reducing flux through IDH and α -KGDH in the TCA cycle. The increased expression and flux of the glyoxylate shunt aligns with its essential function in the metabolism of oleic acid. Additionally,

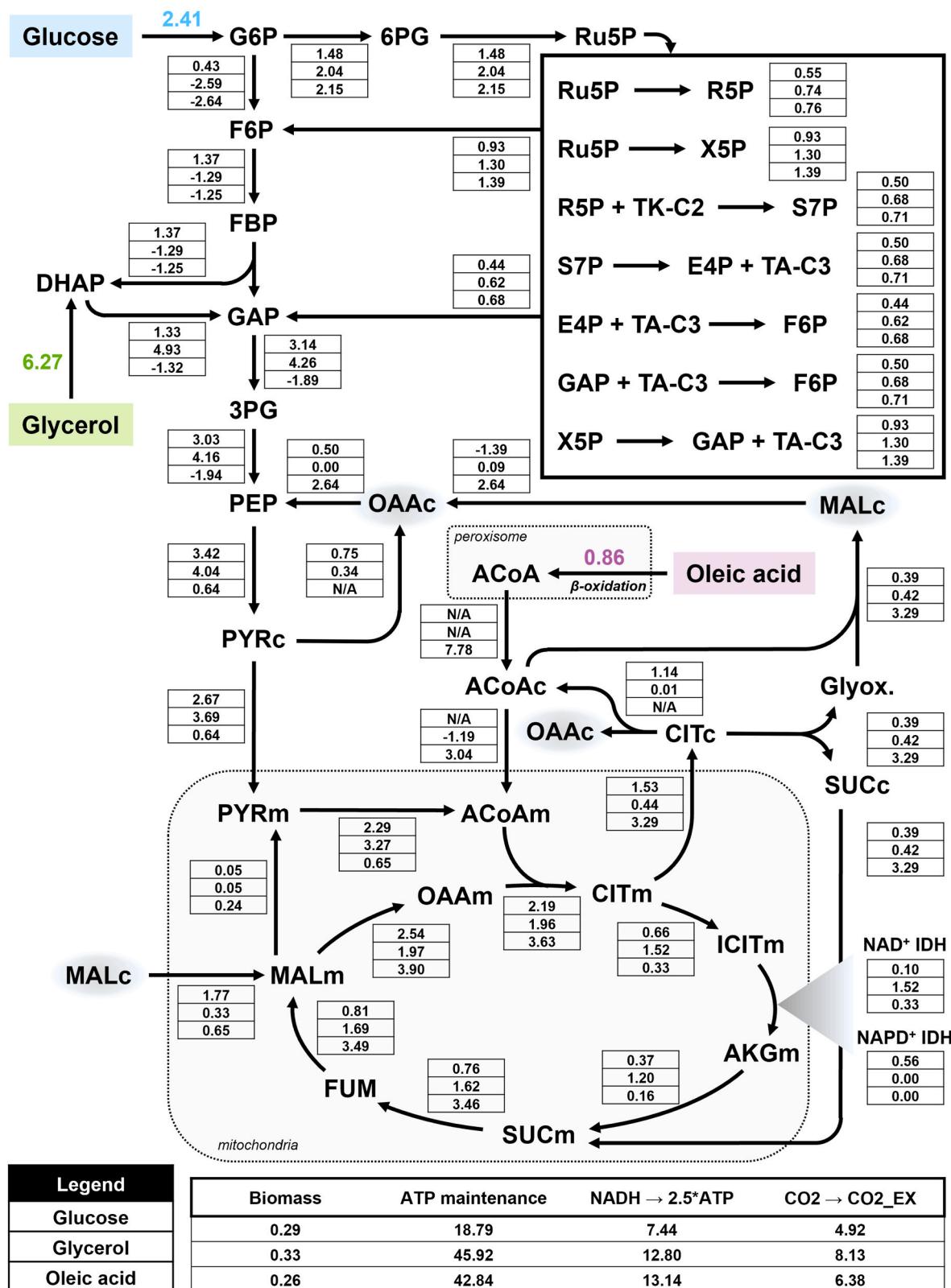


Fig. 4. Metabolic flux distributions. A legend is shown on the bottom left, indicating which flux values represent the respective substrates. Flux values are normalized to the specific substrate uptake rate of each carbon source (Supplementary Table 1, mmol substrate gDCW⁻¹ h⁻¹). Flux values listed as N/A indicate the reaction was omitted from the metabolic model. Negative values denote a net flux in the reverse direction of the arrow. MALc and OAAc, the cytosolic pools for malate and oxaloacetate, respectively, are highlighted to indicate that they appear in two places on the flux map. Values for biomass generation, ATP maintenance, cofactor balance, and CO₂ exchange are shown in the table beneath the flux map. A complete list of best-fit values, 95% CIs, and 68% CIs can be found in Supplementary File 2.

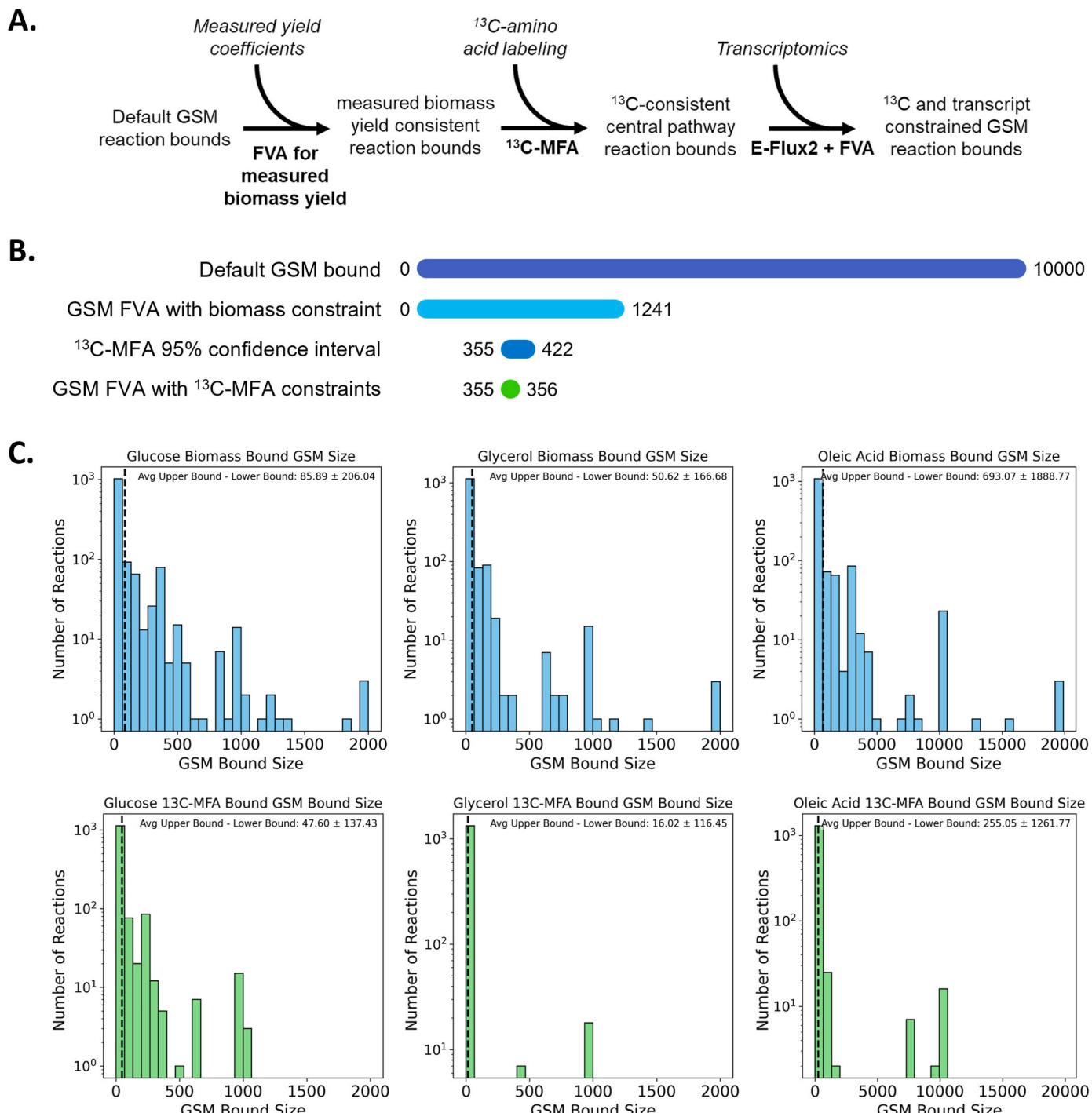


Fig. 5. GSM constraint of GSM bounds. (A) GSM constraint is achieved in a series of stages, beginning with the default reaction bounds and iteratively constrained by measured yield and ^{13}C -MFA via flux variability analysis (FVA). (B) Confidence intervals for the malate synthase reaction with oleic acid demonstrate the effects of constraint implementation. In this example, the uptake is 100 mmol of oleic acid. Bars are not drawn to scale; bounds are denoted by values listed. (C) Distributions of GSM bound sizes for *Y. lipolytica* grown on glucose, glycerol, and oleic acid, with the biomass yield constraint (top) and with ^{13}C -MFA constraints (bottom). For oleic acid, the default flux bound was scaled to 10,000 to prevent artificially restricted fluxes, whereas the default bounds were 1000 for glucose and glycerol.

oleic acid ^{13}C -MFA results showed that the mitochondrial acetyl-CoA pool was primarily supported by transport from the cytosol via the carnitine shuttle, not from pyruvate dehydrogenase (PDH). The transport flux is roughly 4.7x higher than the mitochondrial PDH flux. Transcript data helps validate this claim, showing downregulation of PDH (YAL10E27005g, YAL10F20702g, and YAL10D23683g) in oleic acid cultures compared to the other carbon sources. (Fig. 6B). Additionally, citrate synthase was measured to have the highest flux in oleic acid

cultures, and citrate synthase genes (YAL10E00638g and YAL10E02684g) were upregulated in oleic acid compared to glucose and glycerol cultures (Fig. 6B). These results offer further evidence to support the ^{13}C -MFA findings.

4.2. Energy generation

The oxidative pentose phosphate pathway is the primary source of

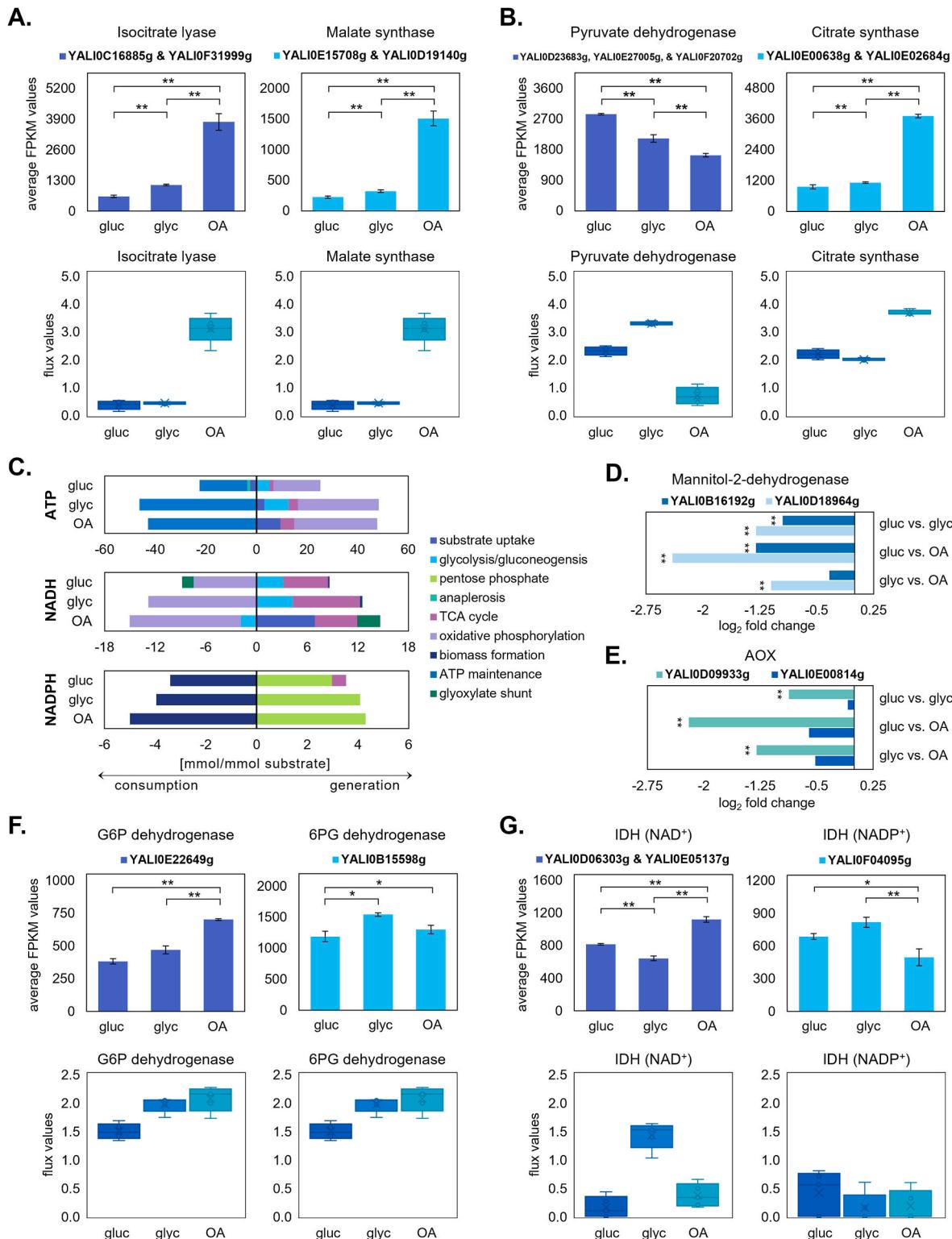


Fig. 6. Comparison of flux and transcript values during *Y. lipolytica* growth. (A) glyoxylate shunt and (B) TCA cycle reactions. (C) Sources and sinks of ATP, NADH, and NADPH as determined by ^{13}C -MFA. (D) DGE analysis of mannitol-2-dehydrogenase (MDH) genes and (E) alternative oxidase (AOX) genes. (F) The oxidative pentose phosphate pathway: G6P dehydrogenase (left) 6 PG dehydrogenase (right), and (G) Isocitrate dehydrogenase, NAD⁺-dependent IDH (right) and NADP⁺-dependent IDH (left). In (A), (B), (F), and (G), the top graphs show the average FPKM counts from RNA transcription data. Error bars represent standard deviation ($n = 3$). One asterisk (*) indicates statistical significance $p_{\text{adj}} < 0.05$ and two asterisks (**) indicate $p_{\text{adj}} < 0.005$ between the reported values. The bottom graphs show flux values and confidence intervals from ^{13}C -MFA. Best-fit values, 68% confidence intervals, and 95% confidence intervals are represented by the lines in the middle, the boxes, and the error bars, respectively. Abbreviations in this figure used to denote the carbon substrate: gluc = glucose, glyc = glycerol, OA = oleic acid, where gluc vs. OA is the differential gene expression comparing glucose to oleic acid, etc.

NADPH generation during fatty acid catabolism as evidenced by high flux through this pathway (Figs. 4 and 6C). However, there are other potential sources of NADPH that ^{13}C -MFA is not able to evaluate, e.g. the mannitol cycle, folate cycle, and NADP^+ -dependent isocitrate dehydrogenase (IDH). Thus, we evaluated the gene expression of these sources. RNA-sequencing suggests that the mannitol cycle and the oxidative pentose phosphate pathway are two major sources of NADPH during fatty acid catabolism (Fig. 6D and F). In the oxidative pentose phosphate pathway, glucose-6-phosphate (G6P) dehydrogenase (YALI0E22649g) is upregulated in oleic acid cultures compared to glucose and glycerol ($p_{\text{adj}} < 0.005$), and 6-phosphogluconate (6 PG) dehydrogenase (YALI0B15598g) is upregulated in glycerol and oleic acid cultures compared to glucose ($p_{\text{adj}} < 0.05$). These findings agree with the ^{13}C -MFA results. Transcript analysis also showed that mannitol-2-dehydrogenase (MDH: YALI0B16192g and YALI0D18964g) is upregulated in glycerol and oleic acid cultures compared to glucose ($p_{\text{adj}} < 0.005$), suggesting this may be a secondary source of NADPH in cells cultured with long-chain fatty acids. Supporting this, we found that an MDH knockout strain (YALI0D18964g, see Supplementary Materials and Methods for details) produced 15% less biomass and had a 16% reduction in the ratio of NADPH/ NADP^+ during oleic acid catabolism compared to the parent strain (Supplementary Fig. 12). NADPH is critical for cell growth in the synthesis of amino acids and nucleotides. During cultivation with either glycerol or glucose, however, there was no observed difference in biomass production between the knockout and parent strains. These findings support the conclusion that the mannitol cycle is likely an important contributor of NADPH during growth with long-chain fatty acids. The mannitol cycle effectively functions as an energy-dependent transhydrogenase reaction, generating one NADPH and consuming one NADH and one ATP. Fructose-6-phosphate (F6P) is converted to mannitol-1-phosphate (M1P), consuming one NADH, and M1P is converted to mannitol. Mannitol is converted to fructose via MDH, generating one NADPH. Fructose is then converted back to F6P with the consumption of one ATP. While cultures growing on oleic acid still need to drive flux through gluconeogenesis to power the mannitol cycle, this cycle offers a way for cells to balance excess NADH generated through the β -oxidation of fatty acids. Additionally, alternative oxidase (AOX: YALI0D09933g and YALI0E00814g) shows upregulation in oleic acid and glycerol cultures (Fig. 6E). This is another route cells may use to reduce the excess NADH pool, helping cells maintain metabolic homeostasis by generating NAD^+ to support the TCA cycle. Importantly, AOX activity does not generate ATP, so energy created in this reaction is released as heat (Rogov and Zvyagilskaya, 2015).

RNA-sequencing shows that NADP^+ -dependent IDH (YALI0F04095g) is downregulated in oleic acid cultures compared to glucose and glycerol ($p_{\text{adj}} < 0.05$), whereas NAD^+ -dependent IDH genes (YALI0D06303g and YALI0E05137g) are upregulated in oleic acid cultures ($p_{\text{adj}} < 0.005$) (Fig. 6G). Due to the identical carbon transitions in both IDH reactions, ^{13}C -MFA is not able to distinguish between them with high resolution. Further, it is important to keep in mind the compartmental pools of NADH and NADPH in the cytosol and the mitochondria. These cofactors cannot cross the mitochondrial membrane directly, rather are regulated indirectly through the compartmental transport of TCA cycle metabolites (Xiao et al., 2017). Therefore, regulation of NADP^+ -dependent IDH is primarily related to the need of NADPH in only the mitochondria. Additionally, we found that the two genes encoding methylenetetrahydrofolate synthase (YALI0E01056g and YALI0F30745g) in the folate cycle were both downregulated in oleic acid compared to glucose ($p_{\text{adj}} < 0.05$). A previous GSM study shows that the folate cycle is not a significant source of NADPH for lipid biosynthesis during glucose based growth (Kavšek et al., 2015).

The ^{13}C -MFA-constrained GSM flux predictions were used to determine which reactions were sources or sinks of cytosolic NADPH (Supplementary Figs. 6–8). With all three carbon sources, the oxidative pentose phosphate pathway was the major contributor of NADPH. The ^{13}C -MFA-constrained GSM predictions indicate that

methylenetetrahydrofolate dehydrogenase is a source of roughly one third of the cytosolic NADPH when glycerol is the substrate, while it is unclear whether this enzyme is a source or sink of NADPH in glucose or oleic acid. Glutamate dehydrogenase is in the primary sink of NADPH with glycerol. This agrees with the ^{13}C -MFA finding that biomass formation is the main pathway for NADPH consumption (Fig. 6C).

4.3. Implication of biosynthesis of acetyl-CoA derived products

Y. lipolytica has emerged as an important host for the production of terpene compounds and other acetyl-CoA derived products because of its native mevalonate pathway. We examined the expression changes in this pathway and other important cellular processes in *Y. lipolytica* during growth on the three carbon substrates (Fig. 7A and B). Glycerol cultures increase expression in terpenoid backbone synthesis genes compared to glucose and oleic acid, especially in the case of mevalonate kinase (MK: YALI0B16038g) (Fig. 7A). Notably, genes involved in coenzyme A (CoA) synthesis are upregulated in both oleic acid and glycerol cultures compared to glucose (Fig. 7B). These findings indicate that TAG substrates may be an excellent choice for terpene production in *Y. lipolytica*.

Using the GSM to estimate theoretical yields, our analysis showed that *Y. lipolytica* could more efficiently produce β -carotene when co-fed with glucose and oleic acid than when fed with either carbon source alone (Fig. 7C, Supplementary Fig. 13). On a gram per gram basis, a higher predicted yield of β -carotene was determined from oleic acid, given that it has a higher degree of reduction, and therefore has more carbon per gram of substrate than glucose. In Fig. 7C, the dashed line represents the theoretical yield if no benefit from co-feeding existed. However, we found that the theoretical yield of β -carotene when co-fed was higher than the sum of the yields from separate cultivation of the two substrates, at any given ratio. This synergistic effect is demonstrated by the difference between the gray and colored dotted lines. This effect can also be seen by tracking how co-feeding glucose and oleic acid reduces the theoretical amount of carbon lost as CO_2 . With co-feeding, glucose catabolism produces NADPH in the pentose phosphate pathway, and the carbon from oleic acid can go towards product formation. This dual-substrate approach can optimize the use of available carbon by balancing the segregated metabolism, enhancing the overall efficiency of β -carotene biosynthesis in *Y. lipolytica*. This effect was also observed in a previous study (Worland et al., 2020b).

Y. lipolytica has gained recent interest to function as a microbial biorefinery converting waste oils to value-added products. These oils are typically ill-defined and can contain oxidative contaminants (Kumar and Negi, 2015). Further, lipid oxidation generates reactive oxygen species (ROS) (Xu et al., 2017), which can impact cellular functions and induce cell morphology changes (Li et al., 2011; Mano et al., 2002). A previous study showed that during *Y. lipolytica* growth on oil with elevated peroxide levels (> 20 meq peroxide/kg), the peroxide value decreased by 50% within the first day of fermentation (Li et al., 2020). *Y. lipolytica* is tolerant to oxidative stress, but previous work has shown when cells were grown with glucose, there was an inverse relationship between lipid biosynthesis and the presence of ROS and aldehydes (Xu et al., 2017). In this study, DGE analysis indicated that genes involved in oxidative stress response were upregulated in glycerol and oleic acid cultures compared to glucose (Fig. 7B). Further, the genes needed for critical cell functions, including genetic processing and meiosis, may be better supported in *Y. lipolytica* cultured with TAGs as opposed to with fatty acids alone, as genes in these pathways are upregulated in glycerol cultures compared to oleic acid (Fig. 7B).

5. Conclusions

This study integrated ^{13}C -MFA, GSM, and transcriptomics data to investigate the metabolic pathways and regulatory mechanisms of *Y. lipolytica* when grown with different carbon sources. By conducting

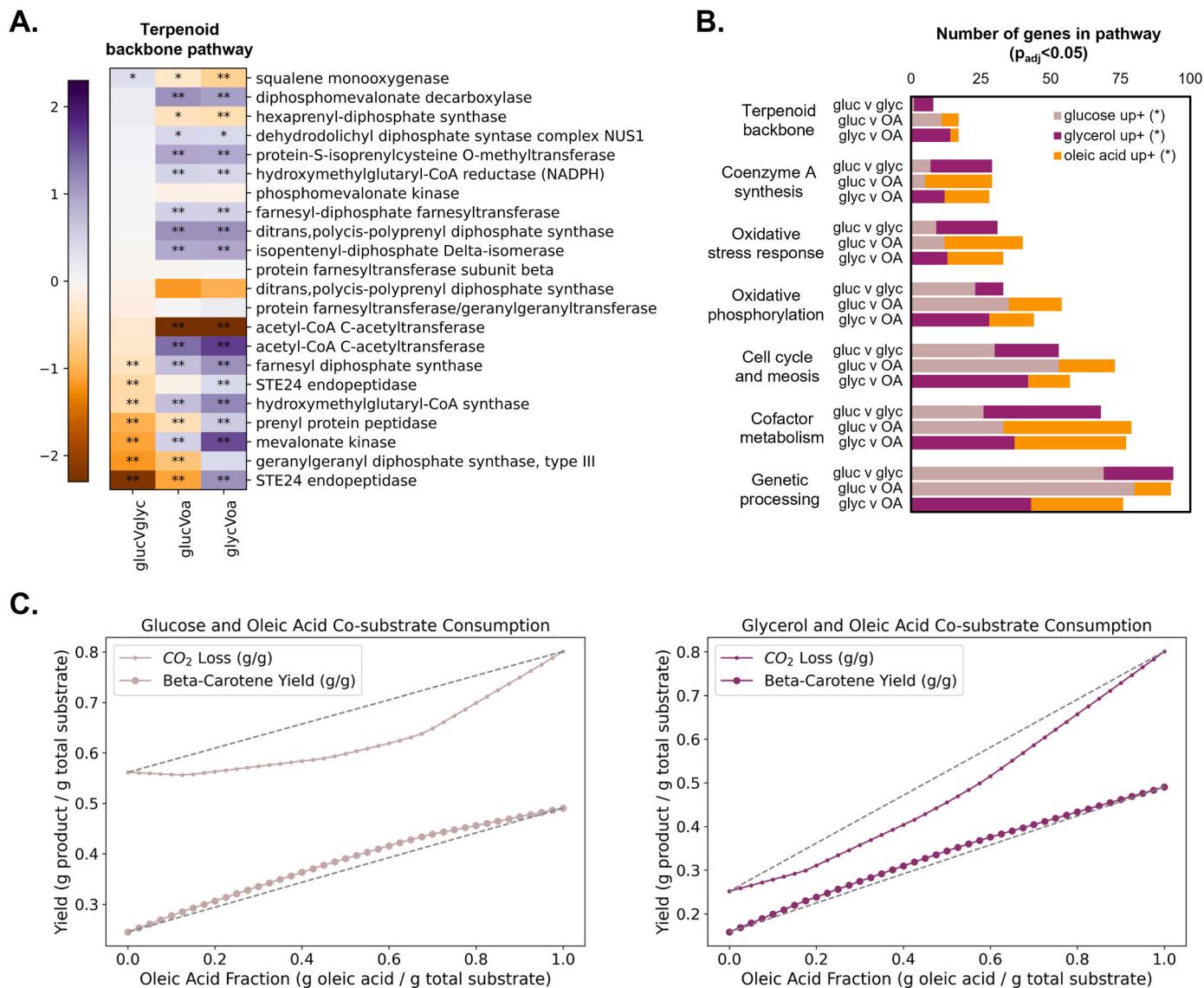


Fig. 7. DGE analysis of cell processes and pathways. (A) Heat map showing the differential expression of genes involved in terpenoid backbone synthesis. A positive log₂ fold change (purple) indicates higher expression in the former substrate, and a negative log₂ fold change (orange) indicates higher expression in the latter substrate. One asterisk (*) indicates $p_{adj} < 0.05$ and two asterisks (**) indicate $p_{adj} < 0.005$. (B) Comparison of differentially expressed genes in the pathways shown. Genes in each pathway were identified by KEGG metabolic pathways and the complete list detailing the specific pathways used in each group is provided in [Supplementary Table 2](#). Tan bars represent glucose upregulation of genes in the pathway compared to the substrate shown adjacent. In the same way, purple bars represent glycerol upregulation and orange bars represent oleic acid upregulation compared to the substrate adjacent. All data shown is statistically significant ($p_{adj} < 0.05$) differential expression. (C) Theoretical yield of β-carotene and associated CO₂ loss with glucose and oleic acid co-feeding (left) and glycerol and oleic acid co-feeding (right). Yield is given in grams of β-carotene per gram of total substrate. The gray dotted line is the predicted sum of the yields when the two substrates are cultivated separately. The colored dotted line is the yield prediction for co-utilization of the two substrates. Abbreviations used in this figure to denote the carbon substrate: gluc = glucose, glyc = glycerol, oa = oleic acid, where glucVoa is the differential gene expression comparing glucose to oleic acid, etc.

the first long-chain fatty acid ¹³C-MFA in *Y. lipolytica* with oleic acid, we found that oleic acid metabolism is dependent on high flux through the glyoxylate shunt while glycerol metabolism relies on the carnitine shuttle. In addition, holistic views of gene expression and GSM further highlight the activation of latent pathways and the increase of metabolic burdens during long-chain fatty acid catabolism. The new understanding of metabolic regulation and flux networks gained in this work offers guidelines for engineering *Y. lipolytica* for industrial biomanufacturing and insights into how eukaryote cells can catabolize TAGs as a sole carbon source.

Code Availability

The code and genome scale model developed for this project are open

source and are available at <https://github.com/UH-MBBE/yarrowia-13C-gsm>.

Declaration of interest

The authors declare no competing interests.

CRediT authorship contribution statement

Alyssa M. Worland: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation. **Zhenlin Han:** Investigation. **Jessica Maruwan:** Investigation. **Yu Wang:** Investigation. **Zhi-Yan Du:** Writing – review & editing, Supervision. **Yinjie J. Tang:** Writing – review & editing, Supervision, Resources,

Funding acquisition, Conceptualization. **Wei Wen Su**: Writing – review & editing, Supervision, Resources, Funding acquisition, Conceptualization. **Garrett W. Roell**: Writing – review & editing, Visualization, Software, Methodology, Investigation, Data curation.

Data availability

Code and GSM: <https://github.com/UH-MBBE/yarrowia-13C-gsm>
 Supplementary File 2: https://raw.githubusercontent.com/UH-MBBE/yarrowia-13C-gsm/main/files/supplementary_file_2.xlsx

Acknowledgements

This work is supported in part by AFRI (award number 2020-67022-31146) from the USDA NIFA to WWS, and by NSF CBET award number 2225809 to YJT.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ymben.2024.06.010>.

Abbreviations

| | |
|---------------------|---|
| ¹³ C-MFA | ¹³ C-metabolic flux analysis |
| 3 PG | 3-phosphoglycerate |
| 6 PG | 6-phosphogluconate |
| ACL | ATP citrate lyase |
| ACoA | acetyl-CoA |
| AcoAc | cytosolic acetyl-CoA |
| AcoAm | mitochondrial acetyl-CoA |
| AKG | α -ketoglutarate |
| α -KGDH | α -ketoglutarate dehydrogenase |
| AOX | alternative oxidase |
| CIT | citrate |
| CFU | colony forming units |
| DHAP | dihydroxyacetone phosphate |
| DGE | differentially expressed gene |
| E4P | erythrose-4-phosphate |
| F6P | fructose-6-phosphate |
| FBA | flux balance analysis |
| FBP | fructose-1,6-bisphosphate |
| FPKM | fragments per kilobase of transcript per million mapped reads |
| FUM | fumarate |
| FVA | flux variability analysis |
| G6P | glucose-6-phosphate |
| GAP | glyceraldehyde-3-phosphate |
| gDCW | grams dry cell weight |
| gluc | glucose |
| glyc | glycerol |
| Glyox | glyoxylate |
| GSM | genome-scale metabolic model |
| ICIT | isocitrate |
| IDH | isocitrate dehydrogenase |
| MAL | malate |
| MDH | mannitol-2-dehydrogenase |
| MFA | metabolic flux analysis |
| MID | mass isotopomer distribution |
| OAA | oxaloacetate |
| OA | oleic acid |
| PDH | pyruvate dehydrogenase |
| PEP | phosphoenolpyruvate |
| PYR | pyruvate |
| R5P | ribose-5-phosphate |
| Ru5P | ribulose-5-phosphate |

| | |
|-------|--|
| SUC | succinate |
| TA-C3 | 3-carbon fragment transferred by transaldolase |
| TK-C2 | 2-carbon fragment transferred by transketolase |
| X5P | xylulose-5-phosphate |

References

Antoniewicz, M.R., Kelleher, J.K., Stephanopoulos, G., 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* 8, 324–337.

Awogbemi, O., Kallon, D.V., 2022. Valorization of agricultural wastes for biofuel applications. *Heliyon* 8, E11117.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrerero, D., Cervera, A., McPherson, A., Szczęśniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol.* 17, 13.

Czajka, J.J., Nathenson, J.A., Benites, V.T., Baidoo, E.E.K., Cheng, Q., Wang, Y., Tang, Y. J., 2018. Engineering the oleaginous yeast *Yarrowia lipolytica* to produce the aroma compound β -ionone. *Microb. Cell Factories* 17, 136.

da Veiga Moreira, J., Jolicœur, M., Schwartz, L., Peres, S., 2021. Fine-tuning mitochondrial activity in *Yarrowia lipolytica* for citrate overproduction. *Sci. Rep.* 11, 878.

Dulermo, T., Lazar, Z., Dulermo, R., Rakicka, M., Haddouche, R., Nicaud, J.-M., 2015. Analysis of ATP-citrate lyase and malic enzyme mutants of *Yarrowia lipolytica* points out the importance of mannitol metabolism in fatty acid synthesis. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1851, 1107–1117.

Edwards, J.S., Palsson, B.O., 2000. The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* 97, 5528–5533.

Erian, A.M., Egermeier, M., Marx, H., Sauer, M., 2022. Insights into the glycerol transport of *Yarrowia lipolytica*. *Yeast* 39, 323–336.

Flores, C.-L., Gancedo, C., 2005. *Yarrowia lipolytica* mutants devoid of pyruvate carboxylase activity show an unusual growth phenotype. *Eukaryot. Cell* 4, 356–364.

Gonçalves, F.A.G., Colen, G., Takahashi, J.A., 2014. *Yarrowia lipolytica* and its multiple applications in the biotechnological industry. *Sci. World J.* 2014, 476207.

Kavšek, M., Bhutada, G., Madl, T., Natter, K., 2015. Optimization of lipid production with a genome-scale model of *Yarrowia lipolytica*. *BMC Syst. Biol.* 9, 72.

Kulkarni, M.G., Dalai, A.K., 2006. Waste cooking oil: an economical source for biodiesel: a review. *Ind. Eng. Chem. Res.* 45, 2901–2913.

Kumar, S., Negi, S., 2015. Transformation of waste cooking oil into C-18 fatty acids using a novel lipase produced by *Penicillium chrysogenum* through solid state fermentation. *3 Biotech.* 5, 847–851.

Li, N., Han, Z., O'Donnell, T.J., Kurasaki, R., Kajihara, L., Williams, P.G., Tang, Y., Su, W. W., 2020. Production and excretion of astaxanthin by engineered *Yarrowia lipolytica* using plant oil as both the carbon source and the biocompatible extractant. *Appl. Microbiol. Biotechnol.* 104, 6977–6989.

Li, Q., Bai, Z., O'Donnell, A., Harvey, L.M., Hoskisson, P.A., McNeil, B., 2011. Oxidative stress in fungal fermentation processes: the roles of alternative respiration. *Biotechnol. Lett.* 33, 457–467.

Liu, N., Qiao, K., Stephanopoulos, G., 2016. ¹³C Metabolic Flux Analysis of acetate conversion to lipids by *Yarrowia lipolytica*. *Metab. Eng.* 38, 86–97.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.

Luévano-Martínez, L.A., Moyano, E., de Lacoba, M.G., Rial, E., Uribe-Carvajal, S., 2010. Identification of the mitochondrial carrier that provides *Yarrowia lipolytica* with a fatty acid-induced and nucleotide-sensitive uncoupling protein-like activity. *Biochim. Biophys. Acta Bioenerg.* 1797, 81–88.

Luttermann, T., Rückert, C., Wibberg, D., Busche, T., Schwarzhans, J.-P., Friehs, K., Kalinowski, J., 2021. Establishment of a near-contiguous genome sequence of the citric acid producing yeast *Yarrowia lipolytica* DSM 3286 with resolution of rDNA clusters and telomeres. *NAR Genom. Bioinf.* 3, lqab085.

Mahadevan, R., Schilling, C.H., 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* 5, 264–276.

Mano, J.i., Torii, Y., Hayashi, S.-i., Takimoto, K., Matsui, K., Nakamura, K., Inzé, D., Babiychuk, E., Kushnir, S., Asada, K., 2002. The NADPH:quinone oxidoreductase P1- ζ -crystallin in *Arabidopsis* catalyzes the α , β -hydrogenation of 2-alkenals: detoxification of the lipid peroxide-derived reactive aldehydes. *Plant Cell Physiol.* 43, 1445–1455.

Messina, E., de Souza, C.P., Cappella, C., Barile, S.N., Scarcia, P., Pisano, I., Palmieri, L., Nicaud, J.-M., Agrimi, G., 2023. Genetic inactivation of the Carnitine/Acetyl-Carnitine mitochondrial carrier of *Yarrowia lipolytica* leads to enhanced odd-chain fatty acid production. *Microb. Cell Factories* 22, 128.

Mishra, P., Lee, N.-R., Lakshmanan, M., Kim, M., Kim, B.-G., Lee, D.-Y., 2018. Genome-scale model-driven strain design for dicarboxylic acid production in *Yarrowia lipolytica*. *BMC Syst. Biol.* 12, 12.

Myers, J.A., Curtis, B.S., Curtis, W.R., 2013. Improving accuracy of cell and chromophore concentration measurements using optical density. *BMC Biophys.* 6, 4.

Nelson, B. S. Stephanie, 2016. Projected Availability of Fats, Oils, and Gases in the U.S. The International Council on Clean Transportation. ICCT. Working Paper 2016-15.

Niehus, X., Casas-Godoy, L., Rodríguez-Valadez, F.J., Sandoval, G., 2018. Evaluation of *Yarrowia lipolytica* oil for biodiesel production: land use oil yield, carbon, and energy balance. *J. Lipids* 2018, 6393749.

Orsavova, J., Misurcova, L., Ambrozova, J.V., Vicha, R., Mlcek, J., 2015. Fatty acids composition of vegetable oils and its contribution to dietary energy intake and

dependence of cardiovascular mortality on dietary intake of fatty acids. *Int. J. Mol. Sci.* 16, 12871–12890.

Papanikolaou, S., Aggelis, G., 2010. *Yarrowia lipolytica*: a model microorganism used for the production of tailor-made lipids. *Eur. J. Lipid Sci. Technol.* 112, 639–654.

Papanikolaou, S., Chevalot, I., Komaitis, M., Aggelis, G., Marc, I., 2001. Kinetic profile of the cellular lipid composition in an oleaginous *Yarrowia lipolytica* capable of producing a cocoa-butter substitute from industrial fats. *Antonie Leeuwenhoek* 80, 215–224.

Rogov, A.G., Zvyagilskaya, R.A., 2015. Physiological role of alternative oxidase (from yeasts to plants). *Biochemistry (Moscow)* 80, 400–407.

Sabra, W., Bommareddy, R.R., Maheshwari, G., Papanikolaou, S., Zeng, A.-P., 2017. Substrates and oxygen dependent citric acid production by *Yarrowia lipolytica*: insights through transcriptome and fluxome analyses. *Microb. Cell Factories* 16, 78.

Sawangkeaw, R., Ngamprasertsith, S., 2013. A review of lipid-based biomasses as feedstocks for biofuels production. *Renew. Sustain. Energy Rev.* 25, 97–108.

Schneider, P., Bekiaris, P.S., von Kamp, A., Klamt, S., 2022. StrainDesign: a comprehensive Python package for computational design of metabolic networks. *Bioinformatics* 38, 4981–4983.

Spagnuolo, M., Shabbir Hussain, M., Gambill, L., Blenner, M., 2018. Alternative substrate metabolism in *Yarrowia lipolytica*. *Front. Microbiol.* 9, 1077.

Wang, N., Chi, P., Zou, Y., Xu, Y., Xu, S., Bilal, M., Fickers, P., Cheng, H., 2020. Metabolic engineering of *Yarrowia lipolytica* for thermostability and enhanced erythritol productivity. *Biotechnol. Biofuels* 13, 176.

Wasylenko, T.M., Ahn, W.S., Stephanopoulos, G., 2015. The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in *Yarrowia lipolytica*. *Metab. Eng.* 30, 27–39.

Worland, A.M., Czajka, J.J., Li, Y., Wang, Y., Tang, Y.J., Su, W.W., 2020a. Biosynthesis of terpene compounds using the non-model yeast *Yarrowia lipolytica*: grand challenges and a few perspectives. *Curr. Opin. Biotechnol.* 64, 134–140.

Worland, A.M., Czajka, J.J., Xing, Y., Harper, W.F., Moore, A., Xiao, Z., Han, Z., Wang, Y., Su, W.W., Tang, Y.J., 2020b. Analysis of *Yarrowia lipolytica* growth, catabolism, and terpenoid biosynthesis during utilization of lipid-derived feedstock. *Metab. Eng. Commun.* 11, e00130.

Xiao, W., Wang, R.-S., Handy, D.E., Loscalzo, J., 2017. NAD(H) and NADP(H) redox couples and cellular energy metabolism. *Antioxidants Redox Signal.* 28, 251–272.

Xu, P., Qiao, K., Stephanopoulos, G., 2017. Engineering oxidative stress defense pathways to build a robust lipid production platform in *Yarrowia lipolytica*. *Biotechnol. Bioeng.* 114, 1521–1530.

You, L., Page, L., Feng, X., Berla, B., Pakrasi, H.B., Tang, Y.J., 2012. Metabolic pathway confirmation and discovery through (13)C-labeling of proteinogenic amino acids. *J. Vis. Exp.* e3583.

Young, J.D., 2014. INCA: a computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics* 30, 1333–1335.

Zhang, H., Wu, C., Wu, Q., Dai, J., Song, Y., 2016. Metabolic flux analysis of lipid biosynthesis in the yeast *Yarrowia lipolytica* using 13C-labeled glucose and gas chromatography-mass spectrometry. *PLoS One* 11, e0159187.

Zhang, Y., Nielsen, J., Liu, Z., 2017. Engineering yeast metabolism for production of terpenoids for use as perfume ingredients, pharmaceuticals and biofuels. *FEMS Yeast Res.* 17, fox080.