

Optimized genome-wide CRISPR screening enables rapid engineering of growth-based phenotypes in *Yarrowia lipolytica*



Nicholas R. Robertson^a, Varun Trivedi^b, Brian Lupish^a, Adithya Ramesh^b, Yuna Aguilar^a, Stephanie Carrera^b, Sangcheon Lee^b, Anthony Arteaga^e, Alexander Nguyen^b, Chase Lenert-Mondou^d, Marcus Harland-Dunaway^c, Robert Jinkerson^{b,e,*}, Ian Wheeldon^{b,e,*}

^a Bioengineering, University of California, Riverside, Riverside, CA, USA

^b Chemical and Environmental Engineering, University of California, Riverside, Riverside, CA, USA

^c Botany and Plant Sciences, University of California, Riverside, Riverside, CA, USA

^d Biochemistry, University of California, Riverside, Riverside, CA, USA

^e Center for Industrial Biotechnology, University of California, Riverside, Riverside, CA, USA

ABSTRACT

CRISPR-Cas9 functional genomic screens uncover gene targets linked to various phenotypes for metabolic engineering with remarkable efficiency. However, these genome-wide screens face a number of design challenges, including variable guide RNA activity, ensuring sufficient genome coverage, and maintaining high transformation efficiencies to ensure full library representation. These challenges are prevalent in non-conventional yeast, many of which exhibit traits that are well suited to metabolic engineering and bioprocessing. To address these hurdles in the oleaginous yeast *Yarrowia lipolytica*, we designed a compact, high-activity genome-wide sgRNA library. The library was designed using DeepGuide, a sgRNA activity prediction algorithm and a large dataset of ~50,000 sgRNAs with known activity. Three guides per gene enables redundant targeting of 98.8% of genes in the genome in a library of 23,900 sgRNAs. We deployed the optimized library to uncover genes essential to the tolerance of acetate, a promising alternative carbon source, and various hydrocarbons present in many waste streams. Our screens yielded several gene knockouts that improve acetate tolerance on their own and as double knockouts in media containing acetate as the sole carbon source. Analysis of the hydrocarbon screens revealed genes related to fatty acid and alkane metabolism in *Y. lipolytica*. The optimized CRISPR gRNA library and its successful use in *Y. lipolytica* led to the discovery of alternative carbon source-related genes and provides a workflow for creating high-activity, compact genome-wide libraries for strain engineering.

1. Introduction

High throughput functional genomic screens are an essential tool for optimizing yeast production strains (Liu et al., 2015). By screening for the genotypes that underpin traits such as antimicrobial resistance (Kwak et al., 2011), ethanol resistance (Fujita et al., 2006), or morphological changes (Lupish et al., 2022), functional genomic screens can identify and deconvolute coding and regulatory sequences associated with specific cellular functions (Lian et al., 2019). Functional genome-wide screens have been especially effective tools for engineering industrially beneficial traits into yeast (Patterson et al., 2018). Yeasts such as *Saccharomyces cerevisiae* (Chen et al., 2016; Fujita et al., 2006), *Yarrowia lipolytica* (Jagtap et al., 2021; Lupish et al., 2022; Patterson et al., 2018; Ramesh et al., 2023), *Candida albicans* (Adames et al., 2019; Gervais et al., 2021; Wang et al., 2020), and *Komagataella phaffii* (Alva et al., 2021; Nishi et al., 2022; Tkachenko et al., 2023) have all

undergone genome-wide screens to better understand industrially significant metabolic pathways. The outputs have led to promising improvements, such as increased stress tolerance (Crook et al., 2016; Fujita et al., 2006), improved protein production and secretion (Huang et al., 2015; Wang et al., 2019), and alternative nutrient utilization (Xu et al., 2019).

Many recent breakthroughs with functional genomic screens were enabled by CRISPR-editing methods, including CRISPR knockout (Horwitz et al., 2015), knock-in (Stovicek et al., 2015), CRISPRi (targeted gene inhibition (Momen-Roknabadi et al., 2020; Schwartz et al., 2017)), and CRISPRa (targeted gene activation (Schwartz et al., 2018)) screens. These techniques provide broad genomic coverage and high specificity to identify gene targets in various host organism cellular functions. The key component of all CRISPR screens is the guide RNA (gRNA or sgRNA) library (Jakočiūnas et al., 2015), which may target either all of the genes in a host genome (Schwartz et al., 2019), or a

* Corresponding author. Chemical and Environmental Engineering, University of California, Riverside, Riverside, CA, USA.
E-mail address: wheeldon@ucr.edu (I. Wheeldon).

specific subset (Thompson et al., 2021). Once a library is assembled, CRISPR screens can be utilized to identify key genes in metabolic pathways or industrially relevant phenotypes (Jakočiūnas et al., 2015; Thorwall et al., 2020), or find gene expression changes that enhance pathway flux (Da Silva and Srikrishnan, 2012).

Functional CRISPR screens are categorized as either arrayed or pooled screens. Individually prepared mutants may be assembled into discrete or overlapping panels, and analyzed with an arrayed screen approach (Cachera et al., 2024). The outputs from arrayed screens are typically directly measurable phenotypes resulting from each mutant, such as production strain survival or accelerated growth rate (Garcia et al., 2021; Gutmann et al., 2021; Lupish et al., 2022). Arrayed screens are excellent for in-depth, quantitative analysis of specific genes predicted to affect phenotypes of interest (Liu et al., 2022). However, they are labor intensive to execute with larger libraries, as they require isolated strain preparations for each mutant to be tested (de Groot et al., 2018).

Alternatively, large, whole genome mutant libraries may be analyzed through a pooled screen (Bowman et al., 2020; Trivedi et al., 2023). Highly efficient transformation methods are used to generate a representative population of library cells, generating all possible single knockouts (or gene inhibitions/activations) of the given organism within a single culture volume (Schwartz et al., 2019). The pooled mutant library cells may then undergo phenotypic divergence when grown under a selection pressure (Smith et al., 2016). While individual genotypes cannot be directly observed or isolated in the pooled culture, deep sequencing can be used to find sgRNA prevalences and by extension, essential genes or genes involved in a phenotype of interest (Ramesh et al., 2023). Pooled genome-wide screens are an ideal high-throughput method for studying unknown relationships between genes, how those relationships affect desired phenotypes, and how those interactions affect metabolite production (Li et al., 2020). Still, challenges remain in the design and utilization of genome-wide sgRNA libraries. They must have maximized coverage of the host genome, ideally with highly efficient guide activities. Guide RNA expression characteristics must also be optimized to ensure sufficient genome coverage (Dalvie et al., 2020). To overcome these challenges, techniques are required to better optimize CRISPR gRNA libraries.

In this work, we design an optimized sgRNA library for CRISPR genome-wide pooled screens in *Y. lipolytica*. We then validated our library for use in growth-based functional genomic screens to improve growth on non-glucose carbon sources, including acetate, fatty acids, and hydrocarbons. To optimize our library, we strategically reduced the size of a previously established *Y. lipolytica* sgRNA knockout library by half, while maintaining coverage of the host genome. In addition to existing experimentally validated sgRNAs, we also incorporated new sgRNAs predicted by DeepGuide (Baisya et al., 2022). To identify phenotypes of interest, we screened *Y. lipolytica* harboring both the sgRNA library and integrated Cas9 on a variety of carbon sources. From our hydrocarbon screens, we found significant overlap in genes essential for growth between similar hydrocarbons. We experimentally validated our acetate hits and found that knockouts of our top hits shortened lag phase and increased growth with acetate as the sole carbon source. This *Y. lipolytica* optimized CRISPR gRNA library provides a workflow for designing efficient libraries for functional genomic screening of metabolic engineering traits.

2. Results and discussion

2.1. Compact CRISPR-Cas9 library contains high-activity sgRNAs

Iterating on our previously created 6-fold coverage library, we sought to create a minimally sized, high activity “v2” library. The high activity gives more confidence in the data and the smaller size reduces the transformation burden to achieve high coverage libraries. We designed a 3-fold coverage CRISPR-Cas9 library to target all protein-

coding sequences in the PO1f strain of *Y. lipolytica*. The library consists of a combination of high-activity sgRNAs from an existing *Y. lipolytica* CRISPR-Cas9 library designed by (Schwartz et al., 2019) (denoted as “v1”) and sgRNAs predicted to be of high activity by an activity prediction tool, DeepGuide (Baisya et al., 2022) (Fig. 1a and b).

Growth screens previously conducted in the *Y. lipolytica* PO1f strain using library v1 defined a metric called cutting score (CS) for each sgRNA — a measure of its cutting efficiency — by comparing sgRNA abundance in the PO1f Cas9 Δ KU70 strain to that in the control strain devoid of Cas9 (Schwartz et al., 2019). Guides that make a cut in the experimental strain become depleted as cells die since disruption of KU70 suppresses non-homologous end joining, the dominant DNA repair mechanism in *Y. lipolytica* (Schwartz et al., 2017), whereas guides that do not make a cut become enriched. By comparing the abundance of each guide in the experimental strain to the abundance of the same guide in the control strain, CS can be calculated. The sgRNA sequences from library v1 along with MNase-seq-derived nucleosome occupancy data for *Y. lipolytica* PO1f were used as input data for DeepGuide sgRNA predictions. Nucleosome occupancy was included in the DeepGuide computation because high nucleosome occupancy can inhibit Cas9 cutting activity by physically blocking the Cas9-sgRNA complex from binding to DNA (Baisya et al., 2022; Schwartz et al., 2019). In order to pick sgRNAs for library v2, library v1 was first filtered to only retain sgRNAs having a CS greater than 4.0 (a value close to the optimum CS threshold for high-activity sgRNAs previously determined by (Ramesh et al., 2023)). Library v2 was thus designed to contain the top two guides with $CS > 4.0$ for each gene from the filtered library v1, and a guide with the highest DeepGuide-predicted CS for that gene. If, for a gene, less than 2 sgRNAs from library v1 had $CS > 4.0$, appropriate numbers of sgRNAs with high DeepGuide-predicted CS were included so as to maintain a library coverage of 3 (see Materials and Methods). The final library consists of three sgRNAs per gene targeting 99.8% of all genes (Fig. S1a). In addition to using design criteria such as DeepGuide predicted CS to pick guides for the library, we ensured that every sgRNA in the final library is unique in the genome, does not target intronic regions, and is spaced apart from other sgRNAs targeting the same gene to maximize diversity of the generated knockouts, further details can be seen in (Ramesh and Wheeldon, 2021; Trivedi et al., 2023). Moreover, next generation sequencing of the cloned, untransformed library revealed a tight normal distribution of sgRNA abundances, indicating a near-uniform and symmetrical (median \approx mean) distribution of guides in the library (Fig. S1b). sgRNA cutting scores for library v2 were generated using the same screening framework as that used for obtaining CS of library v1 guides. The CS for all guides in the v2 library are provided in **Supplementary File 1**.

In addition to gene-targeting guides, library v2 also consists of 360 non-targeting sgRNAs (\sim 1.5% of the total library) that serve as negative controls. The normalized CS (i.e., the difference between raw CS of a guide and the average CS of non-targeting guides) distributions of targeting and non-targeting sgRNAs nearly overlap on day 2 (avg. $CS_{norm, targeting} = 0.12 \pm 0.09$ and avg. $CS_{norm, non-targeting} = 0 \pm 0.06$) due to no observable effect of CRISPR-induced double stranded breaks on cell growth. On day 4, however, the two populations are well separated (avg. $CS_{norm, targeting} = 5.96 \pm 2.04$ and avg. $CS_{norm, non-targeting} = 0 \pm 0.34$), indicating different activity profiles of targeting and non-targeting guides (Fig. 1c). The difference between targeting and non-targeting populations is statistically higher for library v2 (one-tailed unpaired t-test $p < 0.0001$, t-score = 58.81) compared to v1 (one-tailed unpaired t-test $p < 0.0001$, t-score = 43.49), providing evidence that library v2 is a higher-activity library compared to library v1 (Fig. 1d). Also of note is the genomic loci of the experimental and DeepGuide-predicted sgRNAs used to construct library v2, which are evenly distributed across the genome (Fig. 1e). These data show that we built a compact, high activity library that can be used in functional genomic screens in *Y. lipolytica*. The small size enables high library coverage due to lower transformation burden and the high activity allows near complete assurance of genomic

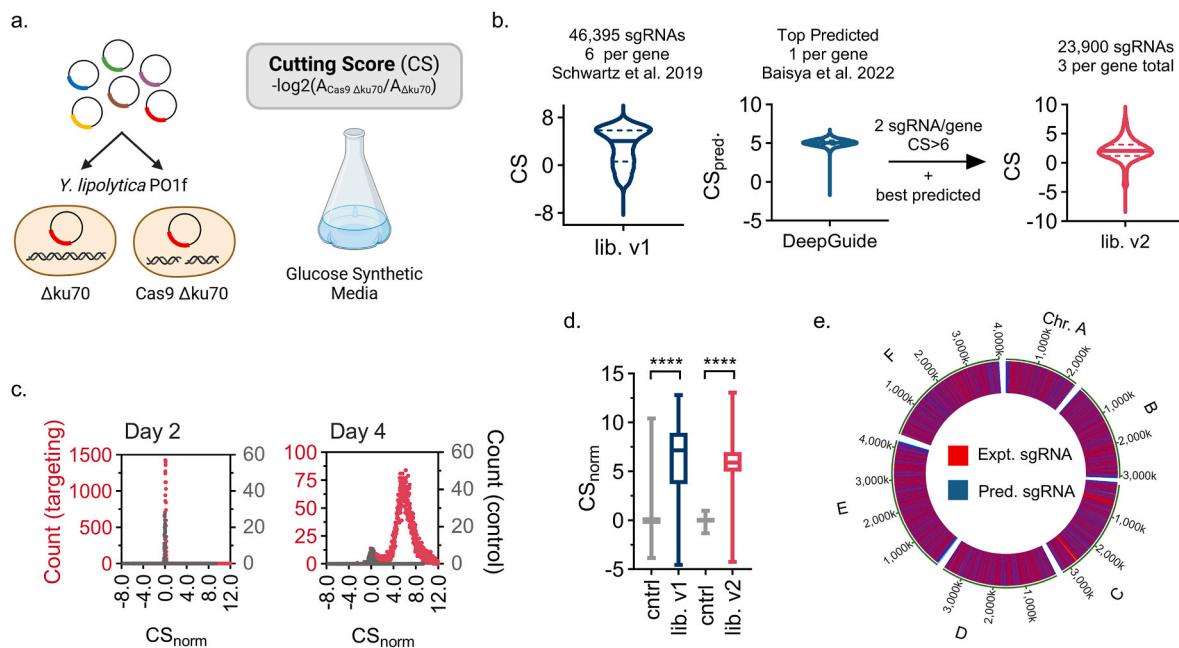


Fig. 1. Optimized CRISPR library generation pipeline and characterization. (a) Cutting score (CS) is calculated as the depletion of a guide in the Cas9 Δku70 knockout strain compared to the Δku70 control. A guide which cuts well in the Cas9 Δku70 knockout strain will become depleted as cells die or have a serious growth penalty. (b) The two best cutting guides from lib. v1 were used in addition with the top predicted guide from DeepGuide to generate lib. v2. (c) Distributions of targeting (red) and non-targeting (gray) sgRNAs after two and four days of growth post transformation. After four days of growth with one subculture on day 2, the high activity sgRNAs diverge from the non-targeting population. (d) Both lib. v1 and v2 contain active guides, however, v2 contains more guides with high activity. (e) DeepGuide predicted and lib. v1 guides are distributed randomly throughout the genome.

coverage.

2.2. Acetate tolerance screens identify gene targets for improved growth

With the optimized genome-wide CRISPR library in hand, we next set out to identify genes essential for growth on acetate. Functional genomic screens were conducted using minimal media supplemented with three

different carbon sources — glucose (0.125 M) as a control condition, and low and high acetate (0.25 and 0.5 M, respectively) as experimental conditions (Fig. 2a). These screens allowed us to calculate a fitness score (FS) for every sgRNA in each media condition. A high FS value indicates that knockout of that gene gives a fitness advantage, provided a cut happens. Conversely, low FS values indicate a loss of cell fitness due to disruption of the target gene. FS is computed as the \log_2 -ratio of sgRNA

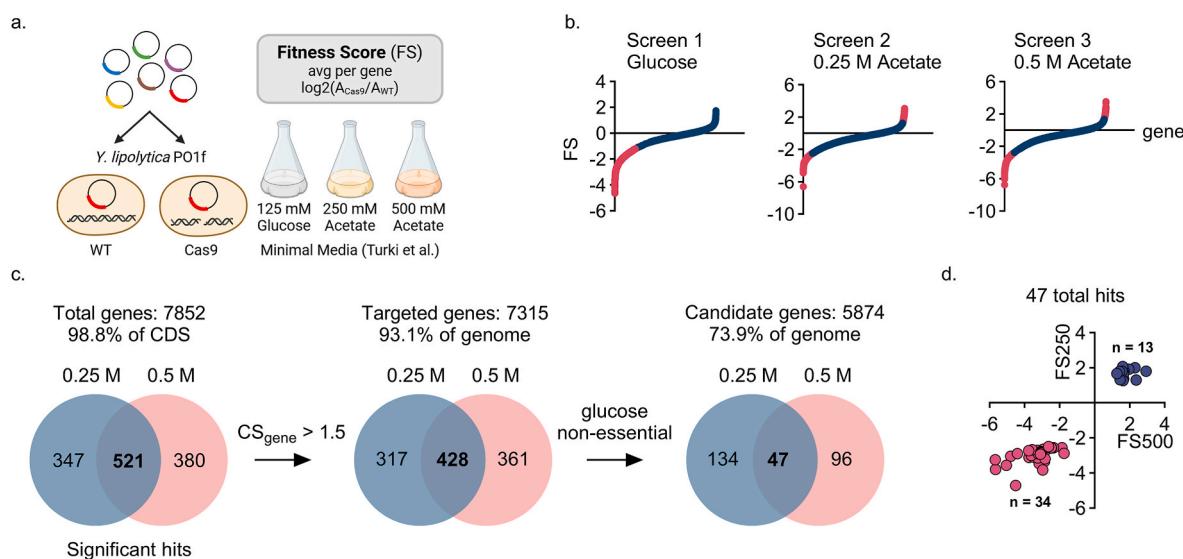


Fig. 2. Growth-based functional genomic CRISPR screening with acetate as the sole carbon source. (a) Growth screens were conducted using *Y. lipolytica* PO1f and a strain containing Cas9 as control and treatment strains respectively in three media conditions - 0.125 M glucose, 0.25 M acetate, and 0.5 M acetate. Fitness score (FS) is calculated as the average enrichment of guides targeting a gene in the Cas9 strain compared to PO1f. Genes with a higher FS indicate that the knockout gave a fitness advantage. (b) Corrected FS values for all genes for the glucose and acetate screens, as determined from the acCRISPR pipeline. Pink portion of the curve represents essential genes for glucose, and high and low FS genes for the two acetate conditions. (c) Venn diagrams representing gene hits in low and high acetate conditions after each filtering step. Low-activity guides and glucose essential genes were removed to condense the pool of hits. (d) Top low and high FS hits in both acetate conditions.

abundance in the PO1f Cas9 strain to that in the PO1f control strain. We used acCRISPR, an analysis pipeline for CRISPR screens that correct screening outcomes based on the activity of each guide used in the screen (Ramesh et al., 2023) to identify essential genes in glucose, and high and low FS value genes in the two acetate conditions. The analysis pipeline identified 1580 essential genes for growth with glucose as the sole carbon source, whereas our v1 library identified 1903; 1214 genes overlap between these essential gene calls (Schwartz et al., 2019). We found 868 and 901 genes that reduce or improve growth on acetate, respectively (Fig. 2b).

Despite lib. v2 containing mostly high activity cutters, we wanted to be sure that very few false positive hits would arise at the end of our screening pipeline. To do this, we used guides with high activity (CS > 1.5) for calling hits (Fig. 2c). This reduced the size of the common significant gene set between the two acetate conditions to 428 (13 with high FS values and 415 with low FS values). Any double stranded break that occurs in a cell causes at least a moderate fitness effect (Schwartz et al., 2019). Additionally, there are likely very few gene knockouts that improve tolerance to acetate in *Y. lipolytica*. These reasons may explain why we demonstrated an 86% loss of positive FS hits (13 with high FS down from 94) when we removed poor targeting guides. These false positives represent control strain cells that did not have a fitness disadvantage of a double stranded break. Next, we wanted to exclude any significant genes that reduce fitness in glucose media. We reasoned that the most relevant hits should be solely due to acetate fitness effects and not to overall fitness. Glucose viability is also important in upstream applications for making further strain modifications and for downstream applications if glucose supplementation is needed for an increase in intracellular reducing power. As a result, we excluded genes that were found to be essential in glucose from the significant gene lists for acetate, further condensing the common significant gene set to 47 genes. Thirteen of the 47 genes had positive FS values while the remaining 34 genes had FS values lower than 0 (Fig. 2d).

2.3. Acetate screen hits shorten lag phase growth

To validate the screening hits, we focused on the top 13 positive FS hits identified in the low and high acetate screening conditions, as knocking out these genes should improve growth on acetate (Fig. 3a). To do this, we transformed PO1f with a plasmid containing Cas9 and the

relevant guide. We obtained 12 of the 13 top positive FS knockouts; functional gene disruptions were confirmed with colony PCR and sanger sequencing. Our top hit, D05956g, encodes SFL1, a flocculation suppressant gene. When knocked out, this gene causes *Y. lipolytica* to flocculate (Fig. S2). This mutation could provide a fitness advantage in the context of our screen as it allows cells to clump together at the bottom of the flask and avoid the high osmotic stress of being a suspended single cell. The flocculation behavior makes growth measurements challenging and thus the D05956g knockout was excluded from further analysis.

For the remaining 11 knockouts, we validated their phenotype in growth assays that mimic the screening conditions, growth in minimal media with 0.5 M acetate as the sole carbon source. Under these growth conditions we observed that at 42 h *Y. lipolytica* PO1f begins to break out of the lag phase and enter the exponential growth phase. This time point gave us the most resolution and repeatability to confirm the hits. Of the 11 characterized knockouts, 8 grew significantly better than the control ($p < 0.05$), including E01193g, D21022g, E37234g, C33332g, E24428g, C02904g, F23331g, and E22029g (Fig. 3b). To validate that there was little to no effect in glucose media, we repeated the same experiment in minimal media with 0.125 M glucose. Only knockout of C30884g grew significantly worse than wild-type (Fig. S3) suggesting that nearly all knockouts affect acetate tolerance or metabolism and not overall cell viability.

We next sought to explore if higher order mutations would further improve *Y. lipolytica* growth in acetate media. To test the effect of stacked mutations, we chose the top 6 knockouts (E01193g, D21022g, E37234g, C33332g, E24428g, and C02904g) and tested all 15 double knockout permutations. Only two double knockouts grew better than our top single knockout, possibly due to a combined reduction in general cell fitness in the higher order mutants (Fig. 3b).

We followed up the preliminary hit validation with more comprehensive characterization of cell growth and quantification of acetate depletion. The top two single and double knockouts were selected as well as C02904g due to the commonality between the top two double knockout hits (Fig. 3c). Limited overall growth and incomplete acetate consumption may be explained by micronutrient or amino acid depletion or pH change caused by consumption of acetate while leaving behind sodium ions. Roughly 25% of acetate was converted to dry cell weight (DCW). All hits have a shorter lag phase (reduced by 24 h in the most potent knockout) and consume more acetate than the PO1f control

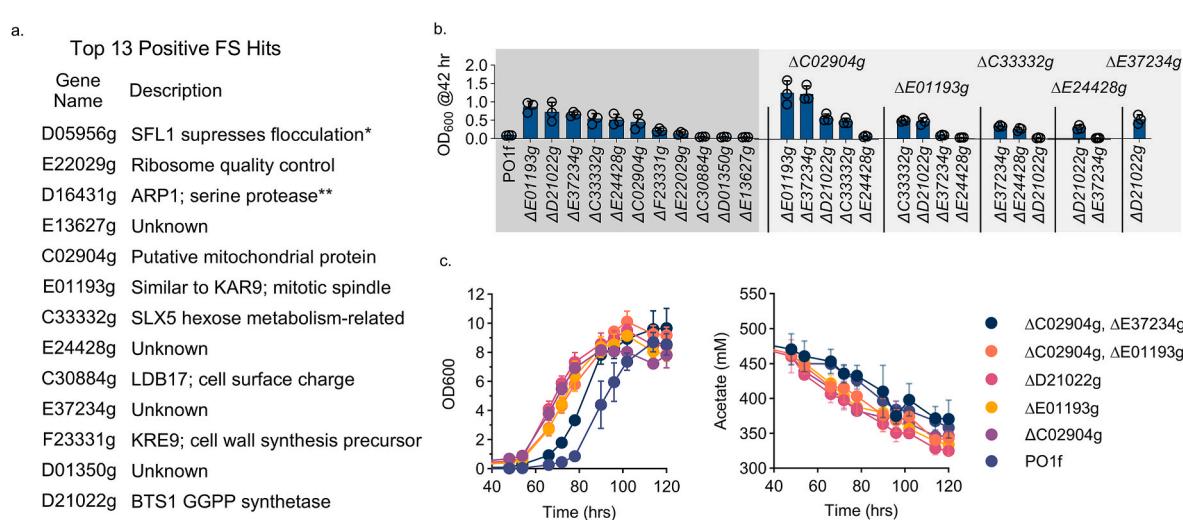


Fig. 3. Validating acetate metabolism hits. (a) Thirteen positive FS value hits in descending order of high acetate FS. Asterisk (*) indicates knockout was not included due to flocculation, double asterisk (**) indicates knockout was not achieved. (b) Single (left dark gray box) and double knockout (5 light gray boxes on the right, common knockout listed above) mutants tested in 96-well 1 mL wells containing 0.5 M acetate in minimal media (30 °C, 42 h, 1000 RPM). Bars represent the mean of three biological replicates, error bars represent the standard deviation. Data points are shown for each replicate. (c) Growth curve and acetate depletion for top three single knockouts and top two double knockouts grown in 250 mL shake flasks with 0.5 M acetate in minimal media (30 °C, 220 RPM). Acetate depletion measured with HPLC. Data points represent the mean of three biological replicates, error bars represent the standard deviation.

strain (Fig. 3c) and Δ C02904g, Δ E01193g, Δ E37234g, and Δ C02904g/ Δ E01193g strains had significantly higher growth rates (Fig. S4). These data provide strong validation of our top positive FS hits as mutations that lead to the improvement of acetate metabolism in *Y. lipolytica*.

Our top four hits to improve growth were E37234g (unknown function), E01193g (KAR9, which plays a role in mitotic spindle positioning in *S. cerevisiae* (Tirnauer and Bierer, 2000)), C02904g (FMP42, putative mitochondrial protein in *S. cerevisiae* (Imamura et al., 2015)), and D21022g (GGPP synthetase homolog (Saikia and Scott, 2009)). GGPP synthetase converts farnesyl pyrophosphate (FPP), the precursor to steroids and N-glycans, to GGPP, a precursor to terpenoids and other metabolites (Liu et al., 2014). Previous work has shown that deletion of this gene improves oxidative stress response in *N. crassa* (Sun et al., 2019). One possibility is that knockout of GGPP synthetase increases intracellular FPP and decreases GGPP, which may reduce oxidative stress from endogenous terpenoid production and increase cell wall integrity by improved flux to membrane essential compounds through FPP.

2.4. Hydrocarbon screening reveals genes necessary for fatty acid metabolism and transport

Given the success of the genome-wide CRISPR approach in identifying genes that improve growth on acetate, we next sought to uncover genes related to hydrocarbon and fatty acid metabolism, substrates on which *Y. lipolytica* has a demonstrated ability to grow (Fickers et al., 2005; Try et al., 2018). Dodecane, 1-dodecene, oleic acid, and margaric acid were chosen as substrates for their diversity in polarity and hydrogenation. Preliminary experiments showed that solid media produced more consistent cell growth than liquid cultures due to uneven mixing of hydrocarbons in the liquid cultures and the fact that margaric acid is solid at 30 °C. We therefore chose to use solid media for

hydrocarbon and fatty acid functional genomic screening. Glucose as a sole carbon source was also screened to control for any putative hits that were solely due to the transition to solid media (Fig. 4a).

Hydrocarbon screen hits were identified in a similar manner to the acetate screen: the acCRISPR analysis pipeline was used to determine FS values from next generation sequencing read counts of the sgRNAs remaining at the end of the screen; genes with fitness advantage or essentiality were called as hits; hits from genes with low CS guides were removed from the hit pool to avoid false positives; and, genes found to be essential in the glucose liquid culture condition were removed from the hit list. In total, 478 high confidence hits were called across the four substrates tested, with 28 gene knockouts with a fitness advantage, and 449 genes essential for growth on the substrates (Fig. 4b).

Upset plot analysis is a visual means of identifying hits that are common across various conditions (Fig. 4c and d). When applied to our screens, this analysis revealed that three hits with positive FS values were called in both the dodecane and dodecene screens, each of which is involved with membrane transport of lipids or carbohydrates (Fig. 4c). The genes common to the dodecane and dodecene screens were a mixture of stress response, endomembrane trafficking, and pro-hyphal membrane and cell wall maintenance genes. There were no positive FS hits for oleic acid and only one for margaric acid (C06486g, a knockout of a transcription factor with homology to *S. pombe* ADN1, believed to promote invasive growth and cell adhesion), which was also found in the dodecane screen. In addition to these hits, ten positive FS hits were identified for dodecane only and eight for 1-dodecene only (see Supp. File 2 for a list of all genes in the upset plots). Additional analysis of screen hits with positive FS is shown in Fig. 5. With respect to the essential genes (i.e., those with low FS values), the overlap between carbon sources was more comprehensive (Fig. 4d). Eleven genes were found to be common across all screens, another 30 were common to three of four screens and 93 were common to two of four screens.

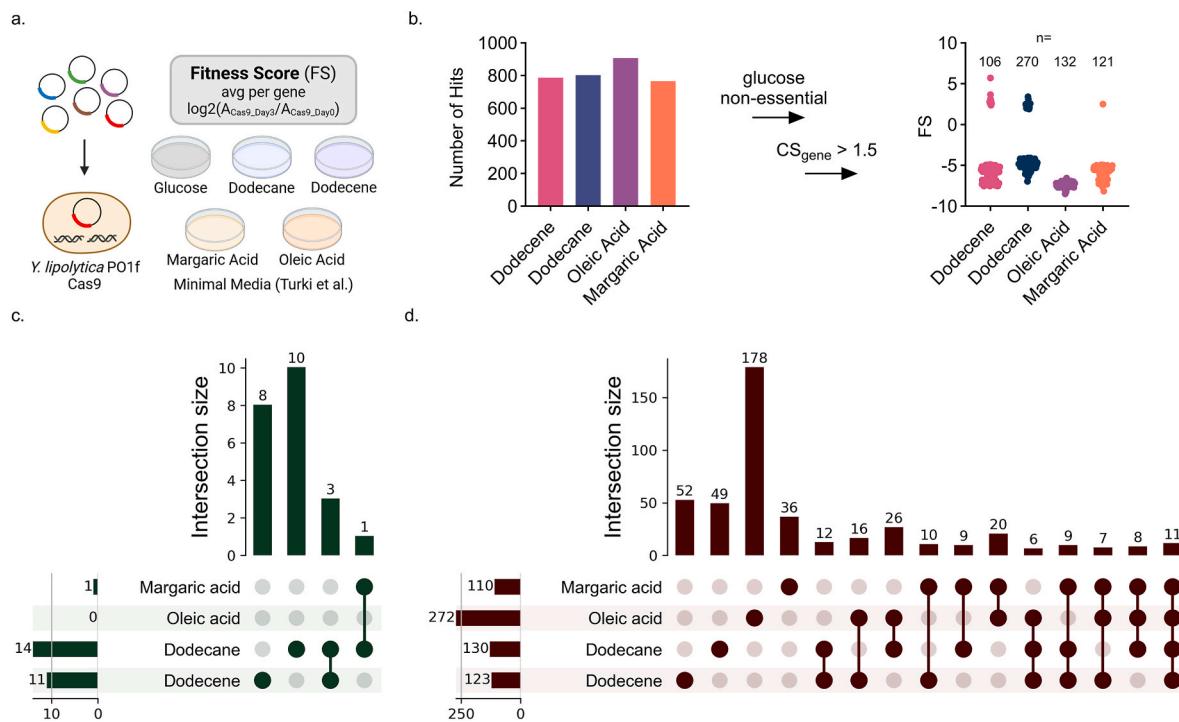
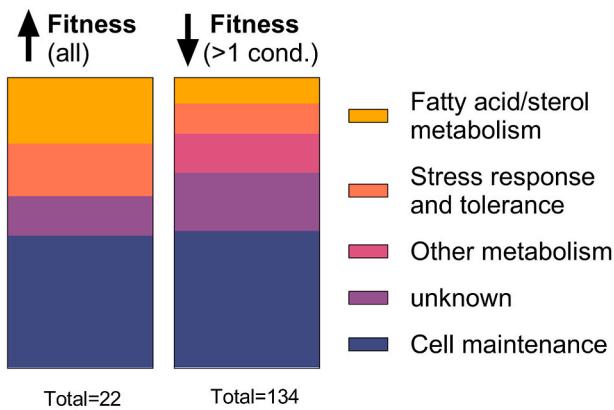


Fig. 4. Growth-based functional genomic CRISPR screening with hydrocarbon substrates. (a) Growth screen, carbon sources screened and fitness score (FS) calculation. Genes with a higher FS indicate that the knockout gave a fitness advantage. (b) Filtering of gene hits obtained for each hydrocarbon condition; gene hits with poor cutting guides and genes essential for growth on glucose were removed to obtain high confidence hits. (c,d) UpSet plot analysis for hits in all hydrocarbon conditions. Hits with a fitness advantage are shown in (c), while essential hits are shown in (d). The values associated with each substrate indicate the number of gene hits that were identified for each substrate. Screen growth conditions were 55.5 mM dodecane solid media grown for 48 h, 55.5 mM 1-dodecene solid media grown for 48 h, 37 mM oleic acid solid media grown for 48 h, and 39.2 mM margaric acid solid media grown for 72 h, all at 30 °C in a static incubator, with two screen replicates of each condition, with each replicate consisting of seven 100 mm plates.

a.



b.

Gene	Homology	Description
E36308g	<i>S.c. HOB1/2</i>	Lipid transport protein
E03584g	<i>A.a. PDI</i>	Triglyceride transfer
A18344g	<i>S.c. ERG5</i>	C22 sterol desaturase
D20202g	<i>S.c. UBP15</i>	Growth defects on oleic acid
C03415g	<i>E.n. n/a</i>	Induced by fatty acids
E32776g	<i>S.c. RSB1</i>	Sphingoid metabolism
D26639g	<i>S.p. PDF1</i>	Fatty acyl metabolism
A20089g	<i>S.c. HTD2</i>	Fatty acid biosynthesis

Fig. 5. Biological function of top hydrocarbon screening hits. (a) Hit distribution by biological function: Fatty acid/sterol metabolism, stress response and tolerance, cell maintenance, and unknown. The left bar includes all positive FS gene hits in all hydrocarbon screens ($n = 22$). The right bar includes all essential gene hits called in two or more screens ($n = 134$). (b) Examples of fatty acid metabolism-related gene hits that improve fitness in at least one hydrocarbon condition and essential genes in at least two hydrocarbon conditions.

Functional analysis of these genes, as well as hits with a fitness advantage, are presented in Fig. 5.

Our screen revealed 134 total genes that reduce fitness in at least two hydrocarbon/fatty acid conditions (Fig. 5a). Twelve hits that reduced fitness were directly related to hydrocarbon metabolism. One of these twelve was C03415, a hit identified in all four hydrocarbon screens and known to be essential for fatty acid metabolism in *Aspergillus nidulans* (*acuH* homolog) (De Lucas et al., 1999). Another 18 hits were related to metabolic pathways, including endomembrane trafficking genes. These include E23880g, an essential gene for all four conditions which is homologous to *S.cerevisiae* VPS29, a membrane-bound vesicle reverse-transport protein (Seaman et al., 1998)), and D14097g that was found to be essential in three of the four screens. D14097g is homologous to *S.cerevisiae* VPS53, which is responsible for retrograde sorting in the golgi apparatus (Conibear and Stevens, 2000). We expected hits such as these two, as there is experimental evidence that these genes are involved with endomembrane trafficking that affect alkane metabolism (Fukuda, 2013). Other hits include the previously mentioned *ADN1* homolog and E07121g found in our dodecane hits (homologous to *PMT4* in *C. albicans* which plays a role in invasive hyphal growth (Lengeler et al., 2008)). We expected a fitness advantage from deletions of pro-hyphal genes like these since there is a known inverse relationship between hyphal morphology and efficient hydrocarbon metabolism (Palande et al., 2014). In addition to the hydrocarbon and metabolic pathway hits, there were 14 hits related to stress response/tolerance genes, a class of genes that are expected to be essential for growth under these conditions. There were 63 hits related to cell maintenance

functions, which we anticipated as they are essential for maintaining cell viability when hydrocarbons are the only available carbon source. Finally, we provided further validation that knockout of four predicted negative FS hits in 1-dodecene all reduce cellular fitness on the carbon source: C03415g (*ACUH* homolog) is induced by fatty acids in yeast (De Lucas et al., 1999), E32776g (*RSB1* homolog) is related to sphingoid metabolism in *S. cerevisiae* (Kihara and Igarashi, 2002), D26639g (*PDF1* homolog) in *Schizosaccharomyces pombe* plays a role in fatty acyl metabolism (Cho and Hofmann, 2004), and A20089g (*HTD2* homolog) is involved in fatty acid metabolism in *S. cerevisiae* (Kastaniotis et al., 2004) (Fig. S5). Taken together, these hits reveal metabolic pathways and cellular functions that are essential for survival in hydrocarbons and fatty acids.

In addition to essential genes, we identified 22 genes that produced positive FS values in the hydrocarbon and fatty acid screens and are beneficial to growth on one or more of the hydrocarbon substrates. Similar to the essential gene hits, roughly half of these hits were related to cell maintenance. This group includes E30725g, identified in the 1-dodecane screen and a homolog of *MNN10* in *S. cerevisiae*, and is related to cell membrane/wall maintenance (Jungmann et al., 1999; Klein et al., 2002; Sacristán-Reviriego et al., 2014). C06486g, an *ADN1* homolog in *S. pombe*, was found in both the margaric acid and dodecane screens. This gene is involved in cell adhesion and hyphal regulation, a trait known to affect hydrocarbon metabolism (Palande et al., 2014). Five genes related to stress response were also found to improve growth on hydrocarbons. Specifically, C14369g, which is homologous to various *Candida* drug resistance exporters in *Candida species* and *S. cerevisiae* (Costa et al., 2013; Dong et al., 2021)) and F01346g a homolog of *MSG5* in *S. cerevisiae* involved in cell signaling (Novodvorska et al., 2016; Sacristán-Reviriego et al., 2014)) were found in the dodecane and 1-dodecene screens, respectively. We uncovered four hits directly related to improved tolerance to hydrocarbons (Fig. 5b). The *HOB2* (E36308g) homolog in *S. cerevisiae* was present in both dodecane and 1-dodecene and is a putative direct membrane contact lipid transporter that moves lipids between different membranes and organelles across the cell, including the ER, peroxisomes, and vacuoles (Castro et al., 2022; Leonzino et al., 2021). As such, it is likely to interface with the membranes of organelles involved with the catabolic assimilation, functionalization, and ultimate β -oxidation of the alka(e)ne, controlling the destination of the lipid (Dabrowski et al., 2023; Fukuda, 2023). *sdATG2* also plays a role in excess peroxisome degradation (Dunn et al., 2005) which may not be beneficial when hydrocarbons are the sole carbon source. The remaining three hits were only found in the 1-dodecene growth condition. They included E03584g, a homolog to an *S. cerevisiae* disulfide isomerase associated with the ER lumen (Gauss et al., 2011). It may also be a component of a triglyceride transfer complex, providing catalytic activity required for neutral lipid transport (Wetterau et al., 1990). Another 1-dodecene hit, A18344g, is homologous to *S. cerevisiae* *ERG5*, a C-22 sterol desaturase localized to the ER and cytochrome P450 enzyme (Skaggs et al., 1996). Similar enzymes contribute to the functionalization of hydrocarbons into fatty alcohols (Fukuda and Ohta, 2013). As this is the first step of hydrocarbon catabolism, A18344g could play a putative role in alkene catabolism. A third 1-dodecene hit, D20202g, is homologous to *S. cerevisiae* *UBP15* homolog found to cause growth defects on oleic acid when deleted (Debely et al., 2011). It helps regulate transport from the ER to the peroxisome and its disruption causes peroxisomal clustering (Debely et al., 2011). Its role putatively places it directly in the hydrocarbon catabolism pathway, in which functionalized hydrocarbons are transferred to the peroxisome for β -oxidation (Fukuda, 2023). Finally, we performed knockouts of all the hits that should improve fitness shown in Fig. 5b. Three of four knockout strains grew larger colonies than wild-type (D20202g, A18344g, and E36303g) on 1-dodecene as spot plates (Fig. S6). In both fitness improvement and essential gene hits, our screen outputs demonstrate how genome-wide CRISPR knockout screens are a vital component of efficient strain engineering.

3. Conclusions

Genome-wide CRISPR screening in industrial hosts like *Y. lipolytica* enables rapid engineering of growth-based phenotypes. We previously developed a genome-wide CRISPR-Cas9 knockout library for use in *Yarrowia* that can be used in functional genetic screens and in rapid design-build-test-learn cycles for strain engineering. The first version of this screening tool contained sgRNAs with a wide range of activity and needed 6-fold coverage to cover nearly all the genes in the genome. Here, we create a new optimized, high activity sgRNA library that is compact in size. This library required only 3-fold genome coverage to target 98.8% of genes in the genome. By creating a library one-half the size, we reduce the burden of generating large libraries that require high transformation efficiency to obtain. Quantifying the activity of each guide in the library revealed that all but a handful of guides are highly active. A tight, normal distribution of guide activity enables high accuracy screening as all (or nearly all) guides can be used to identify gene hits. This optimized CRISPR-Cas9 screening tool enabled us to conduct high throughput strain engineering experiments for growth on various alternative carbon sources, including acetate, dodecane, 1-dodecene, oleic acid, and margaric acid. The acetate screens identified several single and double mutant strains with shortened culture times when consuming acetate as the sole carbon source. The hydrocarbon screens revealed 22 gene targets for improving growth and a set of 134 genes essential for fatty acids and long chain hydrocarbon metabolism. Many of these hits have unknown function or are gene targets for metabolic engineering. Our compact, high activity CRISPR-Cas9 library enables growth-based screens for enhancing carbon source utilization and promises to enable a wide range of screens for improving other growth-based phenotypes.

4. Materials and Methods

4.1. Microbial strains and culturing

All strains used in this work are presented in Table S1. *Yarrowia lipolytica* PO1f (MatA, leu2-270, ura3-302, xpr2-322, axp-2) is the parent for all mutants used in this work. Unless otherwise noted, all yeast culture growth was carried out in 14 mL polypropylene tubes or 250 mL baffled flasks, with incubator conditions of 30 °C and 220 RPM. Under non-selective conditions, *Y. lipolytica* was grown in YPD (1% Bacto yeast extract, 2% Bacto peptone, 2% glucose). Cells transformed with sgRNA-expressing plasmids were initially propagated in synthetic defined media deficient in leucine (SD-leu; 0.67% Difco yeast nitrogen base without amino acids, 0.069% CSM-leu (Sunrise Science, San Diego, CA), and 2% glucose) for two days to allow for genome edits to occur. All plasmid constructions and propagations were conducted in *Escherichia coli* TOP10. *E. coli* cultures grown in Luria-Bertani (LB) broth with 100 mg/L ampicillin at 37 °C in 14 mL polypropylene tubes, at 220 RPM. Plasmids were isolated from *E. coli* cultures using the Zymo Research Plasmid Miniprep Kit II.

4.2. Plasmid construction

All plasmids and primers used in this work are listed in Tables S2 and 3. The plasmids used to knock out genes in *Y. lipolytica* PO1f were constructed by ordering the corresponding sgRNA as a primer (Table S4) with 20 bp homology up- and downstream of the *AvrII* cutsite in the pSC012 plasmid. 60 bp top and bottom strands were ordered and annealed together. The annealed strand and digested plasmid were assembled using Gibson Assembly in a 10:1 M ratio (insert:vector). The assembly product was then transformed directly into electrocompetent *E. coli* TOP10 cells to eventually be propagated and harvested by Miniprep.

4.3. *Y. lipolytica* CRISPR knockouts

A plasmid containing Cas9 and the appropriate sgRNA (pSC012) for the desired gene knockout was transformed into *Y. lipolytica* using a protocol described by Chen et al. (1997). In short, a single colony of the background strain of interest was grown in 2 mL of YPD liquid culture in a 14 mL culture tube at 30 °C with shaking at 220 RPM for 22–24 h (final OD ~30). 300 µL of culture (~10⁸ cells) were pelleted by centrifugation at 4,000g for 2 min and then resuspended in 300 µL of transformation buffer. The transformation buffer contains a final concentration of 45% PEG 4000, 0.1M Lithium Acetate and 100 mM Dithiothreitol. Then, 500 ng of plasmid DNA was added followed by 8 µL 10 mg/L ssDNA (Agilent). The reaction mix was vortexed thoroughly and then incubated for 1 h at 39 °C. 1 mL of water was added and then the cells were pelleted and inoculated into 2 mL SD-ura selective liquid media. After 3 days, the cells were plated at a 10⁻⁶ dilution on YPD. After one day of incubation at 30 °C colony PCR and then sanger sequencing was performed to find frameshift mutation gene knockdowns.

4.4. sgRNA library design

Custom MATLAB scripts were used to design the optimized Cas9 library, and the key elements of the design are reported here. The optimized library had 3 guides designed for all 7919 mRNA coding genes in the *Y. lipolytica* CLIB89 genome (https://www.ncbi.nlm.nih.gov/assembly/GCA_001761485.1) (Magnan et al., 2016). Of these 3 guides, 2 were intended to be picked from the pool of best performing guides in the previous Cas9 screen (Schwartz et al., 2019), while the third guide was designed by DeepGuide predictions (Baisya et al., 2022; Schwartz et al., 2019). First, sgRNAs with CS > 4.0 from the Schwartz et al. Cas9 screen were sorted from highest to lowest and the best two sgRNAs for each gene were identified. The third sgRNA for all genes, as well as guides for any genes that did not have two highly active guides (CS > 4.0), were obtained from DeepGuide's best predictions for that gene. All sgRNAs in the optimized library were verified to contain a unique seed sequence (11 nucleotides closest to the PAM). 360 nontargeting sgRNAs were also included in the library. These guides were confirmed not to target anywhere within the genome by ensuring that the first 12 nucleotides of the sgRNA did not map to any genomic loci.

4.5. sgRNA library cloning

The Cas9 library targeting the protein-coding genes in PO1f was ordered as an oligonucleotide pool from Agilent Technologies Inc. and cloned in-house using the Agilent SureVector CRISPR (Part Number G7556A). The library was subject to a NextSeq run to test for fold coverage of individual sgRNAs and skew. Cloned DNA was transformed into NEB 10-beta *E. coli* and plated. Sufficient electroporations were performed for each library to yield >100X library coverage. The plasmid library was isolated from the transformed cells after a short outgrowth. The optimized Cas9 library was cloned by making use of the Agilent SureVector CRISPR Library Cloning Kit (Part Number G7556A). Briefly, the backbone pCas9yl-GW was linearized and amplified by PCR using the primers InversePCRCas9Opt-F and InversePCRCas9Opt-R. To verify the completely linearized vector, we *Dpn*I digested amplicon, purified the product with Beckman AMPure XP SPRI beads, and transformed it into *E. coli* TOP10 cells. A lack of colonies indicated a lack of contamination from the intact backbone. Library ssDNA oligos were then amplified by PCR using the primers OLS-F and OLS-R for 15 cycles as per vendor instructions using Q5 high fidelity polymerase. The amplicons were cleaned using the AMPure XP beads prior to use in the following step. sgRNA library cloning was conducted in four replicate tubes and subsequently, pooled and cleaned up as per manufacturer's instructions.

One amplification bottle containing 1L of LB media and 3 g of high-grade low-gelling agarose was prepared, autoclaved, and cooled to 37 °C (Agilent, Catalog #5190–9527). Ten transformations of the cloned

library were conducted using Agilent's ElectroTen-Blue cells (Catalog #200159) via electroporation (0.2 cm cuvette, 2.5 kV, 1 pulse). Cells were recovered and with a 1 h outgrowth in SOC media at 37 °C (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, and 20 mM glucose.) The transformed *E. coli* cells were then inoculated into the amplification bottle and grown for two days until colonies were visible in the matrix. Colonies were recovered by centrifugation and subject to a second amplification step by inoculating two 250 mL LB cultures. After 4 h, the cells were collected, and the pooled plasmid library was isolated using the ZymoPURE II Plasmid Gigaprep Kit (Catalog #D4202) yielding ~1.8 mg of plasmid DNA encoding the optimized Cas9 sgRNA library. The library was subject to a NextSeq run to test for fold coverage of individual sgRNAs and skew.

4.6. *Y. lipolytica* library transformation

Transformation of the sgRNA plasmid libraries into PO1f and PO1f Cas9A08 cells was accomplished with the method described in [Chen et al. \(1997\)](#), with modifications ([Chen et al., 1997](#)). For each strain of interest, seven 14 mL culture tubes, all containing 2 mL of YPD, were each inoculated with a single colony of the given strain, and grown in a 14 mL tube at 30 °C with shaking at 225 RPM for 22–24 h (final OD ~30). The cultures were then pooled, and 750 µL aliquots were distributed into 12 1.5 mL centrifuge tubes. Cells were pelleted (centrifuged at 4000 g), washed with ultrapure water, re-pelleted, and each resuspended in the Chen transformation buffer (53% PEG 4000 w/v, 0.1 M Lithium Acetate, 0.1 M DTT). Each tube was then sequentially dosed with 3 µL carrier DNA (ThermoFisher salmon sperm DNA, sheared, 10 mg/mL) and library plasmid stock (1 µg of plasmid per tube). Tubes were gently mixed and transformed by heat shock at 39 °C for 1 h. Each tube was recovered by adding 1 mL of fresh water, pelleting and aspirating supernatant. All 12 pellets were resuspended in 1 mL of fresh water, and then pooled back together. Dilutions of the transformation (0.1%, 0.01%, and 0.001%) were plated on solid SD-leu media to calculate transformation efficiency. The remaining volume of pooled transformants were then inoculated in 500 mL of SD-leu media (in a 2 L baffled flask), and grown for 48 h at 225 RPM and 30 °C. Cells were centrifuged and resuspended in fresh SD-leu media to bring their OD to 20. Glycerol stock aliquots of the full volume of transformants were then prepared by mixing 800 µL of the resuspended transformation culture with 200 µL 100% glycerol, flash freezing in liquid nitrogen, and storing at -80 °C. Four biological replicates of each strain's transformation were performed for pooling as necessary during screen experiments to ensure adequate diversity to maintain library representation and minimize the effect of plasmid instability (>100x coverage, >2.4 × 10⁶ total transformants per biological replicate).

4.7. *Y. lipolytica* acetate screen

CRISPR-Cas9 growth screens in acetate were conducted in synthetic defined minimal media from [Turki et al. \(2009\)](#) deficient in leucine ([Turki et al., 2009](#)). Three media conditions were prepared: Turki-leu 2% (w/v) glucose, Turki-leu 0.25 M acetate, and Turki-leu 0.5 M acetate. The pH of all media was adjusted to 6.3. Acetate was added as a sodium salt. 150 µL (approximately 1 × 10⁷ cells) of 2 day outgrowth glycerol stocks of PO1f Cas9 and PO1f strains transformed with the sgRNA library were used to inoculate 250 mL baffled flasks containing 25 mL of media. Two biological replicates were cultured for each different media and strain condition combinations. Outgrowth following inoculation was done at 30 °C at 220 RPM. The experiment was halted after 4 days of growth, where the OD₆₀₀ of all flasks reached 8–12. On the last day, 1 mL of culture was removed, treated with DNase I, pelleted, and frozen for later library isolation and sequencing.

4.8. Acetate screen library isolation and sequencing

Frozen 1 mL DNase I treated and pelleted culture samples from each CRISPR screen flask was thawed and resuspended in 400 µL sterile Milli-Q water. Each cell suspension was split into two, 200 µL samples. Plasmids were isolated from each sample using a Zymo Yeast Plasmid Miniprep Kit (Zymo Research). Splitting into separate samples here was done to accommodate the capacity of the Yeast Miniprep Kit, specifically to ensure complete lysis of cells using Zymolyase and lysis buffer. This step is critical in ensuring sufficient plasmid recovery and library coverage for downstream sequencing as the gRNA plasmid is a low copy number plasmid. The split samples from a single pellet were pooled, and the plasmid number was quantified using quantitative PCR with qPCR-GW-F and qPCR-GW-R and SsoAdvanced Universal SYBR Green Supermix (Biorad). Each pooled sample was confirmed to contain at least 10⁷ plasmids so that sufficient coverage of the sgRNA library is ensured.

At least 0.2 ng of plasmids (approximately 3 × 10⁷ plasmid molecules) were used as template for PCR and amplified for 16 cycles and not allowed to proceed to completion to avoid amplification bias. The PCR product was purified using SPRI beads and tested on a bioanalyzer to ensure the correct length.

Samples from the screens were prepared as previously described in [Schwartz et al. \(2019\)](#). Briefly, isolated plasmids were amplified using forward (Cr1665-Cr1668) and reverse primers (Cr1669-Cr1673; Cr1709-1711) containing the necessary barcodes, pseudo-barcodes, and adapters ([Table S3](#)). Approximately 1 × 10⁷ plasmids were used as a template and amplified for 22 cycles, not allowing the reaction to proceed to completion. Amplicons at 250 bp were then gel extracted and tested on the bioanalyzer to ensure correct length. Samples were pooled in equimolar amounts and submitted for sequencing on a NextSeq 500 at the UCR IIGB core facility. All fitness scores (FS) for the acetate screens are provided in [Supp. File 1](#).

4.9. Generating sgRNA read counts from raw reads

Next-generation sequencing raw FASTQ files were processed using the Galaxy platform ([Afgan et al., 2018](#)). Read quality was assessed using FastQC v0.11.8, demultiplexed using Cutadapt v1.16.6, and truncated to only contain the sgRNA using Trimmomatic v0.38. Custom MATLAB scripts were written to determine counts for each sgRNA in the library using Bowtie alignment (Bowtie2 v20.4.2; inexact matching) and naïve exact matching (NEM). The final count for each sgRNA was taken as the maximum of the two methods. Parameters used for each of the tools implemented on Galaxy are provided in [Table S5](#), and the demultiplexing primers/barcodes are provided in [Supp. File 3](#).

4.10. Identification of screening hits

Read counts from the glucose screen in SD-leu media in strains containing *KU70* gene knockout and Cas9 were used to calculate cutting score (CS) for each sgRNA by computing log₂ ratio of the total normalized abundances in control and treatment samples, as described in ([Ramesh et al., 2023](#)). Similarly, read counts from glucose and acetate screens in Turki media were used to determine sgRNA fitness scores (FS) as log₂ ratio of total normalized guide abundance in Cas9-containing treatment sample to that in the wildtype control. For the hydrocarbon screens, read counts from the Cas9-containing strain before inoculation onto solid media served as control to calculate sgRNA FS values. The sgRNA FS values from all datasets, along with the CS, were further used to identify significant hits from the respective screens via calculation of gene fitness scores using acCRISPR v1.0.0 ([Ramesh et al., 2023](#)). A CS-threshold of 1.5 was found to be optimum to remove low-activity sgRNAs from the optimized sgRNA library (library v2), and hence, the parameter *cutoff* was set to 1.5. Significance testing in acCRISPR is accomplished using a one-tailed or two-tailed z-test of significance. To identify essential genes for growth in glucose (Turki media), the

parameter for significance testing (*significance*) was set to ‘one-tailed’ since only genes with statistically significantly lower fitness scores were to be called. To identify genes beneficial and detrimental to growth in the two acetate conditions as well as all hydrocarbon conditions, the *significance* parameter was set to ‘two-tailed’. In all cases, genes having FDR-corrected p-value <0.05 were deemed as significant.

4.11. Hydrocarbon and fatty acid solid media preparation

Solid medium hydrocarbon plates were prepared in batches of 20 (20 mL of media per 100 mm Petri dish). First, 500 mL bottles of 2x concentration low melt agarose minimal media were prepared via autoclaving (for each bottle, 1 g MgSO₄ * 7H₂O, 4 g (NH₄)₂, 10 mg FeCl₃ * 6H₂O, 15 g LMP Agarose, 200 mg Uracil, 4 µg Myo-Inositol, 8 µg Biotin, 200 µg Thiamine HCl, 15 g KH₂PO₄, and 5.5 g K₂HPO₄), and kept molten in a 70 °C water bath until ready for use. Meanwhile, for each plate type, bottles of 200 mL ultrapure water were autoclaved and kept heated to approximately 70 °C on a heated stir plate. The amount of hydrocarbon or fatty acid used was calculated to bring the plates to the same molar amount of carbon provided by a 20 g/L glucose formulation, or 5.04 mL of dodecane, 4.93 mL of 1-dodecene, 4.24 g of margaric (heptadecanoic) acid, and 4.67 mL of oleic acid. Hydrocarbon or fatty acid was directly added to the water, briefly allowed to mix (and melt if solid), and sonicated for 10 min (85% amplitude, 3 s on 3 s off cycle). Then, 200 mL of the 2x molten minimal media low melt agarose mix was added to the homogenized fatty acid/hydrocarbon water, briefly mixed on a stir plate, and immediately dispensed into petri dishes. Plates were allowed to harden in a laminar air flow hood for 2 h and then refrigerated until needed, up to 3 days.

4.12. *Y. lipolytica* hydrocarbon and fatty acid solid media culturing

Optimized *Y. lipolytica* cell plating densities were calculated on solid medium hydrocarbon plates. Twelve hours prior to assay start, four glycerol stocks of prepared library cells (see methods 4.6), each from a separate transformation preparation, were pooled into 25 mL of SD-leu media in a 250 mL baffled flask, and grown for 12 h at 225 RPM, 30 °C. Cells were recovered as a pellet via centrifugation (at 4000 g for 5 min). To remove residual glucose, the cell pellet was washed with 15 mL of ultrapure water four times, with cells being recovered via centrifugation each time. Dilutions of the washed cells were then prepared to OD₆₀₀ absorbances of 0.6, 0.4, 0.2, and 0.1. For each plate formulation to be tested, 100 µL of each cell dilution was plated on a single plate using sterile beads (4 plates per formulation). Plated cells were allowed to dry in a laminar air flow hood for 1 h, then they were grown in a static 30 °C incubator for up to 3 days. Starting at 24 h, and every 24 h thereafter, plates were checked for colony formation. Once discrete colonies had grown large enough to be easily visualized, the plates were removed and imaged in a Biorad imager. For each plate formulation, the day at which discrete colonies could be visualized was chosen as the endpoint for the selection screen assays. Finally, we selected the maximum OD for each plate formulation that still formed discrete (instead of fused) colonies to use as the selection screen plating density.

4.13. *Y. lipolytica* hydrocarbon and fatty acid plate screen

Duplicate starter cultures of library cells were prepared 12 h prior to the start of the assay as in section 4.12. Optimal optical density for plating on each hydrocarbon were calculated to be: OD₆₀₀ 0.1 for Oleic acid and OD₆₀₀ 0.2 for Dodecane, 1-Dodecene, and Margaric acid). For each plate formulation replicate, 100 µL of the chosen cell dilution was plated onto 8 pre-prepared plates (see section 4.11) using sterile beads (16 total plates per formulation). Plated cells were allowed to dry in a laminar air flow hood for 1 h, then they were grown in a static 30 °C incubator until the previously determined screen endpoint (see section 4.12; 2 days for Dodecane, 1-Dodecene, and oleic acid, and 3 days for

Margaric acid). Upon removal from the incubator, plates were imaged in the BioRad imager.

The cells were then recovered in a two-part process prior to lysis. First, as much of the colony mass on the plate surface as possible was removed from each plate with 15 mL of dilute Tween 80 buffer (0.01% Tween 80, 0.11 M KH₂PO₄, 0.03 M K₂HPO₄) per plate, and the colonies scraped off with a plastic cell spreader. The samples from all eight plates were pooled for each replicate and centrifuged at 4000 g for 10 min to recover the cells. To recover the remaining cell mass which had burrowed into the solid media, we used a spatula to remove the solid media out of the plates, cut into small pieces, and placed in 250 mL centrifuge vessels. We added 35 g of urea to each vessel, and topped off to 100 mL with Tween 80 buffer. We then sealed and moved the vessels to a 70 °C water bath for 10 min to melt the media and free the cell mass. The vessels were then centrifuged at 4000g for 5 min. The molten agar media was aspirated away from the cell pellets, which were rinsed in ultrapure water and re-centrifuged to remove any residual agar contamination, before being combined with the previously recovered surface colony cells. After pooling washed and embedded cells, we divided the total cell mass from each replicate into three 1.5 mL centrifuge tubes, and stored in a -80 °C freezer until ready to harvest the library plasmids. The plasmid harvesting, quantification, barcode amplification, and sequencing steps were carried out in a similar manner as the acetate screen cells (see section 4.8 and 4.9), except no DNase I was used (as the harvesting process for the gel-embedded cell mass could have prematurely weakened the cell membranes and cultures were thoroughly washed). All fitness scores (FS) for the hydrocarbon screens are provided in Supp. File 1.

4.14. *Y. lipolytica* 1-dodecene hit validation

For positive hit validation, genes were knocked out using pSC012 plasmid containing the appropriate guide as in the acetate hit knockout procedure. Cells were grown overnight in YPD liquid culture (30 °C, 220 RPM) and then diluted to OD₆₀₀ of 1.0. Order of magnitude serial dilutions were performed and then were plated on 1-dodecene plates used in the initial hydrocarbon screening, but not deficient in leucine. Plates were grown in a 30 °C incubator for four days.

For negative hit validation, transformations were performed with either a knockout plasmid containing a high activity guide or a plasmid containing no guide (pSC012). All cells were inoculated into SD -leu liquid media and grown at 30 °C 220 RPM for two days to allow for a cut to occur at the targeted gene. 1-dodecene plates were prepared the same way as in the screen. Transformation cultures were diluted to an OD₆₀₀ of 0.001 and plated. After four days of growth at 30 °C, colonies were counted.

Data availability

The sgRNA raw counts, cutting scores, and fitness scores generated in this study are provided as separate Supplementary Information and Source Data files. The sgRNA library will be made available. The sgRNA sequencing data for the CRISPR-Cas9 screen generated for this study have been deposited in the NCBI SRA database under accession code PRJNA1125018.

CRediT authorship contribution statement

Nicholas R. Robertson: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Varun Trivedi:** Writing – original draft, Visualization, Formal analysis. **Brian Lupish:** Writing – original draft, Investigation, Data curation. **Adithya Ramesh:** Writing – review & editing, Investigation, Data curation. **Yuna Aguilar:** Validation. **Stephanie Carrera:** Methodology, Validation. **Sangcheon Lee:** Validation. **Anthony Arteaga:** Validation. **Alexander Nguyen:**

Formal analysis. Chase Lenert-Mondou: Validation. **Marcus Harland-Dunaway:** Writing – review & editing, Methodology, Conceptualization. **Robert Jinkerson:** Writing – review & editing, Methodology, Conceptualization. **Ian Wheeldon:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by DOE DE-SC0019093, NSF-2225878, NSF-1922642, and NSF-1803630.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.yben.2024.09.005>.

References

Adames, N.R., Gallegos, J.E., Peccoud, J., 2019. Yeast genetic interaction screens in the age of CRISPR/Cas. *Curr. Genet.* 65, 307–327.

Afghan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B.A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Sorzano, N., Goecks, J., Taylor, J., Nekrutenko, A., Blankenberg, D., 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544.

Alva, T.R., Riera, M., Chartron, J.W., 2021. Translational landscape and protein biogenesis demands of the early secretory pathway in *Komagataella phaffii*. *Microb. Cell Factories* 20, 19.

Baisya, D., Ramesh, A., Schwartz, C., Lonardi, S., Wheeldon, I., 2022. Genome-wide functional screens enable the prediction of high activity CRISPR-Cas9 and -Cas12a guides in *Yarrowia lipolytica*. *Nat. Commun.* 13, 922.

Bowman, E.K., Deane, M., Cheng, J.-F., Evans, R., Oberortner, E., Yoshikuni, Y., Alper, H.S., 2020. Bidirectional titration of yeast gene expression using a pooled CRISPR guide RNA approach. *Proc. Natl. Acad. Sci. U.S.A.* 117, 18424–18430.

Cacheria, P., Kurt, N.C., Røpke, A., Strucko, T., Mortensen, U.H., Jensen, M.K., 2024. Genome-wide host-pathway interactions affecting cis-cis-muconic acid production in yeast. *Metab. Eng.* 83, 75–85.

Castro, I.G., Shortill, S.P., Dziurdzik, S.K., Cadou, A., Ganeshan, S., Valentini, R., David, Y., Davey, M., Mattes, C., Thomas, F.B., Avraham, R.E., Meyer, H., Fadel, A., Fenech, E., J., Ernst, R., Zaremberg, V., Levine, T.P., Stefan, C., Comibe, E., Schulziner, M., 2022. Systematic analysis of membrane contact sites in *Saccharomyces cerevisiae* uncovers modulators of cellular lipid distribution. *Elife* 11. <https://doi.org/10.7554/elife.74602>.

Chen, D.C., Beckerich, J.M., Gaillardin, C., 1997. One-step transformation of the dimorphic yeast *Yarrowia lipolytica*. *Appl. Microbiol. Biotechnol.* 48, 232–235.

Chen, Y., Xiao, W., Wang, Y., Liu, H., Li, X., Yuan, Y., 2016. Lycopene overproduction in *Saccharomyces cerevisiae* through combining pathway engineering with host engineering. *Microb. Cell Factories* 15, 113.

Cho, S.K., Hofmann, S.L., 2004. pdf1, a palmitoyl protein thioesterase 1 Ortholog in *Schizosaccharomyces pombe*: a yeast model of infantile Batten disease. *Eukaryot Cell* 3, 302–310. <https://doi.org/10.1128/EC.3.2.302-310.2004>.

Comibe, E., Stevens, T.H., 2000. Vps52p, Vps53p, and Vps54p form a novel multisubunit complex required for protein sorting at the yeast late Golgi. *Mol. Biol. Cell* 11, 305–323.

Costa, C., Henriquez, A., Pires, C., Nunes, J., Ohno, M., Chibana, H., Sá-Correia, I., Teixeira, M.C., 2013. The dual role of *Candida glabrata* drug:H⁺ antiporter CgAqr1 (ORF CAGLOJ09944g) in antifungal drug and acetic acid resistance. *Front. Microbiol.* 4, 170.

Crook, N., Sun, J., Morse, N., Schmitz, A., 2016. Identification of gene knockdown targets conferring enhanced isobutanol and 1-butanol tolerance to *Saccharomyces cerevisiae* using a tunable RNAi screening. *Appl. Microbiol.*

Dabrowski, R., Tulli, S., Graef, M., 2023. Parallel phospholipid transfer by Vps13 and Atg2 determines autophagosome biogenesis dynamics. *J. Cell Biol.* 222. <https://doi.org/10.1083/jcb.202211039>.

Dalvie, N.C., Leal, J., Whittaker, C.A., Yang, Y., Brady, J.R., Love, K.R., Love, J.C., 2020. Host-informed expression of CRISPR guide RNA for genomic engineering in *Komagataella phaffii*. *ACS Synth. Biol.* 9, 26–35.

Da Silva, N.A., Srikrishnan, S., 2012. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 12, 197–214.

De Lucas, J.R., Domínguez, A.I., Valenciano, S., Turner, G., Laborda, F., 1999. The acuH gene of *Aspergillus nidulans*, required for growth on acetate and long-chain fatty acids, encodes a putative homologue of the mammalian carnitine/acylcarnitine carrier. *Arch. Microbiol.* 171, 386–396. <https://doi.org/10.1007/s002030050725>.

Debelyy, M.O., Platta, H.W., Saffian, D., Hensel, A., Thoms, S., Meyer, H.E., Warscheid, B., Girzalsky, W., Erdmann, R., 2011. Ubp15p, a ubiquitin hydrolase associated with the peroxisomal export machinery. *J. Biol. Chem.* 286, 28223–28234.

de Groot, R., Lüthi, J., Lindsay, H., Holtackers, R., Pelkmans, L., 2018. Large-scale image-based profiling of single-cell phenotypes in arrayed CRISPR-Cas9 gene perturbation screens. *Mol. Syst. Biol.* 14, e8064.

Dong, Z., Yang, S., Lee, B.H., 2021. Bioinformatic mapping of a more precise *Aspergillus Niger* degradome. *Sci. Rep.* 11, 693.

Dunn Jr., W.A., Cregg, J.M., Kiel, J.A.K.W., van der Klei, I.J., Oku, M., Sakai, Y., Sibirny, A.A., Stasyk, O.V., Veenhuis, M., 2005. Pexophagy: the selective autophagy of peroxisomes. *Autophagy* 1, 75–83.

Fickers, P., Benetti, P.-H., Waché, Y., Marty, A., Mauersberger, S., Smit, M.S., Nicaud, J.-M., 2005. Hydrophobic substrate utilisation by the yeast *Yarrowia lipolytica*, and its potential applications. *FEMS Yeast Res.* 5, 527–543.

Fujita, K., Matsuyama, A., Kobayashi, Y., Iwahashi, H., 2006. The genome-wide screening of yeast deletion mutants to identify the genes required for tolerance to ethanol and other alcohols. *FEMS Yeast Res.* 6, 744–750.

Fukuda, R., 2023. Utilization of n-alkane and roles of lipid transfer proteins in *Yarrowia lipolytica*. *World J. Microbiol. Biotechnol.* 39, 97.

Fukuda, R., 2013. Metabolism of hydrophobic carbon sources and regulation of it in n-alkane-assimilating yeast *Yarrowia lipolytica*. *Biosci. Biotechnol. Biochem.* 77, 1149–1154.

Fukuda, R., Ohta, A., 2013. Utilization of hydrophobic substrate by *Yarrowia lipolytica*. In: Barth, G. (Ed.), *Yarrowia Lipolytica: Genetics, Genomics, and Physiology*. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 111–119.

Garcia, J.M., Schwabe, M.J., Voelker, D.R., Riekhof, W.R., 2021. A functional genomic screen in *Saccharomyces cerevisiae* reveals divergent mechanisms of resistance to different alkylphosphocholine chemotherapeutic agents. *G3* 11. <https://doi.org/10.1093/g3journal/jkab233>.

Gauss, R., Kanehara, K., Carvalho, P., Ng, D.T.W., Aeby, M., 2011. A complex of Pdi1p and the mannosidase Htm1p initiates clearance of unfolded glycoproteins from the endoplasmic reticulum. *Mol. Cell* 42, 782–793.

Gervais, N.C., Halder, V., Shapiro, R.S., 2021. A data library of *Candida albicans* functional genomic screens. *FEMS Yeast Res.* 21, foab060.

Gutmann, F., Jann, C., Pereira, F., Johansson, A., Steinmetz, L.M., Patil, K.R., 2021. CRISPRi screens reveal genes modulating yeast growth in lignocellulose hydrolysate. *Biotechnol. Biofuels* 14, 41.

Horwitz, A.A., Walter, J.M., Schubert, M.G., Kung, S.H., Hawkins, K., Platt, D.M., Hernday, A.D., Mahadejkul-Meadows, T., Szeto, W., Chandran, S.S., Newman, J.D., 2015. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-cas. *Cell Syst* 1, 88–96.

Huang, M., Bai, Y., Sjostrom, S.L., Hallström, B.M., Liu, Z., Petranovic, D., Uhlén, M., Joensson, H.N., Andersson-Svahn, H., Nielsen, J., 2015. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc. Natl. Acad. Sci. U.S.A.* 112, E4689–E4696.

Imamura, Y., Yu, F., Nakamura, M., Chihara, Y., Okane, K., Sato, M., Kanai, M., Hamada, R., Ueno, M., Yukawa, M., Tsuchiya, E., 2015. RSC chromatin-remodeling complex is important for mitochondrial function in *Saccharomyces cerevisiae*. *PLoS One* 10, e0130397. <https://doi.org/10.1371/journal.pone.0130397>.

Jagtap, S.S., Bedekar, A.A., Singh, V., Jin, Y.-S., Rao, C.V., 2021. Metabolic engineering of the oleaginous yeast *Yarrowia lipolytica* PO1f for production of erythritol from glycerol. *Biotechnol. Biofuels* 14, 188.

Jakociunas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., Keasling, J.D., 2015. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metab. Eng.* 28, 213–222.

Jungmann, J., Rayner, J.C., Munro, S., 1999. The *Saccharomyces cerevisiae* protein Mnn10p/Bed1p is a subunit of a Golgi mannosyltransferase complex. *J. Biol. Chem.* 274, 6579–6585.

Kastaniotis, A.J., Autio, K.J., Sormunen, R.T., Hiltunen, J.K., 2004. Htd2p/Yhr067p is a yeast 3-hydroxyacyl-ACP dehydratase essential for mitochondrial function and morphology. *Mol. Microbiol.* 53, 1407–1421. <https://doi.org/10.1111/j.1365-2958.2004.04191.x>.

Kihara, A., Igarashi, Y., 2002. Identification and characterization of a *Saccharomyces cerevisiae* gene, RSB1, involved in sphingoid long-chain base release. *J. Biol. Chem.* 277, 30048–30054. <https://doi.org/10.1074/jbc.M203385200>.

Klein, A.T.J., van den Berg, M., Bottger, G., Tabak, H.F., Distel, B., 2002. *Saccharomyces cerevisiae* acyl-CoA oxidase follows a novel, non-PTS1, import pathway into peroxisomes that is dependent on Pex5p. *J. Biol. Chem.* 277, 25011–25019.

Kwak, Y.-S., Han, S., Thomasow, L.S., Rice, J.T., Paultz, T.C., Kim, D., Weller, D.M., 2011. *Saccharomyces cerevisiae* genome-wide mutant screen for sensitivity to 2,4-diacetylphloroglucinol, an antibiotic produced by *Pseudomonas fluorescens*. *Appl. Environ. Microbiol.* 77, 1770–1776.

Lengeler, K.B., Tielker, D., Ernst, J.F., 2008. Protein-O-mannosyltransferases in virulence and development. *Cell. Mol. Life Sci.* 65, 528–544.

Leonzino, M., Reinisch, K.M., De Camilli, P., 2021. Insights into VPS13 properties and function reveal a new mechanism of eukaryotic lipid transport. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* 1866, 159003.

Lian, J., Schultz, C., Cao, M., Hamedirad, M., Zhao, H., 2019. Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. *Nat. Commun.* 10, 1–10.

Li, S., Jendresen, C.B., Landberg, J., Pedersen, L.E., Sonnenschein, N., Jensen, S.I., Nielsen, A.T., 2020. Genome-wide CRISPRi-based identification of targets for decoupling growth from production. *ACS Synth. Biol.* 9, 1030–1040.

Liu, J., Liu, M., Shi, T., Sun, G., Gao, N., Zhao, X., Guo, X., Ni, X., Yuan, Q., Feng, J., Liu, Z., Guo, Y., Chen, J., Wang, Y., Zheng, P., Sun, J., 2022. CRISPR-assisted rational

flux-tuning and arrayed CRISPRi screening of an L-proline exporter for L-proline hyperproduction. *Nat. Commun.* 13, 891.

Liu, R., Bassalo, M.C., Zeitoun, R.I., Gill, R.T., 2015. Genome scale engineering techniques for metabolic engineering. *Metab. Eng.* 32, 143–154.

Liu, Z., Zhou, J., Wu, R., Xu, J., 2014. Mechanism of assembling isoprenoid building blocks 1. Elucidation of the structural motifs for substrate binding in geranyl pyrophosphate synthase. *J. Chem. Theory Comput.* 10, 5057–5067. <https://doi.org/10.1021/ct500607n>.

Lupish, B., Hall, J., Schwartz, C., Ramesh, A., Morrison, C., Wheeldon, I., 2022. Genome-wide CRISPR-Cas9 screen reveals a persistent null-hyphal phenotype that maintains high carotenoid production in *Yarrowia lipolytica*. *Biotechnol. Bioeng.* <https://doi.org/10.1002/bit.28219>.

Magnan, C., Yu, J., Chang, I., Jahn, E., Kanomata, Y., Wu, J., Zeller, M., Oakes, M., Baldi, P., Sandmeyer, S., 2016. Sequence assembly of *Yarrowia lipolytica* strain W29/CLIB89 shows transposable element diversity. *PLoS One* 11, e0162363.

Momen-Roknabadi, A., Oikonomou, P., Zegans, M., Tavazoie, S., 2020. An inducible CRISPR interference library for genetic interrogation of *Saccharomyces cerevisiae* biology. *Commun. Biol.* 3, 723.

Nishi, T., Nakazawa, H., Hasunuma, T., Asai, K., 2022. A streamlined strain engineering workflow with genome-wide screening detects enhanced protein secretion in *Komagataella phaffii*. *Communications*.

Novodvorska, M., Stratford, M., Blythe, M.J., Wilson, R., Beniston, R.G., Archer, D.B., 2016. Metabolic activity in dormant conidia of *Aspergillus Niger* and developmental changes during conidial outgrowth. *Fungal Genet. Biol.* 94, 23–31.

Palande, A.S., Kulkarni, S.V., León-Ramírez, C., Campos-Góngora, E., Ruiz-Herrera, J., Deshpande, M.V., 2014. Dimorphism and hydrocarbon metabolism in *Yarrowia lipolytica* var. *indica*. *Arch. Microbiol.* 196, 545–556.

Patterson, K., Yu, J., Landberg, J., Chang, I., Shavarebi, F., Bilanchone, V., Sandmeyer, S., 2018. Functional genomics for the oleaginous yeast *Yarrowia lipolytica*. *Metab. Eng.* 48, 184–196.

Ramesh, A., Trivedi, V., Lee, S., Tafrishi, A., Schwartz, C., Mohseni, A., Li, M., Lonardi, S., Wheeldon, I., 2023. acCRISPR: an activity-correction method for improving the accuracy of CRISPR screens. *Commun. Biol.* 6, 617.

Ramesh, A., Wheeldon, I., 2021. Guide RNA design for genome-wide CRISPR screens in *Yarrowia lipolytica*. *Methods Mol. Biol.* 2307, 123–137.

Sacristán-Reviriego, A., Madrid, M., Cansado, J., Martín, H., Molina, M., 2014. A conserved non-canonical docking mechanism regulates the binding of dual specificity phosphatases to cell integrity mitogen-activated protein kinases (MAPKs) in budding and fission yeasts. *PLoS One* 9, e85390.

Saikia, S., Scott, B., 2009. Functional analysis and subcellular localization of two geranylgeranyl diphosphate synthases from *Penicillium paxilli*. *Mol. Genet. Genomics* 282, 257–271. <https://doi.org/10.1007/s00438-009-0463-5>.

Schwartz, C., Cheng, J.-F., Evans, R., Schwartz, C.A., Wagner, J.M., Anglin, S., Beitz, A., Pan, W., Lonardi, S., Blenner, M., Alper, H.S., Yoshikuni, Y., Wheeldon, I., 2019. Validating genome-wide CRISPR-Cas9 function improves screening in the oleaginous yeast *Yarrowia lipolytica*. *Metab. Eng.* 55, 102–110.

Schwartz, C., Curtis, N., Löbs, A.-K., Wheeldon, I., 2018. Multiplexed CRISPR activation of cryptic sugar metabolism enables *Yarrowia lipolytica* growth on cellobiose. *Biotechnol. J.* 13, e1700584.

Schwartz, C., Frogue, K., Ramesh, A., Misa, J., Wheeldon, I., 2017. CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*. *Biotechnol. Bioeng.* 114, 2896–2906.

Seaman, M.N., McCaffery, J.M., Emr, S.D., 1998. A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. *J. Cell Biol.* 142, 665–681.

Skaggs, B.A., Alexander, J.F., Pierson, C.A., Schweitzer, K.S., Chun, K.T., Koegel, C., Barbuch, R., Bard, M., 1996. Cloning and characterization of the *Saccharomyces cerevisiae* C-22 sterol desaturase gene, encoding a second cytochrome P-450 involved in ergosterol biosynthesis. *Gene* 169, 105–109.

Smith, J.D., Suresh, S., Schlecht, U., Wu, M., Wagih, O., Peltz, G., Davis, R.W., Steinmetz, L.M., Parts, L., St Onge, R.P., 2016. Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design. *Genome Biol.* 17, 45.

Stovicek, V., Borja, G.M., Forster, J., Borodina, I., 2015. EasyClone 2.0: expanded toolkit of integrative vectors for stable gene expression in industrial *Saccharomyces cerevisiae* strains. *J. Ind. Microbiol. Biotechnol.* 42, 1519–1531.

Sun, X., Wang, F., Lan, N., Liu, B., Hu, C., Xue, W., Zhang, Z., Li, S., 2019. The Zn(II) 2Cys6-type transcription factor ADA-6 regulates conidiation, sexual development, and oxidative stress response in *Neurospora crassa*. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.00750>.

Thompson, N.A., Ranzani, M., van der Weyden, L., Iyer, V., Offord, V., Droop, A., Behan, F., Gonçalves, E., Speak, A., Iorio, F., Hewinson, J., Harle, V., Robertson, H., Anderson, E., Fu, B., Yang, F., Zagnoli-Vieira, G., Chapman, P., Del Castillo Velasco-Herrera, M., Garnett, M.J., Jackson, S.P., Adams, D.J., 2021. Combinatorial CRISPR screen identifies fitness effects of gene paralogues. *Nat. Commun.* 12, 1–11.

Thorwall, S., Schwartz, C., Chartron, J.W., Wheeldon, I., 2020. Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. *Nat. Chem. Biol.* 16, 113–121.

Tirnauer, J.S., Bierer, B.E., 2000. Eb1 proteins regulate microtubule dynamics, cell polarity, and chromosome stability. *J. Cell Biol.* 149, 761–766.

Tkachenko, A.A., Borshchhevskaya, L.N., Sineokiy, S.P., Gordeeva, T.L., 2023. CRISPR/Cas9-mediated genome editing of the *Komagataella phaffii* to obtain a phytase-producer markerless strain. *Biochemistry* 88, 1338–1346.

Trivedi, V., Ramesh, A., Wheeldon, I., 2023. Analyzing CRISPR screens in non-conventional microbes. *J. Ind. Microbiol. Biotechnol.* 50. <https://doi.org/10.1093/jimb/kuad006>.

Try, S., De-Coninck, J., Voilley, A., Chunhieng, T., Waché, Y., 2018. Solid state fermentation for the production of γ -decalactones by *Yarrowia lipolytica*. *Process Biochem.* 64, 9–15.

Turki, S., Kraein, I.B., Weeckers, F., Thonart, P., Kallel, H., 2009. Isolation of bioactive peptides from tryptone that modulate lipase production in *Yarrowia lipolytica*. *Bioresour. Technol.* 100, 2724–2731.

Wang, G., Björk, S.M., Huang, M., Liu, Q., Campbell, K., Nielsen, J., Joensson, H.N., Petranovic, D., 2019. RNAi expression tuning, microfluidic screening, and genome recombining for improved protein production in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 9324–9332.

Wang, Y., Zhang, Z., Lu, X., Zong, H., Zhuge, B., 2020. Genetic engineering of an industrial yeast *Candida* glycerinogenes for efficient production of 2-phenylethanol. *Appl. Microbiol. Biotechnol.* 104, 10481–10491.

Wetterau, J.R., Combs, K.A., Spinner, S.N., Joiner, B.J., 1990. Protein disulfide isomerase is a component of the microsomal triglyceride transfer protein complex. *J. Biol. Chem.* 265, 9800–9807.

Xu, Q., Bai, C., Liu, Y., Song, L., Tian, L., Yan, Y., Zhou, J., Zhou, X., Zhang, Y., Cai, M., 2019. Modulation of acetate utilization in *Komagataella phaffii* by metabolic engineering of tolerance and metabolism. *Biotechnol. Biofuels* 12, 61.