UNIVERSITY OF CALIFORNIA
RIVERSIDE

Industrial Strain Optimization through Genome-Wide Knockout Screens in *Yarrowia Lipolytica*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Bioengineering

by

Brian Alexander Lupish

September 2024

Dissertation Committee:
Dr. Ian Wheeldon, Chairperson
Dr. Katherine Borkovich
Dr. Joshua Morga

The Dissertation of Brian Alexander Lupish is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

First, I would like to thank my advisor, Dr. Ian Wheeldon. As chaotic as things were with my projects, I really appreciate your patience and honest feedback in working with me to get me to where I needed to be.

I would also like to thank the rest of my committee, Dr. Joshua Morgan and Dr. Katherine Borkovich. You both gave me lots of good advice and occasional emotional support throughout this journey and helped me keep pushing forward towards this point.

I would like to thank our postdoc Dr. Clifford Morrison, for all of his mentorship and patience in the lab. It really helped me through the most difficult times of my PhD and got me through to the other end.

I would like to acknowledge two of my published works which were utilized as the bulk of the first and third chapters of this dissertation. The first chapter was primarily drawn from "Genome-wide CRISPR-Cas9 screen reveals a persistent null-hyphal phenotype that maintains high carotenoid production in Yarrowia lipolytica", published in *Biotechnology and Bioengineering* (Lupish et al, 2022), and the third chapter was primarily drawn from "Optimized genome-wide CRISPR screening enables rapid engineering of growth-based phenotypes in Yarrowia lipolytica", pre-published in ***b**ioRxiv* (Robertson et al, 2024).

I would like to acknowledge all of my co-authors on the two published manuscripts included in this dissertation, including Dr. Robert Jinkerson, Dr. Adithya Ramesh, Dr Corey Schwartz, Jordan Hall, Nick Robertson, Varun Trivedi, Yuna Aguilar, Anthony Artega, Alexander Nguyen, Sangcheon Lee, Chase Lenert-Mondou, and Marcus Harland-

Dunaway. Your hard work and dedication was essential to seeing this work to completion.

I would like to acknowledge funding sources for the work presented in this dissertation. Specifically, the work presented in Chapter 1 was supported by NSF 1706545 and the U.S. Department of Energy (DOE) Joint Genome Institute (JGI) grant CSP-503076. Said work conducted by the JGI, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. The work presented Chapters 2 and 3 was supported by DOE DE-SC0019093, NSF-2225878, NSF-1922642, and NSF-1803630.

I would like to thank my lab mates, both former and current, for their substantial contributions to this work, as well as for the support and friendship you all provided to me during my time here. In addition to my co-authors, I also want to thank current and former lab members Mario Lopez, Dominic Biondo, Aida Tafrishi, Eva Ottum, Adrian Garcia, Greg Douhan, Xiao Hong, Jason Rosales, Jesus Sosa, Leonardo Sanchez Zamora, Norman Seder, Reyhane Ghorbani Nia, Xuye Lang, Stephanie Carrera, Trinity Nguyen, Zach Hartley, Sarah Thorwall, Shuang Wei, Mary Wan, and Troy Alva.

I would like to thank my professional colleagues Monique Quinn, Megan Holms and Dr. Jessica Trinh, for their generosity and willingness to write and help me execute code to sort and process my data, and to work with me to design rational strategies for analyzing my data.

And finally, I would like to thank my family, especially my parents, Ronald Lupish and Lynn Vroblick, and my stepmother, Leslie Herock, for all of the love and support you've

shown me along this journey. After all these years, I finally made it, and I'm so happy that you can finally see me attain this degree after all the years leading up towards it.

ABSTRACT OF THE DISSERTATION

Industrial Strain Optimization through Genome-Wide Knockout Screens in *Yarrowia Lipolytica*

by

Brian Alexander Lupish

Doctor of Philosophy, Graduate Program in Bioengineering
University of California, Riverside, September 2024
Dr. Ian Wheeldon, Chairperson

Yarrowia lipolytica is a versatile oleaginous yeast used for bioprocessing and bioproduction with uses ranging from supplement production to bioremediation. Its ability to metabolize a wide range of carbon sources, generate and store surplus Acetyl-CoA and lipids, and grow in a wide range of environmental conditions make it a promising candidate for a wide range of novel bioprocessing applications. Engineering optimized strains of *Y. lipolytica* remains a challenge due to an incomplete understanding of its genome and biological function under industrially-relevant conditions. An efficient approach to address these knowledge gaps and design optimized strains is a genome-wide CRISPR knockout screen. Here, we use optimized whole-genome gRNA libraries and innovative bioinformatic pipelines to carry out functional knockout screens and identify genes to target for strain engineering. First, we conducted a qualitative screen to identify knockouts that abolish hyphal formation, multicellular filaments that can interfere with industrial bioreactor function. We identified a benign null-hyphal knockout Δ*RAS2* with equivalent growth and production titer characteristics of existing strains. We also developed a pipeline to identify potential gene knockouts that may optimize the usage of

Y. lipolytica for solid state fermentation, a growing need in industrial bioprocessing. We used GO-term analysis of the screen results to elucidate putative differences in the relative importance of different cellular systems between solid and liquid environments. We identified genes responsible for endomembrane system and mitochondrial function as more essential under solid conditions, and genes responsible for biosynthesis, metal ion/redox, and ribosomal function as more essential under liquid conditions. Finally, we screened for gene knockouts with improved metabolism of non-glucose carbon sources and identified several with improved acetate catabolism and growth with the knockouts ΔE37234g, ΔE01193g, ΔC02904g, and ΔD21022g exhibiting the best growth and acetate catabolism characteristics. We also identified gene knockouts with preliminary improvements to hydrocarbon and fatty acid catabolism including ΔE36308g, ΔE03584g, ΔA18344g and ΔD20202g. These discoveries validate our genome wide knockout libraries and screen pipelines while providing novel gene targets for industrial *Y. lipolytica* strain engineering.

Table of Contents

List of Figures

Chapter 1 Figures:

Chapter 2 Figures:

Chapter 3 Figures:

Supplemental Figures:

List of Tables and Files

Supplemental Tables:

Supplemental Files (hyperlinks):

Introduction: Addressing Unmet Industrial Needs of the yeast *Yarrowia Lipolytica* with CRISPR Screens

Industrial Bioprocessing and *Yarrowia lipolytica*

The biotechnology and bioprocessing sectors are continuing to advance, with a constant need for new biomanufacturing and bioprocessing technologies. The global next-generation biomanufacturing market has been predicted to keep growing, with a compound annual growth rate (CAGR) of 7.5% between 2024 and 2032, reaching 39.4 billion dollars at the end of that period (Next-gen Biomanufacturing Market Size...). Furthermore, the synthetic biology market may grow to as much as 100 billion dollars by 2030 (U.S. Government Accountability Office). In order to sustain these trends and meet unmet needs in the space, an expanded repertoire of nonconventional bioproduction and bioprocessing hosts will be required (Garvey 2022; Seppälä et al. 2017).

One microorganism that has already seen some adoption to meet the evolving needs of bioproduction and bioprocessing is *Yarrowia lipolytica*. It is an aerobic, oleaginous yeast with GRAS (generally regarded as safe) classification, enabling its use in food products (Gonçalves et al. 2014; Mamaev and Zvyagilskaya 2021). It has an excellent capacity to synthesize and store a wide range of fatty acids and related metabolites, in part due to its high Acetyl-CoA flux (Liu et al. 2019). It also can utilize a wide range of carbon sources

beyond glucose, including , glycerol, triglycerides, and alkanes (Tenagy et al. 2015; Fickers et al. 2005; Papanikolaou et al. 2002). Its great flexibility coupled with its rapid growth and ease of use have made it a prime candidate for broad industrial use.

Y. lipolytica has already shown promise across several biotechnological applications. It is inherently well suited to the production of industrial lipases and biosurfactants, given its oleaginous characteristics (Yu et al. 2007; Csutak et al. 2015). By extension, it can also be used to produce a wide range of lipids and Acetyl-CoA derived products such as lycopene (Fontanille et al. 2012; Matthäus et al. 2014). Other, less conventional uses include bioremediation and contaminant biosensing, due to its tolerance for and ability to metabolize hydrocarbons (Żogała et al. 2005; Alkasrawi et al. 1999). Despite its wide range of current and potential uses, there are unresolved difficulties that limit *Y. lipolytica*'s broader adoption.

There are several remaining challenges that must be addressed before *Y. lipolytica* can reach its full potential as a biotechnology host. Even with recent progress, more work needs to be done in annotating the genes and proteins of unknown function in the various strains of *Y. lipolytica* (Ganesan et al. 2019). Furthermore, *Y. lipolytica* is strictly aerobic, and has extremely high oxygen requirements; meeting those requirements becomes increasingly difficult as a bioprocess' scale increases (Kar et al. 2012). Bioprocess scaling is further complicated by *Y. lipolytica*'s unpredictable dimorphism, in which both a budding yeast and multicellular hyphal state may occur, disrupting nutrient, pH, and oxygen diffusion characteristics (Vandermies and Fickers

2019). In order to address these issues, improved approaches for studying the remaining unknowns in the *Y. lipolytica* genome are required.

Genome Wide Screens as a Strain Engineering Tool

Genome-wide screens are a systematic approach used to study entire genomes of host organisms. These screens identify coding sequences that are associated with specific traits, functions, or processes (Friedman and Perrimon 2004). While they can theoretically be carried out in many different species, including multicellular animals such as *Caenorhabditis elegans* and *Drosophila melanogaster*, single celled eukaryotes such as yeast are the simplest to conduct genome wide screens on (Friedman and Perrimon 2004). Generally, the screen method either knocks out genes, or alters their expression levels. Early genome-wide yeast studies used mitotic recombination to delete genes one by one, which led to the first characterization of most possible *S. cerevisiae* single gene knockouts (Winzeler et al. 1999; Giaever et al. 2002). Such methods are cumbersome to scale up, especially in species with larger genomes and less recombination activity (Guha and Edgell 2017). Overexpression of specific genes may be accomplished through high-copy plasmid transformations (Sandoval et al., 2011), which may better elucidate important phenotypes than knockouts. Unfortunately, this method is also time consuming and resource intensive to scale up. RNA interference techniques were adopted as a higher throughput and more flexible screening option, in which interfering RNA molecules could partially or fully suppress gene transcription and activity(Wang et al. 2019; Crook et al. 2016; Müller et al. 2005). Even with improved

coverage and efficiency, RNAi screens may suffer from off-target interference (Neumeier and Meister 2020) and incomplete silencing of targeted genes (Drinnenberg et al. 2009). Transposon libraries, in which insertional mutagenesis inserts a marker into a gene loci, allows for simple selection of knockout species (Liu et al., 2022). However, insertion biases can lead to open reading frame ORFs not being disrupted, leading to an incomplete knockout (Liu et al., 2022). More recently, CRISPR-derived whole-genome screening approaches such as CRISPR knockout, CRIPSRi (targeted gene inhibition), and CRISPRa (targeted gene activation) have provided comprehensive genomic coverage and high specificity in comparison to older methods (Liu et al. 2022; Misa and Schwartz 2021). CRISPRi and CRISPRa screens give valuable insight into expression level effects on genes, but there can be some ambiguity in output data regarding screen hits and gene essentiality (Liu et al. 2022; Misa and Schwartz 2021). Nevertheless, CRISPR screens have become one of the most prolific genome wide screening technologies.

Further advances in high throughput sequencing and CRISPR-editing systems led CRISPR knockout screening methods to excel at elucidating novel gene targets in various yeasts (Ramesh et al., 2023). CRISPR systems utilize Cas protein to make double-strand breaks at precise sites in host organism chromosomes in tandem with customizable guide nucleotides(Raschmanová et al. 2018). These cuts can be used to generate knockouts via non-homologous end joining (NHEJ) (Horwitz et al. 2015), or gene insertions via homology directed repair (HDR) (Stovicek et al. 2015). Either way, to work in a multiplex or genome-wide screen, a guide RNA (gRNA) library is required (Friedman and Perrimon 2004). These libraries consist of plasmids containing gRNA sequences,

each coding for a target cut site in the host organism chromosome(s). Generally, libraries target either all of the genes in a host genome (Schwartz et al. 2019), or a specific subset (Thompson et al. 2021). Several considerations must go into their design, including host species compatibility and genome coverage(Adames et al. 2019). Once assembled, CRISPR knockout screening systems can be utilized to call essential genes in specified conditions or find gene knockouts that enhance fitness.

Genome-wide CRISPR screens may be designed and analyzed in several ways. Individually prepared mutants may be analyzed as part of a panel in an arrayed screen (Fujita et al. 2006). The outputs from arrayed screens are typically measurable phenotypes. Often, the measured output is a change in growth, ranging from death to enhanced growth (Garcia et al. 2021). Alternatively, other observable changes may be measured, such as cell morphology or substrate utilization (Lupish et al. 2022; Gutmann et al. 2021). While arrayed screens are highly effective for the quantification of rare or desirable phenotypes, they are labor intensive and require isolated strain preparations for each mutant to be tested (de Groot et al. 2018).

Unlike arrayed screens, pooled screens combine large numbers of mutants into a single culture (Bowman et al. 2020). A highly efficient transformation method must be used to ensure full library representation in the transformant population (Schwartz et al., 2019). This theoretically creates all possible single knockouts (or gene inhibitions/activations) of an organism within a small culture volume, which enables the cells to diverge in survival or phenotype based on selection pressures (Smith et al.

2016). While individual genotypes cannot be directly observed in the pooled culture, deep sequencing can be used to find gRNA prevelances and by extension, essential genes or genes involved in a phenotype of interest (Ramesh et al., 2023). Still, there are remaining challenges in designing and utilizing genome-wide CRISPR screen guide libraries. They must have complete (or near complete) coverage of the host genome, as well as efficient guide activities to ensure that knockouts occur consistently. gRNA expression characteristics must also be taken into account to ensure sufficient genome coverage (Dalvie et al. 2020). To overcome these challenges, CRISPR gRNA libraries need to be highly optimized to the species and strain that they are targeting.

The broad flexibility of genome wide screens lends them well to the study of non conventional yeast such as *Yarrowia lipolytica*, with several types of these screens having been utilized successfully with *Yarrowia lipolytica* in recent years. One approach tested putative transcription factors with an arrayed screen that tested various transcription factor constructs against various integrated promoter sites tied to a fluorescent protein (Leplat et al. 2015; Leplat et al. 2018). Another method utilized saturation mutagenesis, generating a transposon library to analyze relative essentiality of genes (Patterson et al. 2018). A third protocol cloned digested fragments of *Y. lipolytica* gDNA into vectors to generate a gDNA fragment library, which was re-transformed into cells to screen for propionate tolerance (Park and Nicaud 2020). While all of these approaches showed promise, the precision and comprehensiveness of the CRISPR knockout screen makes it an ideal method for engineering beneficial strains of *Y. lipolytica*.

Engineering Industrially Relevant *Yarrowia lipolytica* Strains with CRISPR Knockout Screens

CRISPR screens have proven to be especially effective for strain engineering, and with the development of optimized Cas9 and gene integration systems for *Y. lipolytica* in our research group, it has become feasible to conduct these types of screens in *Y. lypolytica* (Schwartz et al. 2017; Schwartz et al. 2017; Schwartz et al. 2016). Our group has conducted several CRISPR screens using these tools, helping to generate improved lycopene production strains and enabling cellobiose catabolism (Schwartz et al. 2018; Schwartz et al. 2017). More recently, we developed a whole genome knockout sgRNA library for use with our CRISPR screens, which enabled us to screen for knockouts with improved lipid accumulation (Schwartz et al. 2019). Output from these whole genome knockout screens has been further improved with guide activity prediction and correction methods (Baisya et al. 2022; Ramesh et al. 2023). With this expansive toolset, we could then proceed to a new series of whole genome knockout CRISPR screens conducted with novel methods and refined guide libraries.

In the following studies, we utilized two optimized whole genome guide libraries to conduct CRISPR knockout screens to identify and engineer several industrially beneficial traits into *Yarrowia lipolytica*. We first carried out a pooled whole genome knockout screen to identify single gene knockouts that prevent the formation of hyphae, a trait which can interfere with bioreactor production runs. We then carried out a pooled knockout screen to identify deletions that improve and inhibit growth on solid media, for

the purpose of expanding existing knowledge of strain engineering for solid state fermentation applications. Finally, we carried out pooled knockout screens with non-glucose carbon sources, including acetate, two alkanes, and two fatty acids, so we could begin engineering strains of *Y. lipolytica* optimized for bioremediation and the use of inexpensive industrial waste as their carbon source for bioproduction. These experiments generated promising knockout candidates for the desired purposes of the screens, while expanding our knowledge of the genes involved with the relevant cellular functions. In doing so, we demonstrated the efficacy of using pooled whole genome knockout screens for engineering industrially relevant strains of *Y. lipolytica*.

Chapter 1: Genome-wide CRISPR-Cas9 screen reveals a persistent null-hyphal phenotype that maintains high carotenoid production in *Yarrowia lipolytica*

1.1 Chapter 1 Abstract

Yarrowia lipolytica is a metabolic engineering host of growing industrial interest due to its ability to metabolize hydrocarbons, fatty acids, glycerol and other renewable carbon sources. This dimorphic yeast undergoes a stress-induced transition to a multicellular hyphal state, which can negatively impact biosynthetic activity, reduce oxygen and nutrient mass transfer in cell cultures, and increase culture viscosity. Identifying mutations that prevent the formation of hyphae would help alleviate the bioprocess challenges that they create. To this end, we conducted a genome-wide CRISPR screen to identify genetic knockouts that prevent the transition to hyphal morphology. The screen identified five mutants with a null-hyphal phenotype – ΔRAS2, ΔRHO5, ΔSFL1, ΔSNF2, and ΔPAXIP1. Of these hits, only ΔRAS2 suppressed hyphal formation in an engineered lycopene production strain over a multi-day culture. The RAS2 knockout was also the only genetic disruption characterized that did not affect lycopene production, producing more than 5 mg L^{-1} OD^{-1} from an heterologous pathway with enhanced carbon flux through the mevalonate pathway. These data suggest that a ΔRAS2 mutant of *Y. lipolytica* could prove useful in engineering a metabolic engineering host of the production of carotenoids and other biochemicals.

1.2 Chapter 1 Introduction

Nonconventional filamentous and dimorphic fungi are of growing interest for bioproduction due to their abilities to metabolize a range of carbon sources and to produce biomolecules with high titers. One such fungus is *Yarrowia lipolytica*, an oleaginous dimorphic yeast that has desirable traits for industrial applications (Abdel-Mawgoud et al. 2018; Löbs et al. 2017; Zhu and Jackson 2015). It is well suited for lipid biosynthesis, the production of fatty acids, carotenoids, and other acetyl-CoA-derived molecules (Blazeck et al. 2014; Morgunov et al. 2004; Schwartz et al. 2017a; Wang et al. 2022; Xue et al. 2013) and advanced genome editing tools are available to enable rapid pathway and strain design (Baisya et al. 2022; Ramesh et al. 2020; Schwartz et al. 2017b; Schwartz et al. 2018). Despite these successes, a number of technical challenges must be overcome prior to the widespread use of *Y. lipolytica* for industrial bioproduction (Czajka et al. 2018; Sabra et al. 2017). Among these challenges are the obstacles posed by dimorphism and hyphal formation (Worland et al. 2020).

Hyphae are filamentous multicellular structures with contiguous parallel cell walls separated by septa, structures that form in multiple types of fungi including several budding yeasts (Crampin et al. 2005; Kiss et al. 2019). While the formation of hyphae is a default state for many fungal species, several yeasts including *Y. lipolytica* are dimorphic, that is, they are able to exist in a single cell free-floating state as well as a hyphal or pseudohyphal state (Ruiz-Herrera and Sentandreu 2002; Vallejo et al. 2013). Transition into a hyphal or pseudohyphal state may result as a response to stressors such as starvation, high temperature, high pH, or low concentrations of dissolved oxygen

(Bellou et al. 2014; Cullen and Sprague 2000; Gimeno et al. 1992; Kawasse et al. 2003; Lee and Elion 1999; Ruiz-Herrera and Sentandreu 2002; Sudbery et al. 2004; Szabo 1999). Although some bioproduction strategies benefit from hyphal morphology (Fickers et al. 2009), hyphal formation in a bioreactor can be problematic due to reduced mass transfer of dissolved oxygen and nutrients, increased culture viscosity, or increased stress response due to hyphal shearing (Ahamed and Vermette 2010; Cai et al. 2014; Harvey and McNeil 1994; Li et al. 2002; Martin and Bushell 1996; Müller et al. 2003). Hyphal growth is also associated with lower biosynthetic activity and product yields in some yeasts including ethanol from *S. cerevisiae* (Reis et al. 2013) and lipid accumulation in *Y. lipolytica* (Bellou et al. 2014; Gajdoš et al. 2016). Therefore, it would be beneficial to engineer a null-hyphal phenotype strain of *Y. lipolytica*.

Compared to filamentous fungi, yeasts have fewer hyphae generating pathways and a simplified process of formation (Kiss et al. 2019). In addition, the transition from unicellular to hyphae often terminates in pseudo-hyphal morphology, with some yeasts never forming true hyphae (Berman and Sudbery 2002; Pomraning et al. 2018). Nevertheless, the yeast-to-hyphae transition in dimorphic yeast still functions as a survival mechanism in nutrient-limited or stressful environments (Pomraning et al. 2018). While the full process of the yeast-to-hyphal transition is not completely understood in *Y. lipolytica*, two types of signaling cascades are known to be involved: two mitogen activated phosphorylation kinase (MAPK) cascades and a protein kinase A (PKA) pathway, both of which share notable conservation across dimorphic yeasts (Gancedo 2001; Tisi et al. 2014). Despite the known relevance of these pathways in the yeast to

hyphal transition, there remain many unidentified upstream and downstream components and poorly understood regulatory functions. Therefore, a broader coverage approach is needed to identify viable genetic targets for a null-hyphal phenotype.

Here, we used a pooled CRISPR-Cas9 knockout screen to identify null-hyphal knockout candidates in *Y. lipolytica*; several phenotypic knockouts were identified by selection of colonies with a smooth appearance, which is indicative of a loss of hyphal formation. We subsequently characterized the knockout strains' growth rates and retention of the null hyphal phenotype. In doing so, we identified RAS2 disruption as the most promising for industrial use. Finally, we tested the knockout's effect on lycopene production, demonstrating production without a loss of product titer in a strain devoid of hyphae.

1.3 Chapter 1 Material and Methods

1.3.1. Strains construction

All strains were derived from *Yarrowia lipolytica* PO1f (MatA, leu2-270, ura3-302,xpr2-322, axp-2). The Cas9 strain used in the genome-wide screen and growth assays was created by integrating a codon optimized copy of Cas9 from *S. pyogenes* into the A08 locus of *Y. lipolytica* by markerless integration. As described in (Schwartz et al. 2019), the genome copy of Cas9 was expressed using a UAS1B8-TEF(136) promoter (Blazeck et al. 2011) with a ScCYC terminator. PO1f-HMEBI, the lycopene overproduction strain,

was engineered as previously described in (Schwartz et al. 2017c). Briefly, the HMG1, MVD1, CrtE, CrtB, and CrtI were integrated in the genome as follows: two copies of HMG1 were inserted, one into the D17 and a second into the XDH site; one copy of MVD1 was inserted into the disabled LEU2 site, one copy of CrtE was inserted into the A08 locus; one copy of CrtB was inserted into the AXP locus, and one copy of CrtI was inserted into the XPR2 site. In addition to these insertions, the leucine and uracil auxotrophies were alleviated by randomly integrating function copies of the *Y. lipolytica* LEU2 and URA3 genes.

1.3.2 Media and culture conditions.

Unless otherwise noted, all *Y. lipolytica* strains were culture YPD media (1% Bacto yeast extract, 2% Bacto peptone, 2% glucose). For the genome-wide screen and RAS2 rescue assays, cell were grown in synthetic defined SC Leucine Deficient media (0.17% Yeast Nitrogen Base, 0.2% Leu deficient amino acid mix (5.6% of all 19 L-Amino acids except for Leucine, 5.6% inositol, 5.6% uracil, 1.3% adenine, and 1.3% para-aminobenzoic acid), 0.5% Ammonium Sulfate). When grown for lycopene biosynthesis, we used YPD10 media (1% Bacto yeast extract, 2% Bacto peptone, 10% glucose). All *Y. lipolytica* cultures were conducted in 25 mL of media in 250 mL baffled shake flasks at 30^{o}C and 225 RPM, inoculated to an OD_{600} of 0.1 using an overnight starter culture. All *Y. lipolytica* transformations as previously described (Schwartz et al. 2017a; Schwartz et al. 2019; Schwartz et al. 2017c)

DH5α *E. coli* cells were used to clone and propagate the plasmids used in the whole genome screen and subsequent gene knockouts. NEB TOP10 *E. coli* cells were used to clone and propagate the plasmids used in the RAS2 rescue assay. All *E. Coli* cultures were grown in Lysogeny Broth (1% Tryptone, 0.5% Yeast Extract, 1% NaCl) with 100 mg/L ampicillin for selective pressure.

1.3.3 CRISPR-Cas9 screening for null-hyphal phenotypes.

To screen for genetic knockouts that lack hyphal morphology we generated upward of 50,000 colonies representing members of the genome-wide knockout library. This library was previously generated (Schwartz et al. 2019) and includes the functional disruption of more than 94% of all protein coding sequences in the PO1f strain. All colonies with a smooth phenotype (indicative of a loss of hyphal morphology) were visually identified and subjected to colony PCR. Sequencing of the CRISPR plasmids contained in each hit revealed five unique hits – YALI1_E35305g, YALI1_D05956g, YALI1_D30097g, YALI1_E30639g, and YALI_F04690g.

1.3.4 Mutant strain growth rate.

Each of the five null-hyphal mutants identified in the genome wide screen were characterized in terms of growth rate in shake flask cultures. Growth rates were determined by linearizing the mean OD values via natural logarithm calculations and subsequent linear regression, generating a slope representative of the growth rate.

1.3.5 Plasmid Construction.

To generate genetic knockouts in the lycopene overexpression strain, we digested pCRISPRyl (Addgene #70007) with AvRII and then integrated a double stranded sgRNA insert targeting one of the following genes, RAS2, RHO5, SFL1 and MHY1. We generated the sgRNA inserts by annealing complementary oligos, and integrated each insert into the pCRISPRyl linearized backbone via Gibson assemblies. To rescue RAS2 function, we replaced the hrGFP ORF in pIW209 with the RAS2 ORF cloned from wild type PO1f genomic DNA via Golden Gate Assembly. All guide sequences and primers used here are provided in **Supplemental Table S3**.

1.3.6 Hyphal phenotype characterization.

To measure the prevalence of hyphal phenotypes, the four generated. PO1f-HMEBI knockouts (PO1f-HMEBI ΔRAS2, PO1f-HMEBI ΔRHO5, PO1f-HMEBI ΔSFL1, and PO1f-HMEB IΔMHY1) were grown in duplicate for 10 days, along with PO1f and PO1f-HMEBI, which served as controls. On days 1, 3, 6, 9, and 10, one slide from each duplicate of each strain was made using a 2 μL sample of the culture. Six photographs of each slide were taken (for a total of twelve photographs per strain per measurement day) with an Olympus BX51 Microscope on brightfield settings while using a 100x oil objective. ImageJ software was used to count the total number of cells in each image, and any cell with a length greater than twice its width was classified as exhibiting hyphal behavior.

1.3.7 RAS2 rescue assay.

PO1f ΔRAS2 was transformed with the RAS2 rescue vector to generate a rescued RAS2 phenotype. PO1f and PO1f ΔRAS2 were transformed with an empty vector to serve as positive and negative controls, respectively. Three mL cultures of all 3 strains were grown overnight at 30 ^{o}C. The next day, with the cultures in log phase, they were diluted to ODs of 0.2. A 4 uL droplet from all three dilutions was spotted onto a 1.2 % Agar SC leucine deficient plate. The plate was grown for ~45 hours, until the spots had matured. The spots were then photographed and their morphologies analyzed. Cells from each mutant were scraped from the plates and resuspended in 1 mL of SC leucine deficient media. A 2 µL sample of each solution was then visualized via confocal microscopy at 60x magnification.

1.3.8 Lycopene Quantification.

Lycopene production cultures were grown as described in Schwartz et al. 2017, using 10% glucose media. Lycopene was extracted and quantified using a method detailed by Chen et al. 2016 (Chen et al. 2016) for carotenoid extraction, with a few adaptations. At each measurement time point, triplicate 1mL aliquots were withdrawn from each assay culture, and dry cell weights (DCW) of each were measured through centrifugation at 5000 g for 3 min, media aspiration, and subsequent pellet drying at 80^{o}C until stable weights were measurable. The pellets were then washed with water, re-pelleted, resuspended in 1mL of 3M HCl, and boiled for 2min. The boiled pellets were then cooled in an ice bath for 3 min. After another water wash, the pellets were resuspended in 1 mL of acetone. 200 µl of 500-750 µm glass beads were added to each acetone

resuspension, and the cells were lysed (while achieving liposome disruption) by vortexing the mixtures for 2 minutes. The resulting supernatant was collected, and lycopene titers were quantified by measuring the supernatant absorbance at 472 nm comparing measurements to a standard curve of purchased lycopene (Sigma – Aldrich); see **Supplemental Figure S1.4**.

1.4 Chapter 1 Results and Discussion

1.4.1 Phenotype Genome Wide Knockout Screen for Null-Hyphal mutants

To identify genes associated with hyphal formation, we used a pooled library of CRISPR-Cas9 sgRNAs to target nearly every gene in the genome of *Y. lipolytica* PO1f (Schwartz et al. 2019). The previously designed library covers 7,854 coding sequences (CDS) with ~6-fold coverage. Unique sgRNAs were designed to target the first 300 exon base pairs in each CDS, then scored and ranked based on their predicted on-target cutting efficiency (Doench et al. 2014). The final library contained the 6 highest scoring sgRNAs for each

CDS, along with a negative control set of 480 nontargeting sgRNAs. Oligos encoding each sgRNA were commercially synthesized and subsequently cloned into an expression vector with sgRNA expression driven by a synthetic RNA polymerase III (Pol III) promoter (Schwartz et al. 2016), while Cas9 expression was accomplished through a genome-integrated expression cassette.

Gene name and function of the morphology screen hits were identified through a BLASTp search, the results of which are shown in **Supplemental Table S1**. Notably, many of the screening hits correspond to a gene known to be associated with regulating cell morphology or cell stress: *RAS2* (YALI1_E35305g) encodes for a GTP binding protein involved in starvation response and cell morphology (Li et al. 2014; Mösch et al. 1996; Mösch and Fink 1997); *SFL1* (YALI1_D05956g) bears similarity to known heat shock transcription factors (Pan and Heitman 2002; Patterson et al. 2018) ; *SNF2* (YALI1_D30097g) is a chromatin remodeling protein in the SWI/SNF transcription complex (Hirschhorn et al. 1992); and *RHO5* (YALI1_E30639g) is involved in cell integrity, helping to propagate heat and oxidative stress signals leading to induced cell death. The final gene identified in our screen was (YALI_F04690g), whose encoded protein had a 40.9% uniprot BLAST identity score match with the Ptip protein (part of the histone H3K4 methyltransferase complex) in *Drosophila melanogaster* and is referred to here as *PAXIP1* (Fang et al. 2009). Each of these genes are either putatively gene expression regulators regulation, or putatively related to a signal cascade that may affect the hyphal transition (**Figure 1.1**)

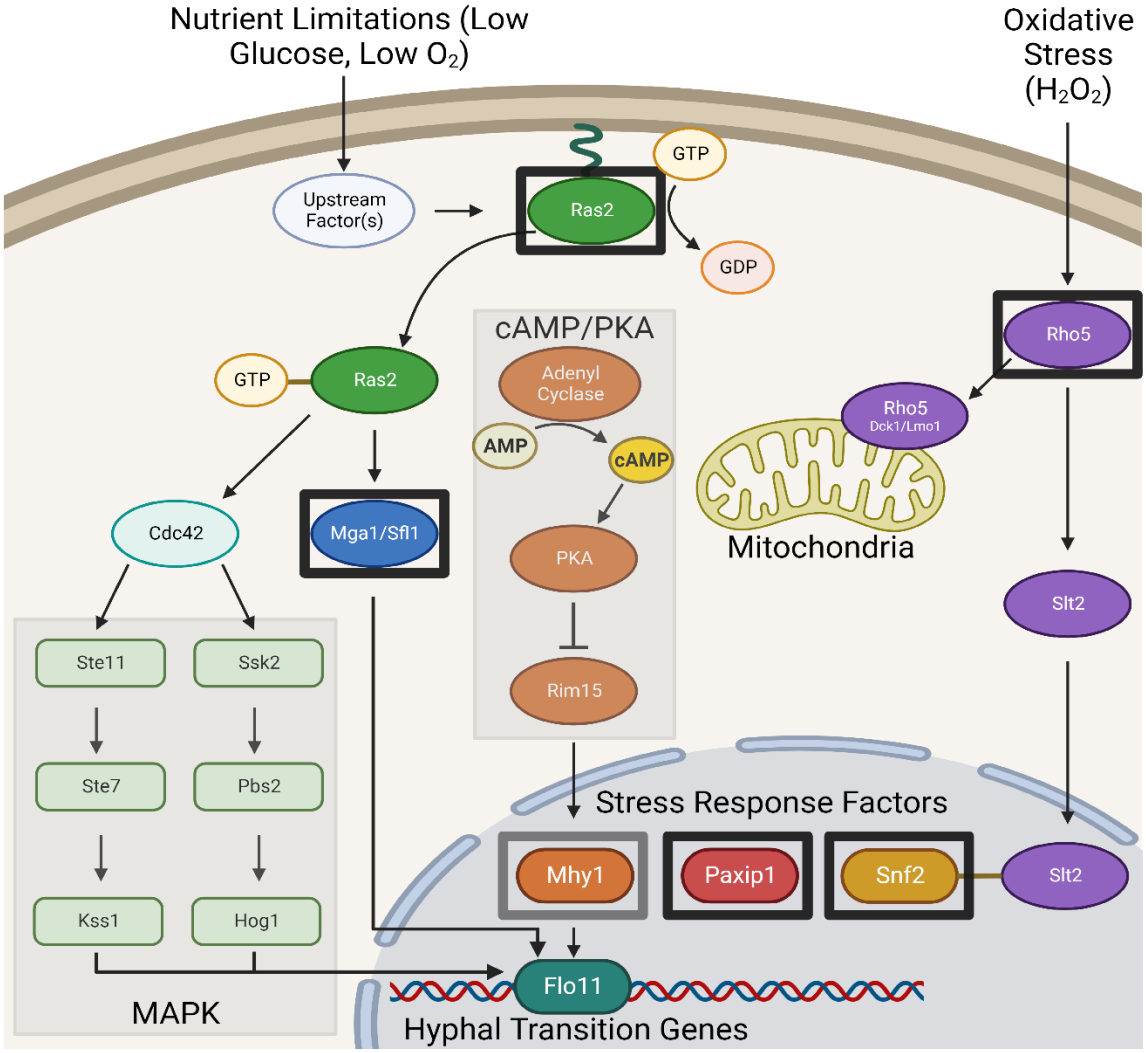


Figure 1.1: Putative roles of knockout hits in the hyphal morphology transition. Null-hyphal knockout hits from the genome wide screen are highlighted by black rectangles, and literature-derived null-hyphal knockouts are highlighted by a gray rectangle.

1.4.2 Growth characteristics and morphological phenotypes of Null-Hyphal Hits

A goal of this study was to identify one or more genetic knockouts that eliminate or suppress hyphal formation and that have minimal or no effect on the yeast's ability to perform as a biochemical production host. In addition to the resulting increase in time

and resources required for a production run, a reduced growth rate can signify other metabolic burdens which may compromise the synthesis of the desired product. As such, we measured growth rates of the mutant strains (**Figure 1.2**) and observed that the ΔRAS2, ΔRHO5, and ΔSFL1 strains had similar growth rates to unmodified PO1f, thus leaving these mutants as potential host candidates. The ΔSNF2 and ΔPAXIP1 strains, however, showed impeded growth, ruling out their use as potential null-hyphal hosts for industrial use. In addition to these mutants, we also characterized a ΔMHY1 (YALI_B28150g) strain; MHY1 functions downstream of RAS2, disruption of which has been shown to reduce hyphal formation without a reduced growth phenotype (Morgunov et al. 2004; Konzock and Norbeck 2020). Given these results, the ΔRAS2, ΔRHO5, ΔSFL1, and ΔMHY1 knockout strains were selected for further investigation.

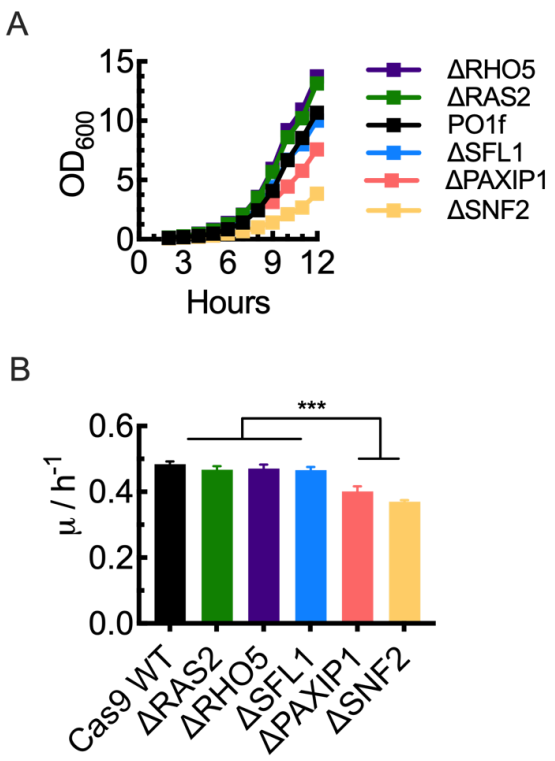


Figure 1.2: The effect of putative hyphal knockout on growth rate. (**a**) Time course of cell growth as measured by the optical density at 600 nm (OD_{600}). *Y. lipolytica* cultures were grown in shake flask cultures with YPD media. (**b**) Growth rate, μ, calculated from the slopes of the natural logarithms of the growth curve measurements in part a. Data points and bars represent the mean of triplicate measurements. Error bars represent the standard deviation. *** represents $p < 0.001$ from Dunnett's multiple comparisons test, post-hoc of a one-way ANOVA.

With the growth rate of the hyphal knockout strains characterized, we proceeded to test the knockouts that had no effect on growth rate in a lycopene overproduction host. Previously, we reported a series of genetic manipulations to *Y. lipolytica* that introduce and enhance a lycopene biosynthesis pathway (Schwartz et al. 2017a; Schwartz et al. 2017c). These manipulations include the overexpression of a series of bacterial enzymes – CrtE, CrtB, and CrtI – that convert farnesyl pyrophosphate into lycopene, and the homologous overexpression of HMG1 and MVD1, both of which are known to increase mevalonate pathway flux to isopentenyl diphosphate (IPP), a precursor to lycopene

biosynthesis. The lycopene production strain was designated as PO1f-HMEBI with each overexpression represented by H (HMG1), M (MVD1), E (CrtE), B (CrtB), and I (CrtI). Using PO1f-HMEBI as the parent strain, we generated the four most viable null-hyphal gene deletion candidates, ΔRAS2, ΔRHO5, ΔSFL1 and ΔMHY1, and characterized the percentage of cells that underwent a transition from yeast to hyphal morphology over a ten day culture, a time course selected based on our experience in producing lycopene in engineered strains over a similar time period. At each time point, we withdrew an aliquot from two independent cultures, mounted the live samples on microscope slides, and generated six images per slide for a total of twelve images per strain per time point (examples provided in **Supplemental Figure S1.1**). We defined any cells with a length greater than twice the widest point as hyphal/pseudohyphal for the purposes of identifying all cells transitioning from the yeast state. Examples of both hyphal and pseudohyphal structures meeting this criteria are shown in **Figure 1.3**.

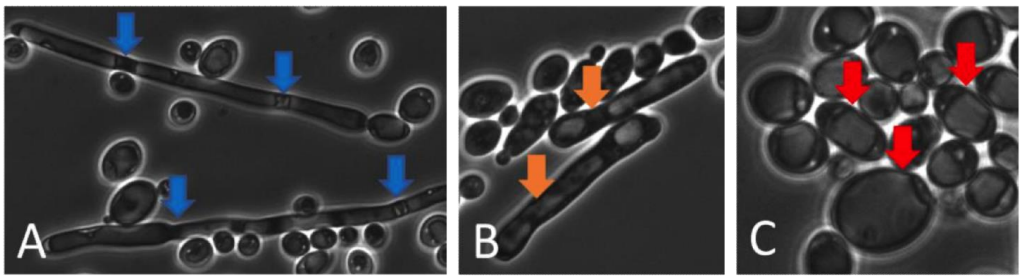


Figure 1.3: Images of hyphal, pseudohyphal, and budding phenotypes in *Yarrowia lipolytica.* (**A**) Image of the hyphal state, with blue arrows indicating septa between cells, which are characteristic of hyphae. (**B**) Image of the pseudohyphal state, with orange arrows indicating elongated cells with signs of incomplete division. (**C**) Image of the budding state, with red arrows pointing to unbranched and rounded yeast cells.

Both PO1f and PO1f-HMEBI showed hyphal transition behavior in approximately 15% of their cells across the full time series, establishing a baseline occurrence rate (**Figure**

1.4). ΔRHO5 proved to be the least effective knockout with an ~3% hyphal occurrence after one day of culture, but jumping to ~20% after 3 days of growth and increasing to >28% after day nine. Disruption of SFL1 also resulted in early suppression of (pseudo)hyphal formation, maintaining ~5% elongated structures for six days of culture, increasing to >23% at day nine and ten. ΔMHY1 began with ~3% hyphal occurrence, but steadily increased over subsequent days, reaching nearly 30% on day ten. Unlike the other knockouts, ΔRAS2 had low (<5%) hyphal occurrence across all ten days, without the rebound of hyphal behavior noted in all of the other knockouts. We therefore identified ΔRAS2 as the most promising benign and non-transient hyphal knockout generated.

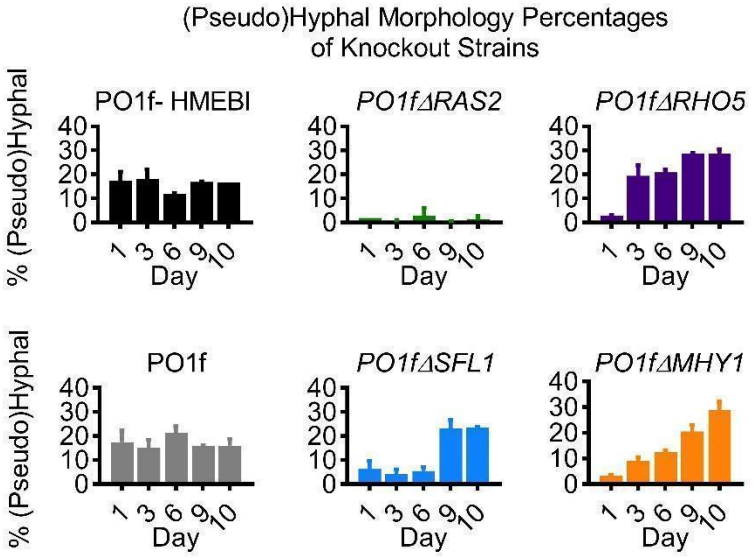


Figure 1.4: Hyphal morphology percentages of PO1f, PO1f-HMEBI, and putative hyphal knockout strains. The knockout strains include PO1f-HMEBI ΔRAS2, PO1f-HMEBI ΔRHO5, PO1f-HMEBI ΔSFL1, and PO1f-HMEBI ΔMHY1. Bars represent the average number of cells in a hyphal or pseudohyphal state observed in 12 different images across two biological replicates. At least 250 cells were characterized at each time point for each strain. Error bars represent the standard deviation across the 12 different images.

One explanation for the lasting effect of the RAS2 disruption is that Ras2 is an early control node for the transition to hyphal morphology, existing upstream of several hyphae-regulating pathways. Primarily responsive to glucose conditions outside the cell, Ras2 is a plasma-membrane GDP binding protein until being activated by Cdc25. Once activated, Ras2 has a broad cascade of interactions that lead to various stress responses through cyclic AMP and PKA as well as through the rho-like GTPase Cdc42, which in turn activates the MAPK pathway (Li et al. 2014; Mösch et al. 1996). The signal pathways modulated by the interaction of Ras2 and Cdc25 have been implicated in a wide range of downstream cellular processes, including transition to a filamentous state (Mösch and Fink 1997). Ras2 is involved directly in sensing glucose in the environment and receives regulatory feedback from glucose metabolism through activation of Cdc25 by fructose-1-6-bisphosphate. Commonly acting to down-regulate stress response genes, RAS2 knockouts display a heightened basal stress tolerance in *S. cerevisiae*, which is another industrially useful trait (Zacharioudakis et al. 2017; Shama et al. 1998). In this anticipatory state RAS2 knockouts have shown greater thermotolerance and resistance to oxidative stress, while also exhibiting life span extension akin to that seen in response to calorie restriction. The ΔRAS2 mutant is therefore a putative null hypal strain without other observed negative traits and some potentially unexpected potential benefits.

1.4.3 Validation of PO1f ΔRAS2 mutant

In order to ensure that Ras2 functions as a control node for hyphae formation, we conducted a rescue experiment with RAS2 expressed from an episomal plasmid in a PO1f ΔRAS2 strain. As shown in **Figure 1.5**, the rescued ΔRAS2 cells regained the rough colony phenotype present in PO1f cells, as well as cellular filaments when viewed under a microscope. This indicates that the hyphae-forming wild type phenotype was restored in tandem with RAS2 expression, even as the ΔRAS2 cells retained the null hyphal smooth colony phenotype. As such, we confirmed RAS2 as a necessary gene for hyphal formation, where its presence has a direct relationship to the hyphal forming phenotype of *Y. lipolytica*.

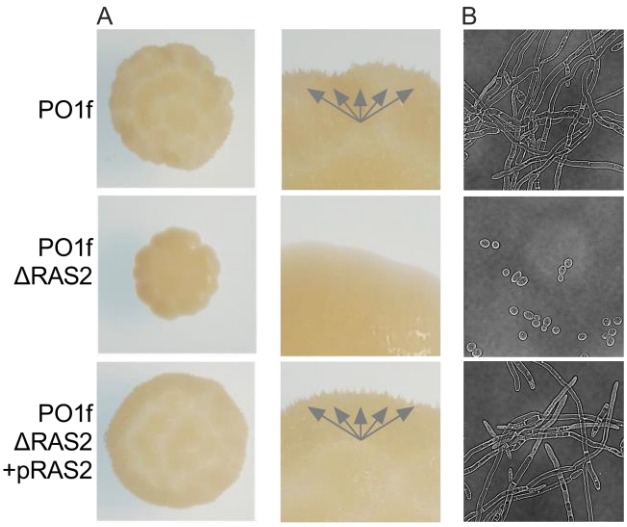


Figure 1.5: ΔRAS2 rescue assay. Rows represent different experimental groups, with the left column showing the full cell spot, the center column showing a zoomed in image of the top edge of the same spot, and the right column showing microscopy images (60x) of cells resuspended in liquid media. Both the wild type PO1f cells and the rescued ΔRAS2 cells produced hyphal morphology with characteristic rough edges and hyphal structures, while the ΔRAS2 knockout strain produced colonies with a smooth phenotype and round cells, indicative of a loss of the hyphal structures. Gray arrows indicate the rough or spiked edge indicative of hyphae.

1.4.4 Lycopene Production with PO1fΔRAS2 mutant

A primary goal of this study was to identify genetic manipulations that reduce or eliminate hyphal formation in a production host. To this end, we generated ΔRAS2, ΔRHO5, ΔSFL1 and ΔMHY1 mutant strains in a PO1f-HMEBI background and measured the effect of each knockout on lycopene production (**Figure 1.6**). As expected, the wild type HMEBI strain produced 2.8 mg of lycopene per gram of dry cell weight (mg/gDCW) by day 3, and 4.9 or more mg lycopene/gDCW on days 6 and 8, results equivalent to those previously reported for this strain (Schwartz et al. 2017a). The ΔRHO5, ΔSFL1, and ΔMHY1 strains produced comparatively little lycopene, reaching 1.4 mg lycopene/gDCW or less by the end of an 8 day culture. The ΔRAS2 strain, however, produced similar amounts of lycopene to the HMEBI strain, at 1.8 mg lycopene/gDCW on day 3, around 4.0 mg lycopene/gDCW on day 6, and 5.1 mg lycopene/gDCW on day 8. It therefore functioned just as well as HMEBI as a lycopene production platform despite lacking the lycopene production enhancements of the HMEBI strain.

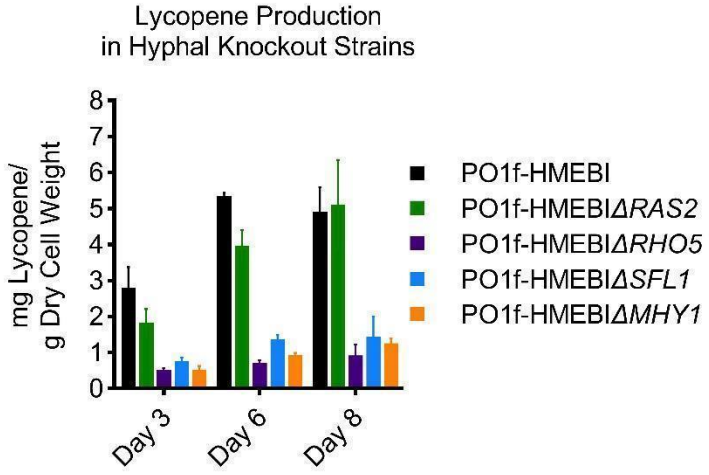


Figure 1.6: The effect of hyphal knockout on lycopene in *Y. lipolytica*. The HMEBI lycopene production strain and HMEBI ΔRAS2 strain consistently produced two to four times as much lycopene as the HMEBI ΔRHO5, HMEBI ΔSFL1, and HMEBI ΔMHY1 strains.

While the RAS2 knockout did not produce an increase in lycopene titer at the benchtop scale, the potential industrial benefits of its use cannot be overlooked. By eliminating hyphae formation in a fully functional production strain, associated labor costs and malfunctions from bioreactor fouling may be avoided, and cells that would have otherwise ended up on the bioreactor vessel could instead be harvested for more product. Likewise, the challenges of oxygen and nutrient diffusion due to the presence of hyphae is inherently worse in the large volumes of industrial bioreactors, where oxygen and nutrient diffusion is already a major challenge. It is even possible that an increased product titer could still be gained from a null-hyphal strain when grown at the scale of an industrial bioreactor through the reduction of low oxygen and nutrient regions in the vessel. Moreover, the known role of RAS2 in the hyphal transition of *S. cerevisiae* and *Candida albicans* (Mösch and Fink 1997; Parveen et al. 2019; Chow et al. 2019) speaks

to the potential utility of this mutation for bioprocessing with other hosts. While more work must be done to quantify any theoretical titer increases from the ΔRAS2 at an industrial scale, the practical value of a fully functional but lower maintenance strain should not be underestimated.

1.5 Chapter 1 Conclusion

These results illustrate the utility of our CRISPR-Cas9 knockout library screen for identifying desirable phenotypes. Through a simple visual screen, we were able to rapidly identify and subsequently characterize a benign and industrially useful single gene knockout. Indeed, many industrially beneficial mutations have been and will continue to be identified through screens of whole genome knockout libraries. Furthermore, the knowledge we gain about what genes serve as effective regulatory and control points for our goals will inform future engineering in the next generation of bioproduction strains.

Chapter 2: Insights from Genome wide Knockout CRISPR-Cas9 Screens in Solid and Liquid Media for Identifying Growth-enhancing Candidate Genes in *Yarrowia lipolytica*

2.1 Chapter 2 Abstract

Solid state fermentation (SSF) is an increasingly important area of research in industrial bioprocessing and strain engineering. *Yarrowia lipolytica* is well suited to SSF due to its morphological characteristics and tolerance to lower moisture environments. However, ongoing difficulties persist with using *Y. lipolytica* in SSF, such as oxygen and temperature control, and large scale SSF agitation. Engineering *Y. lipolytica* strains to address these difficulties is challenging, as there is a knowledge gap on *Y. lipolytica* genes that are essential or detrimental to growth on solid media. In order to identify genes and cellular systems of importance to growth on solid media, we conducted a whole-genome knockout CRISPR screen, comparing growth of *Y. lipolytica* on solid and liquid agarose media. We identified a few genes, such as F09433, involved with vacuole autophagy, that were advantageous knockouts on solid media. We also utilized gene ontology (GO) terms to elucidate types of gene knockouts with greater essentiality on solid media compared to liquid media, such as endomembrane and mitochondrial function genes. The results of this screen provide a starting point to derive a more complete understanding of the genetic underpinnings of the differences between solid and liquid media growth in *Y. lipolytica.*

2.2 Chapter 2 Introduction

Solid state fermentation (SSF) processes are a unique form of bioprocessing which uses a moist solid scaffold and nutrient source to grow a given microorganism, in lieu of cells being suspended in liquid media (Lizardi-Jiménez and Hernández-Martíne...). Typically, filamentous fungi and dimorphic yeasts (such as *Aspergillius niger* and *Yarrowia lipoytica*) are best suited to growth on solid substrates, as they have lower moisture requirements than bacteria and a wide range of viable solid substrates (Thomas et al. 2013; Pal and Khanum 2010). SSF has been used to prepare fermented foods for millennia, as the process could increase nutrition and prevent spoilage (Paredes-López and Harry 1988). More recently, SSF has been adopted for several biotechnological applications, including organic acid production (Mao et al. 2020), enzyme production (Wang et al. 2019; Abdullah et al. 2015; El-Naggar et al. 2015), industrial waste processing/product conversion (Cerda et al. 2019; Leite et al. 2021; El Sheikha and Ray 2023), and environmental remediation/toxin degradation (Asemoloye and Marchisio 2022; Zinjarde et al. 2014). As SSF has become more widespread in bioprocessing, it has proven to be ideal for a number of applications.

Liquid/submerged bioproduction has been a staple in industry for decades, but SSF brings several unique advantages. It better mimics the natural habitat of several industrially relevant molds and fungi, (which evolved to grow on solid substrates), taking advantage of the higher concentrations of nutrients in solid substrate to increase bioproduction activity (Farinas 2015; Lizardi-Jiménez and Hernández-Martíne...) and produce more concentrated product titers (Rodríguez Couto and Sanromán 2005;

Viniegra-González et al. 2003; Khanahmadi et al. 2006). This is likely due to a lack of catabolite repression (the suppression of many desired metabolites in the presence of abundant nutrients), which is notably reduced in SSF (Thomas et al. 2013; Viniegra-González and Favela-Torres 2006). SSF is also well suited for bioremediation tasks, with some SSF species having high tolerance for certain pollutants and otherwise dangerous compounds (Ziganshin et al. 2010), as well as their ability to produce high concentrations of biosurfactants and enzymes required for degrading the contaminants (at the direct physical interface with the targeted compounds) (Martins et al. 2009). For example, *Yarrowia lipolytica* has demonstrated potential for removing fuel pollution from soil (Żogała et al. 2005). By extension, many agricultural and industrial wastes and byproducts may be either utilized as a SSF growth substrate (Orzua et al. 2009), or converted into valuable products through SSF methods (Vastrad and Neelagund 2012). There are also substantial economic benefits to using SSF. The lack of free water reduces contamination risk and energy input required for mixing (Karimi et al. 2021). And by replacing costly growth media and substrate with agricultural and industrial waste, a major expense of bioproduction may be greatly reduced (Hölker et al. 2004). Overall, there are many applications where SSF may prove advantageous, especially given its inherent advantages.

Even with its upsides, SSF has a number of technical challenges. Since SSF bioreactors consist of solid and gas phases (with varying moisture in the liquid phase), several complications can arise. One of the most prevalent difficulties is the management and control of oxygen and CO_2 gradients. Extreme oxygen gradients may form across the

growth substrate, (often due to heterogeneous substrate particle size), altering cellular metabolic activity (Chilakamarry et al. 2022; Oostra et al. 2001). Another common difficulty with SSF is temperature control and management. The low specific heat of air and the relatively low moisture content in the substrate inhibit efficient heat transfer in SSF bioreactors, leading to unacceptable temperature gradients and difficulty lowering bioreactor temperature (Raghavarao et al. 2003). Both of these difficulties become even more pronounced when scaling up SSF. As SSF bioreactor scale increases, additional mechanisms such as forced aeration and substrate mixing must be employed to address gas and heat exchange (Durand 2003; Mitchell et al. 2000). While several bioreactor designs are able to aerate and agitate solid substrates, the process is energy intensive (Mitchell et al. 2006) and may damage the microorganisms being used (Manan and Webb 2017). Addressing these issues is difficult across multiple bioreactor types, but selecting the correct type of microorganism can minimize the design challenges of SSF applications.

Yarrowia lipolytica is able to grow effectively in and on both hydrophilic and hydrophobic substrates, enabling a wider range of nutrient sources and products (Nascimento et al. 2022). Viable growth substrates include waste products such as cooking oil (Xiaoyan et al. 2017), and oil cakes (Imandi et al. 2013). It can even grow in the presence of toxic organic solvents (Walker et al. 2019). It is also dimorphic, able to transition between yeast and hyphal form depending on environmental conditions, enabling its hyphal form to predominate on solid substrates (de Souza et al. 2019). This dimorphism also enables a changing gradient of hyphal cells across the growth mass on

solid media, which optimizes oxygen utilization (Rahardjo et al. 2002). Overall, *Y. lipolytica* is well suited to utilization for solid state bioprocessing. However, the many difficulties in engineering effective SSF bioprocesses necessitate further *Y. lipolytica* strain optimization for improved growth on solid substrate.

Compared to liquid bioprocesses, fewer advancements have been made toward engineering *Y. lipolytica* strains for a solid-state bioproduction environment. To address this deficit, we may use a genome wide CRISPR knockout screen to identify both beneficial knockouts and essential genes for a solid media environment. There are precedents for screening fungi strains on solid media for various purposes. For example, several strains of the filamentous fungus *Aspergillus niger* were isolated from the environment and grown on a solid scaffold to measure secreted xylanase activity over time (Abdullah et al. 2015). Additionally, a SSF screen using agar embedded with cellulose (to be tested as an alternative carbon feedstock) was also carried out using individually plated *Trichoderma reesei* strains, where local cellulose consumption could be directly measured (Florencio et al. 2012). An especially relevant example was an alcohol sensitivity screen, carried out on a *Saccharomyces cerevisiae* gene deletion library whose members were individually plated on agar plates containing different types of alcohols (Fujita et al. 2006). While this example demonstrates the theoretical viability of a SSF screen for a knockout library, it highlights limitations as well. Through multiple iterations, Fujita et al tested 4500 knockout strains, requiring the individual generation, isolation, and testing of each one. Not only is such an approach highly labor intensive, its

results are less quantitative than a pooled CRISPR knockout screen, preventing precise quantitative rankings and comparisons of knockout environmental sensitivities.

Given the size of our genome wide knockout library of just under 25,000, individual mutant screening is not a realistic approach. Instead, we must use a pooled library approach, in which the entire library is plated on solid media, grown, and harvested into a pooled sample to be analyzed with quantitative high throughput sequencing. In order to identify potential engineering targets for solid state fermentation applications of *Yarrowia lipolytica*, we used a refined sgRNA guide library to identify essential genes and beneficial gene knockouts with respect to growth in both solid and liquid glucose minimal media, and analyzed existing literature to elucidate functionality of both types of hits.

2.3 Chapter 2 Materials and Methods

2.3.1. Microbial strains and culturing

All strains used in this work are presented in **Supplemental 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.

2.3.2. Plasmid construction

All plasmids and primers used in this work are listed in **Supplemental Table S2** and **Supplemental Table S3**, respectively. The plasmids used to knock out genes in *Y. lipolytica* PO1f were constructed by ordering the corresponding sgRNA as a primer (**Supplemental Table S3**) 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 molar ratio (insert:vector). The assembly product was then transformed directly into electrocompetent *E. coli* TOP10 cells to eventually be propagated and harvested by Miniprep.

2.3.3. *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. (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 hours (final OD ~30). 300 µL of culture (~10^8 cells) were pelleted by centrifugation at 4,000 g for 2 minutes 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 hour 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.

2.3.4. 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., 2019a), while the third guide was designed by DeepGuide predictions (Baisya et al., 2022; Schwartz et al., 2019a). First, sgRNAs with a cutting score (CS) greater than 4 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.

2.3.5. 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 kit (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 TOP10 *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 DpnI 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 hr 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 hr, 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.

2.3.6. *Y. lipolytica* Library Transformation

Transformation of the sgRNA plasmid libraries into PO1f and PO1f Cas9::A08 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 hours (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 hour. 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 hours 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 x 10^6 total transformants per biological replicate).

2.3.7. *Y. lipolytica* solid/liquid glucose screen

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 hours at 225 RPM, 30°C. Cells were recovered as a pellet via centrifugation (at 4000 g for 5 minutes). 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.

25 mL of liquid glucose minimal media (see Turki et al., 2009) cultures were seeded with washed cells up to an OD600 of 0.3. Two biological replicates were cultured for each strain, resulting in 4 flasks. Cells were then grown at 30 °C at 220 RPM. The experiment was halted after 42 hours of growth, with the OD600 of all flasks ranging from 8 to 12.

Optimal optical density for plating on glucose was calculated to be: OD600 0.05 (See section 2.3.12). For each plate formulation replicate, 100 µL of the washed cells, diluted to OD600 0.05, were plated onto 16 pre-prepared plates (see section 2.3.11) using sterile beads (16 total plates per replicate or 64 plates total). Plated cells were allowed to dry in a laminar air flow hood for 1 hour, then they were grown in a static 30°C incubator for 42 hours (see section 2.3.12 for endpoint calculation). Upon removal from the incubator, plates were imaged in the BioRad imager and analyzed to ensure colonies were discrete and representative of individual transformants.

The entire volume of each liquid culture replicate was centrifuged at 4000 g for 10 minutes, with the media aspirated off of the pellet. Then, each replicates' cells were resuspended in 3 mL of sterile water, split into three 1 mL tubes, re-pelleted and aspirated, and frozen at -80°C until ready to process further. Solid media cells were harvested in a two-part process. First, as much of the colony mass on the plate surface as possible was removed from each plate with 15 mL of Tween 80 buffer (0.01% Tween 80, 0.11 M KH_2PO_4, 0.03 M K_2HPO_4) 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 minutes 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 minutes to melt the media and free the cell mass. The vessels were then centrifuged at 4000 g for 5

minutes. 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.

2.3.8. Solid Liquid Screen Library isolation and sequencing

Frozen pelleted culture samples from the solid and liquid cultures were thawed and resuspended in 400 µL sterile, Ultrapure H_2O. 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 copy 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 10^7 plasmids so that sufficient coverage of the sgRNA library is ensured.

At least 0.2 ng of plasmids (approximately $3x10^7$ 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 Cas9 screens were prepared as previously described in Schwartz et al., 2019. Briefly, isolated plasmids were amplified using forward (Cr1667-Cr1669) and reverse primers (Cr1669, Cr1670, Cr1672, and Cr1710) containing the necessary barcodes, pseudo-barcodes, and adapters (**Supplemental Table S3**). Approximately $1x10^7$ 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 magnetic bead 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.

2.3.9. 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 v2..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 **Supplemental Table S4**, and the demultiplexing primers/barcodes are provided in **Supplemental Table S5**

2.3.10. 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). Read counts from the Cas9-containing strain sampled before inoculation into solid and liquid media conditions 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 genes beneficial and detrimental to growth in the solid and liquid glucose growth conditions, the *significance* parameter was set to 'two-tailed'. In all cases, genes having FDR-corrected p-value < 0.05 were deemed as significant.

2.3.11. Solid and liquid glucose media preparation

Liquid glucose minimal media was prepared in 1 L batches. First, 500 mL of 2x concentration minimal media was prepared via autoclaving (for each bottle, 1 g $MgSO_4$ * $7H_2O$, 4 g $(NH_4)_2$, 10 mg $FeCl_3$ * $6H_2O$, 200 mg Uracil, 4 µg Myo-Inositol, 8 µg Biotin, 200 µg Thiamine HCl, 15 g KH_2PO_4, and 5.5 g K_2HPO_4). Simultaneously, a 500 mL

mixture of ultrapure water and 20 g D-glucose was also autoclaved. The glucose-water mixture was added to the 2x minimal media mixture, briefly mixed on a stir plate, and allowed to cool to room temperature.

Solid medium glucose plates were prepared in batches of 50 (20 mL of media per 100 mm petri dish). First, 500 mL of 2x concentration low melt agarose minimal media was prepared via autoclaving (for each bottle, 1 g $MgSO_4$ * $7H_2O$, 4 g $(NH_4)_2$, 10 mg $FeCl_3$ * $6H_2O$, 15 g LMP Agarose, 200 mg Uracil, 4 µg Myo-Inositol, 8 µg Biotin, 200 µg Thiamine HCl, 15 g KH_2PO_4, and 5.5 g K_2HPO_4), and kept molten in a 70°C water bath until ready for use. Meanwhile, a 500 mL mixture of ultrapure water and 20 g D-glucose was autoclaved and kept heated to approximately 70°C on a heated stir plate. The glucose-water mixture was added to the 2x low melt agarose minimal media mix, and 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 hours and then refrigerated until needed, up to 3 days.

2.3.12. *Y. lipolytica* Glucose Culturing Optimization

Prior to selection screens, optimized *Y. lipolytica* cell plating densities were calculated on solid medium glucose plates. Twelve hours prior to assay start, four glycerol stocks of prepared library cells (see methods 2.2.7), 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 hours at 225 RPM, 30°C. Cells were recovered as a pellet via centrifugation (at 4000 g for 5 minutes). 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 OD-600 absorbances of 0.4, 0.2, 0.1, and 0.05. Then, 100 µL of each cell dilution was plated on a glucose minimal media plate (see methods 2.2.11) using sterile beads. Plated cells were allowed to dry in a laminar air flow hood for 1 hour, then they were grown in a static 30°C incubator for two days. Starting at 24 hours, and every 24 hours 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, adjusted downwards by up to 6 hours if colonies had begun to merge. 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.

2.4 Chapter 2 Results and Discussion

2.4.1 Solid and Liquid State Screening Identifies genes involved with growth on different phased substrates

In order to identify genes essential for growth on solid and liquid media, we used the acCRISPR pipeline (Ramesh et al. 2023) to calculate significant fitness score (FS) outliers, with negative outliers representing essential genes for the given condition, and positive outliers representing gene knockouts that convey a fitness advantage. We conducted functional screens using 2% glucose minimal media, with library cells harvested prior to seeding (as a control) and after 42 hours of growth. We were thus able to calculate FSs for all of the library's single knockouts grown in a shaken liquid culture, as well as a static solid culture (**Figure 2.1a**). The FS values represent the $\log_2$-ratio of the ratio of recovered gRNA abundances between a given growth condition and the pre-treatment time 0 sample, from which FS values for genes may be calculated. The FS values of the knockout strains in the two environments have a wide distribution. The acCRISPR pipeline was used to correct our FS values for the activities of the guides used to generate them, enabling us to calculate the statistically significant positive and negative FS outliers, signifying advantageous knockouts and essential genes, respectively (**Fig 2.1b**). It was noted that there was a wider range of FS outlier magnitude for solid media than for liquid media, especially for negative FS values. While the number of essential gene hits for solid and liquid media were almost identical each other, (563 solid essential hits and 559 liquid essential hits) the solid media essential gene hits had negative FS values with magnitudes 4 times greater than the liquid media hits.

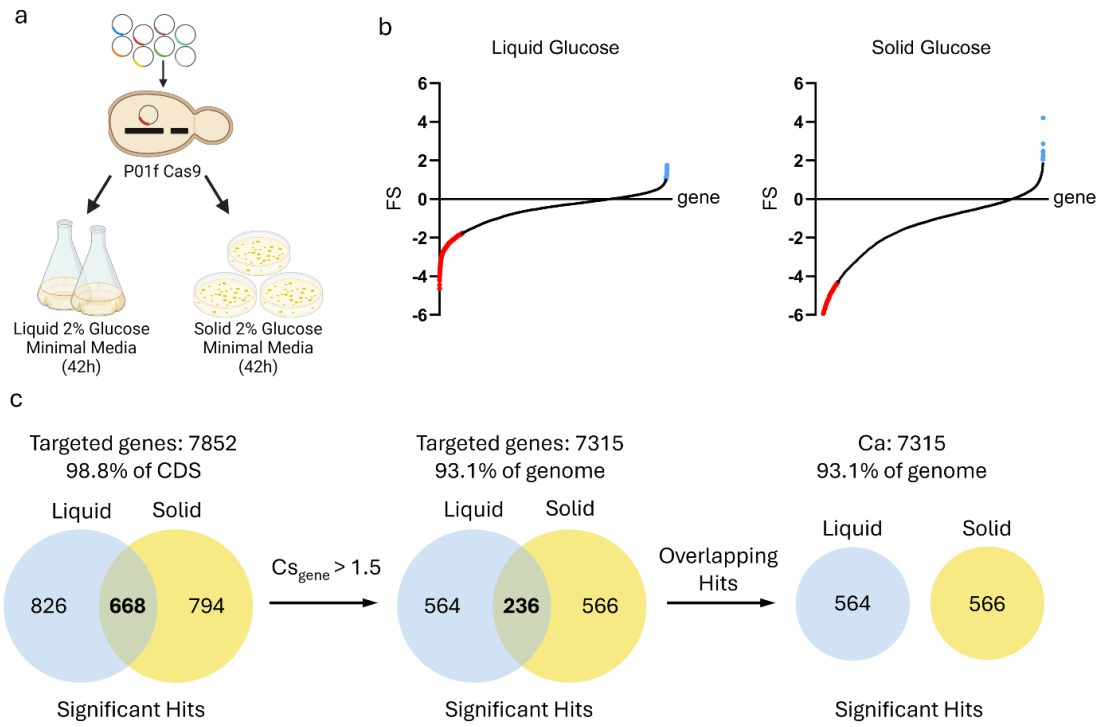


Figure 2.1. Solid vs Liquid glucose functional genomic CRISPR screening. (a) Growth screens were conducted using *Y. lipolytica* PO1f Cas9 with time 0 h and time 42 h control and treatment conditions respectively in two media conditions – liquid 2% (w/v) glucose minimal media and solid 2% (w/v) glucose minimal media. Fitness score (FS) is calculated as the average enrichment of guides targeting a gene after 42 hours of growth compared to 0 hours of growth. Genes with a higher FS indicate that the knockout gave a fitness advantage. (b) Corrected FS values for all genes in the solid and liquid glucose screens, as determined from the acCRISPR pipeline. Red portions of the curves represent essential genes for glucose while blue portions of the curves represent knockouts with a fitness advantage (c) Venn diagrams representing gene hits in solid and liquid conditions after each filtering step. Low-activity guides and glucose essential genes were removed to condense the pool of hits.

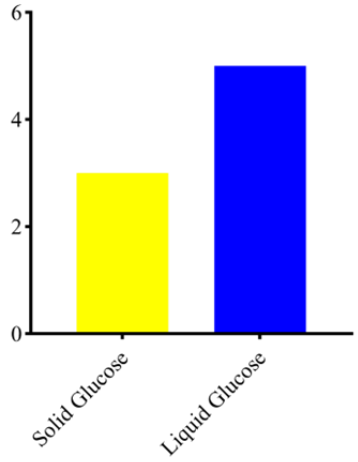
With 7852 targeted genes in the library (98.8% coverage), a total of 826 significant outliers were found in the liquid condition, and 794 significant outliers were found from the solid sample. There were also 668 significant outliers that occurred under both conditions. Still, in order to reduce the risk of false positives, we removed hits derived from gRNAs with CS < 1.5, leaving us with 7315 targeted genes (93.1% of the genome)

and 564 significant liquid hits, 566 significant solid hits, and 236 hits from both categories. Hits that occur in both solid and liquid conditions are likely to be either essential genes or advantageous knockouts regardless of media state, so we chose to focus on solid or liquid exclusive hits to elucidate further knowledge about growth on different types of substrate (**Fig 2.1c**).

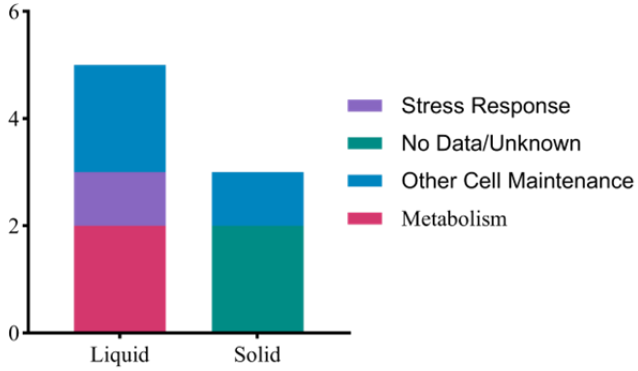
2.4.2 Solid and Liquid State essential gene and beneficial hits identify knockouts that favor solid or liquid conditions

We then quantified the significant hits as GOF and LOF hits from the CS-filtered unique solid and liquid hits and sorted the deleted genes by putative cellular function. There were 3 unique solid hits with a fitness advantage, and 5 unique liquid hits with a fitness advantage (**Figure 2.2a**). The low number of advantageous knockouts is not surprising, as single gene deletions rarely convey a measurable benefit for cells under unique growth conditions. The number of hits was further reduced by our removal of low cutting score guides, eliminating many likely false positive advantageous hits (30 hits were trimmed from the liquid fitness advantage hits, and 3 hits were trimmed from the solid fitness advantage hits). The few advantageous knockout hits that we found were subjected to literature-based functional analysis, beginning with utilizing the *Yarrowia lipolytica* CLIB89 (YALI1) annotated genome (Magnan et al. 2016) to identify gene homologies with other yeasts and fungi, followed up with BLAST and paper-BLAST searches. We used the existing annotations and literature on gene homologues to manually classify the advantageous knockouts into broad cellular function categories.

a



b



c

Liquid Fitness Advantage Hits		
Gene	**Homology**	**Description**
B27723g	s.c. ART2	Arrestin-like protein
C08882g	s.c. MIX23	Mitochondrial membrane import regulator
C20224g	s.c. YJU3	fatty acid ethyl ester catabolism
E02000g	s.c. PLM2	promoters G1/S transition
E36098g	s.c. TRM1	tRNA modification enzyme
Solid Fitness Advantage Hits		
Gene	**Homology**	**Description**
A09325g	N/A	Unknown
E37234g	N/A	Unknown
F09433g	s.c. ATG15, s.p ATG15, a.n. ATG15, n.c. ATG15	Lipase involved with vacuole targeted body regulation and lysis

Figure 2.2. Biological Function of top Solid vs Liquid GOF screening hits. (a) Post hit filtering of improved fitness hit counts for liquid and solid. (b) Liquid and solid advantageous knockout hit distribution by biological function: Fatty acid/sterol metabolism, stress response and tolerance, cell maintenance, and unknown. (c) Analysis of liquid and solid GOF hits, their homologies, and functions.

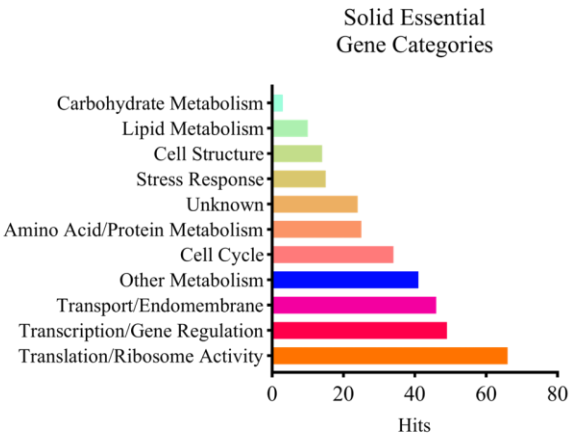
We then grouped hits by broad gene function categories derived from literature review. Due to existing literature tying hyphal formation to improved growth on solid substrate (Rahardjo et al. 2002), we include genes with known relationships to cellular morphology as their own category. Functional analysis of the fitness advantage hits indicated two unknown genes with no obvious homologies for solid media, and one gene classified under cellular maintenance (**Figure 2.2b**). This gene, *F09433g*, has homology with ATG15 in *S. cerevisiae*, and *atg15 S. Pombe*, *A. niger*, and *N. crassa* (**Figure 2.2c**). All of them code for a phospholipase vital for vacuole autophagy (Mukaiyama et al. 2009; Nitsche et al. 2012). One putative reason for an autophagy-suppression knockout to provide a fitness advantage on solid media might be a release of more resources to go towards cell division and growth, if autophagy wasn't as important on solid media in the presence of glucose, though the use of minimal media and cannot be ruled out as a contributing factor.

The fitness advantage hits for liquid cultures were more widespread. Two each of cell maintenance and metabolism genes were found, as well as a stress response gene. The metabolism related genes were E36098g, a tRNA methyltransferase (Liu et al. 1998), and C20224g, which degrades monoacylglycerol (Heier et al. 2016). The cell maintenance genes were C08882g, a mitochondrial membrane import regulator (Zöller et al. 2020) and E02000g, a G1/S cell cycle regulator (Horak et al. 2002). The last hit was an arrestin-like

protein, involved with endomembrane transport under stress or nutrient change conditions (Baile et al. 2019). Further experiments are needed to determine why deleting these genes may provide a fitness advantage in liquid glucose minimal media.

In order to gain further insights on genes essential to growth in a solid media environment, we analyzed essential gene hits for growth on solid media. As there were many hits across a wide range of different categories, we organized them into a broader grouping of categories to classify the types of essential genes for solid media growth (**Figure 2.3a**) by using the existing YALI1 genome annotations to find homologous genes via paperBLAST (Magnan et al. 2016). Genes related to translation, ribosome biogenesis, and ribosome activity were the most common solid media essential gene category, with transcriptional regulation, endomembrane/transport, and miscellaneous metabolism making up other highly prevalent categories.

a



b

Representative Solid Essential Gene Hits			
Gene	**Homology**	**Description**	**Source**
F30254g	*C.a. cdc39*	Pro-hyphal transcription factor	Uhl et al, 2003
F37890g	*S.c.* HCR1 *N.c.* 07954	eIF3 translation factor, deletion reduces biofilm formation in *S.c.*	Valásek et al, 1999
D20061g	*S.c.* ABZ2	Folic acid biosynthesis	Botet et a, 2007
C26544g	*S.c.* PET8	Mitochondrial transmembrane transporter	Lalo et al, 1994
B20242g	*S.c.* CDC16	Component of anaphase promoting complex for cell cycle	Hartwell et al, 1973
B07809	*S.c.* TDA3	Endosome-to-Golgi retrograde transport	Reid et al, 2011
B28744g	*S.c.* LAS17	Actin filament organization/polymerization	Naqvi et al, 1998
B20416g	*S.c.* RTT107	DNA Damage Repair	Scholes et al, 2001
B05185g	*S.c.* UBR1	ERAD E3 Ubiquitin Ligase	Bartel et al, 1990
A21509g	*S.c.* GPI3	UDP-Glycotransferase subunit	Leidich et al, 1995

Figure 2.3. Biological Function of Solid LOF screening hits. (a) Broad categorical filtering of all 327 essential gene hits for solid media, derived from GO-terms and BLAST alignments from the literature. (b) Representative solid essential gene hits, with homology-derived putative function.

Representative genes from each of the broad categories, along with their *S. cerevisiae* homologs, were cataloged as well; the lowest FS value from each category was chosen (**Figure 2.3b**). Of these genes, F30254g was noteworthy as homologous to the *Candida albicans* gene *cdc39*, a transcription factor putatively linked to deficits in filamentous growth (Uhl et al. 2003; Epp et al. 2010). F37890g is another hit of interest, as A homolog to an elF3 translation factor component in *Neurospora crassa* and *S. cerevisiae*. In the case of *S. cerevisiae*, the knockout decreased biofilm formation (Vandenbosch et al. 2013). Previous work has shown a putative benefit of yeast biofilm formation for solid medium growth, which may explain the essentiality of this knockout on solid media (Lima-Pérez et al. 2018). In both cases, the knockouts may disrupt phenotypes that are believed to be advantageous for yeast growth on solid media. The other hits affect a wide range of cellular mechanisms, including *para*-aminobenzoic acid synthesis (Botet et al. 2007), cell cycle progression (Hartwell et al. 1973), and endomembrane trafficking (Reid et al. 2011). There weren't always clear connections in the literature between the hits and growth on solid media, so in order to find more consistent patterns in the relationship between cellular function and media type, we refocused our analysis on GO-term differences between solid and liquid media.

2.4.3 Solid and Liquid State Essential Hit FunCat Analysis identifies putative essential cellular systems for growth on solid media

To further elucidate categories of genes that are uniquely essential for solid media growth, the essential gene hits for both the solid and liquid conditions were analyzed using FunCat term categories, a gene annotation scheme that serves as a more hierarchical alternative to gene ontology (GO) terms (Ruepp et al. 2004). We utilized the open source analysis tools in the FungiFun2 online platform (Priebe et al. 2015), in which statistically significant ($p < 0.05$) FunCat term enrichments and depletions for submitted sets of genes were calculated using one-sided Fisher's exact tests, comparing FunCat term prevalence across the YALI0 genome to their occurrences in the submitted gene sets. We then arranged the significantly enriched and depleted terms into their respective term hierarchies, enabling us to visualize biological trends unique to the two growth conditions (**Figure 2.4**, **Supplemental Figure S2.1**, **Supplemental Figure S2.2**, **Supplemental Figure S2.3**, **Supplemental File S2.6**).

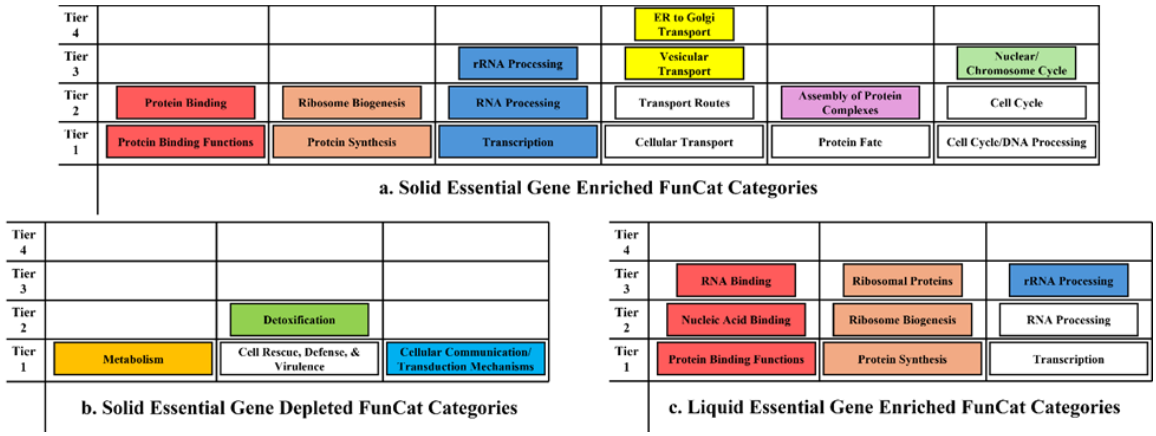


Figure 2.4. Significant FunCat term enrichments and depletions between solid and liquid essential hits. Colored squares represent statistically significant occurrences of the FunCat annotations in the given gene set, while white squares categories did not have significant occurrences of the listed FunCat term (significance was determined through one-sided Fisher's exact tests). (a) Solid Essential Gene Enriched FunCat Term Categories. There were some shared categories of enrichment with the liquid essential gene hits, but vesicular and ER to Golgi transport categories were unique to the solid essential gene hits, as were protein complex assembly and nuclear/chromosomal cycle categories. (b) Solid Essential Gene Depleted FunCat Term Categories. Metabolism, detoxification, and cellular communication were under-represented terms across the solid essential gene hits. (c) Liquid Essential Gene Enriched FunCat Term Categories. Here, we see substantial overlap in the protein binding function, protein synthesis, and transcription categories with the solid essential gene hits, but with some more specific categories emphasized, including RNA binding and ribosomal proteins.

The FunCat analysis of our solid and liquid essential gene sets identified significantly enriched and depleted annotation categories across nine broad biological function classifications (**Figure 2.4**). Some terms were significantly enriched across both solid and liquid essential gene sets (**Figure 2.4a,c**). They tended to fall under broad categories, such as "protein binding", though "RNA binding" was only a significant term hit for liquid essential genes (**Figure 2.4a,c**). "Protein synthesis" and its sub-category of "ribosome biogenesis" also were significant hits for both gene sets, though a third tier sub term, "ribosomal proteins", was only a significant hit for liquid (**Figure 2.4a,c**). Interestingly, the third tier term "rRNA" processing was significantly enriched for both

gene sets, but its parent terms of "RNA processing" and "Transcription" (second and first tier respectively) were only significantly enriched for the solid essential gene set (**Figure 2.4a,c**). Overall, there were few terms that were enriched only in the liquid essential gene set, but the few that were RNA and ribosome related (**Figure 2.4c**).

While no FunCat terms were significantly depleted in the liquid essential gene set, a few terms were significantly depleted from the solid essential gene set. All fell into three categories without any significant enrichment in either gene set (**Figure 2.4b**). The broad, first tier terms of "Metabolism" and "Cellular Communication/Signal Transduction Mechanism" were depleted, as was the second tier term of "Detoxification" (though its parent first tier term, "Cell Rescue, Defense and Virulence" was not significantly depleted). On their own, these depleted terms don't indicate a pattern that matches trends in existing literature.

The most relevant FunCat results showed several distinct cellular functions with significant enrichment only in the solid essential gene knockouts, indicating the putative importance of those functions for growth on solid media. Two of the significant term hits fell under the broad "cellular transport, transport facilitation, and transport route" first tier category, also falling under the "Transport Route" second tier category. The third tier term "Vesicular Transport", and its own fourth tier sub-term "ER to Golgi Transport" were both enriched only for the solid essential gene hits (**Figure 2.4a**).

There are some possible reasons why endomembrane and cellular transport-related genes could be essential more often on solid media. Some genes related to the endomembrane

system, such as the Ras superfamily proteins *RAB4* and *YPT1*, and the peroxisome membrane gene *Y1INP1* have previously been shown to be upregulated or beneficial during hyphal growth in SSF (Chang et al. 2007; Salgado-Bautista et al. 2020; Swennen and Beckerich 2007). Since cell wall components must be transported over greater distances for hyphae to grow on solid media, endomembrane and cellular transport genes are likely to play a key role in hyphal growth on solid media (Pakula et al. 2005). Since genes that promote hyphal growth or transition are already known to be beneficial for SSF (Gomes et al. 2018), it is not surprising that endomembrane system genes, which may play a role in hyphal growth, were often found to be essential in solid media, though further studies will be required to validate the specific essentialities of these genes for growth on solid media.

For both solid and liquid essential genes, we demonstrated possible trends in essential cellular functions that vary between them. While further studies will need to be conducted to validate and further refine the connections between endomembrane system and mitochondrial genes and essentiality on solid media, these results provide individual knockouts to study in more detail with a focus on how their loss affects those functions.

2.5 Chapter 2 Conclusion

The preceding results demonstrate various categories of gene that appear to be especially important for growth in solid and liquid conditions for *Yarrowia lipolytica*. The refined CRISPR knockout library was well represented in the samples, and provided broad coverage of possible knockouts, even after experimental filters were applied. We also demonstrated the viability of pooled solid-media knockout screens for *Yarrowia* lipolytica. By also applying GO-term analysis, we were able to identify additional patterns within our knockout FS values. Further investigation of the suppression of vacuole autophagy may yield further understanding which cellular functions should be suppressed to maximize cellular efficiency for growth on solid substrates. Conversely, strain alterations that may interfere with mitochondrial function and aerobic respiration appear to be incompatible with efficient growth on solid media. By utilizing our refined CRISPR knockout library coupled to various types of pooled growth screens, we may identify valuable strain engineering targets for growth under increasingly broad ranges of conditions and for a wider range of industrial bioprocessing tasks.

Chapter 3: Optimized genome-wide CRISPR screening enables rapid engineering of growth-based phenotypes in *Yarrowia lipolytica*

3.1 Chapter 3 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.

3.2 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* (Lupish et al. 2022; Patterson et al. 2018; Jagtap et al. 2021; 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., 2019a), or a 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). 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., 2022). 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.

3.3 Materials and Methods

3.3.1. Microbial strains and culturing

All strains used in this work are presented in **Supplemental 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.

3.3.2. Plasmid construction

All plasmids and primers used in this work are listed in **Supplemental Table S2** and **Supplemental Table S3**, respectively. The plasmids used to knock out genes in *Y. lipolytica* PO1f were constructed by ordering the corresponding sgRNA as a primer (**Supplemental Table S3**) 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 molar ratio (insert:vector). The assembly product was then transformed directly into electrocompetent *E. coli* TOP10 cells to eventually be propagated and harvested by Miniprep.

3.3.3. *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. (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 hours (final OD ~30). 300 µL of culture (~10^8 cells) were pelleted by centrifugation at 4,000g for 2 minutes 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 hour 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.

3.3.4. 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., 2019a), while the third guide was designed by DeepGuide predictions (Baisya et al., 2022; Schwartz et al., 2019a). 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.

3.3.5. 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 DpnI 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 hr 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 hr, 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.

3.3.6. *Y. lipolytica* Library Transformation

Transformation of the sgRNA plasmid libraries into PO1f and PO1f Cas9::A08 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 hours (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 hour. 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 hours 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 x 10^6 total transformants per biological replicate).

3.3.7. *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 0.5 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 uL (approximately 1x10^7 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 OD_{600} 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.

3.3.8. 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 10^7 plasmids so that sufficient coverage of the sgRNA library is ensured.

At least 0.2 ng of plasmids (approximately $3x10^7$ 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 Cas9 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 (**Supplemental Table S3**). Approximately $1x10^7$ 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 hydrocarbon screens are provided in **Supplemental File S3.7**.

3.3.9. 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 v2..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 **Supplemental Table S4** and the demultiplexing primers/barcodes are provided in **Supplemental Table S5**.

3.3.10. 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 shown in **Figure 3**.1a and 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. The sgRNA FS values from the three 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 used to remove low-activity sgRNA from the original library v2, as the ac-coefficient was found to be maximum at this threshold. To identify essential genes for growth in glucose (Turki media), a one-tailed test of significance was performed. To identify genes beneficial and detrimental to growth (GOF and LOF) in the two acetate conditions, a two-tailed test was performed. In all cases, genes having FDR-corrected $p < 0.05$ were deemed as significant.

3.3.11. 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) (**Figure 3.1**). First, 500 mL bottles of 2x concentration low melt agarose minimal media were prepared via autoclaving (for each bottle, 1 g $MgSO_4 * 7H_2O$, 4 g $(NH_4)_2$, 10 mg $FeCl_3 * 6H_2O$, 15 g LMP Agarose, 200 mg Uracil, 4 µg Myo-Inositol, 8 µg Biotin, 200 µg Thiamine HCl, 15 g KH_2PO_4, and 5.5 g K_2HPO_4), 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 minutes (85% amplitude, 3 second on 3 second 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 hours and then refrigerated until needed, up to 3 days.

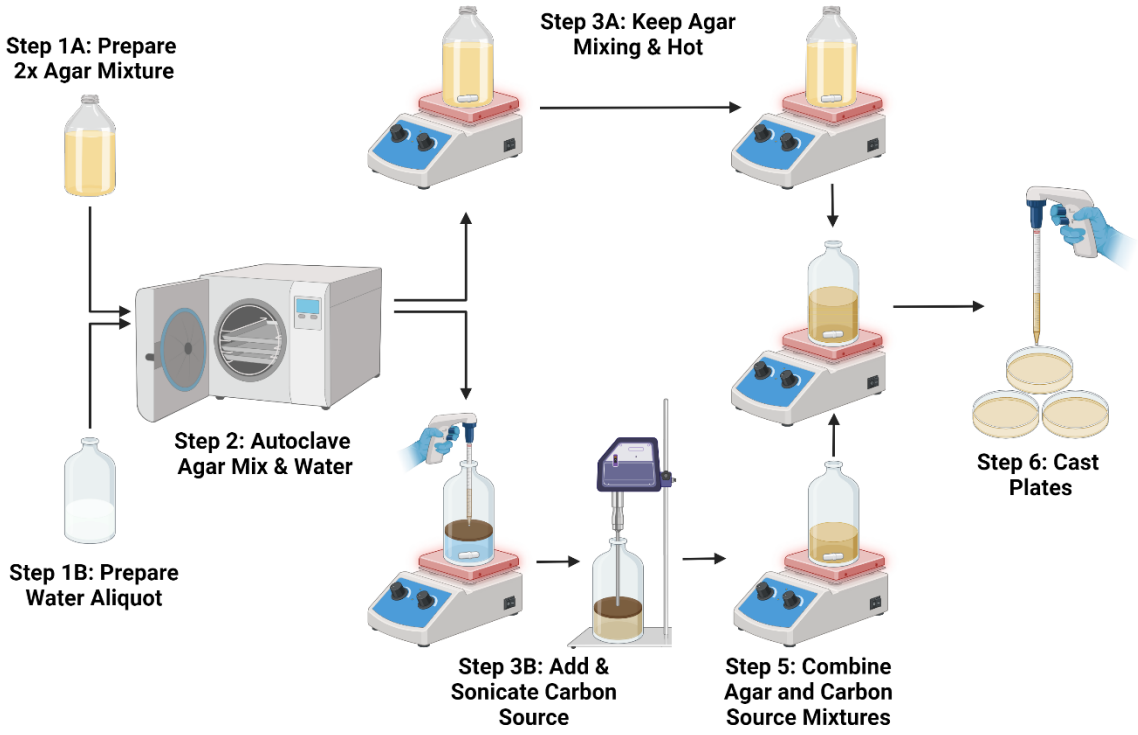


Figure 3.1 Hydrocarbon Plate Preparation Methodology: Method utilized to prepare Turki Minimal media low melt agarose plates with homogenized hydrocarbon and fatty acid carbon sources.

3.3.12. *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 hours at 225 RPM, 30°C. Cells were recovered as a pellet via centrifugation (at 4000 g for 5 minutes). 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 OD_{600} 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 hour, then they were grown in a static 30°C incubator for up to 3 days. Starting at 24 hours, and every 24 hours 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.

3.3.13. *Y. lipolytica* Hydrocarbon and Fatty Acid Plate Screen

Duplicate starter cultures of library cells were prepared 12 hours prior to the start of the assay as in section 3.3.12. 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 3.3.11) using sterile beads (16 total plates per formulation). Plated cells were allowed to dry in a laminar air flow hood for 1 hour, then they were grown in a static 30°C incubator until the previously determined screen endpoint (see section 3.3.12; 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 KH_2PO_4, 0.03 M K_2HPO_4) 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 minutes 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 minutes to melt the media and free the cell mass. The vessels were then centrifuged at 4000g for 5 minutes. 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 3.3.8 and 3.3.9), 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 **Supplemental File S3.7**.

3.4 Chapter 3 Results and Discussion

3.4.1 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., 2019b) (denoted as "v1") and sgRNAs predicted to be of high activity by an activity prediction tool, DeepGuide (Baisya et al., 2022) (**Figure 3.1 a,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., 2019b). 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 (**Supplemental Figure S3.1a**). 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 sgRNA targeting the same gene to maximize diversity of the generated knockouts; further details can be seen in Trivedi, Ramesh, and Wheeldon 2023 and Ramesh and Wheeldon 2021. Moreover, next generation sequencing of the cloned, untransformed library revealed a tight normal distribution of sgRNA abundances, indicating a near-uniform and symmetrical (median ≈ mean) distribution of guides in the library (**Supplemental Figure S3.1b**). 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 **(Supplemental File S3.7**).

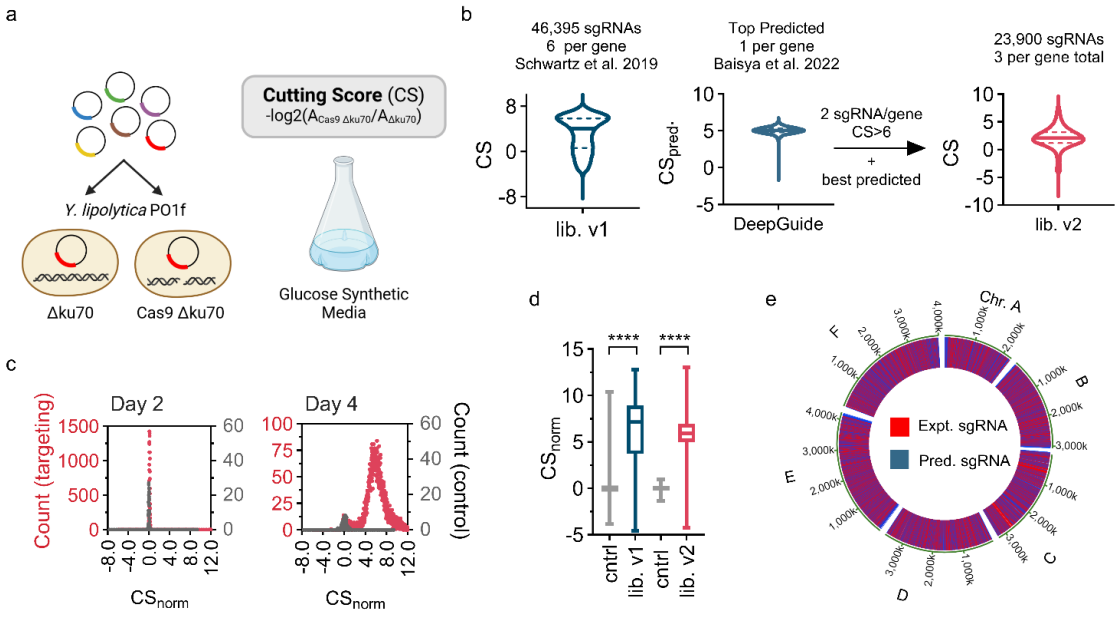


Figure 3.1. Optimized CRISPR library generation pipeline and characterization. (a) Cutting score (CS) is calculated as the depletion of a guide in the Cas9 *ku70* knockout strain compared to the *ku70* knockout control. A guide which cuts well in the Cas9 *ku70* knockout strain will become depleted as 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.

In addition to gene-targeting guides, library v2 also consists of 360 non-targeting sgRNAs (~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±0.09 and avg. $CS_{norm,\ non\text{-}targeting}$ = 0±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±2.04 and avg. $CS_{norm,\ non\text{-}targeting}$ = 0±0.34), indicating different activity profiles of targeting and non-targeting

guides (**Figure 3.1c**). 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 (**Figure 3.1d**). 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 (**Figure 3.2e**). 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 coverage.

3.4.2. Acetate tolerance screens identify gene targets for improved growth

Utilizing the optimized genome-wide CRISPR library, 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 M and 0.5 M, respectively) as experimental conditions (**Figure 3.2a**). 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 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 (**Figure 3.2b**).

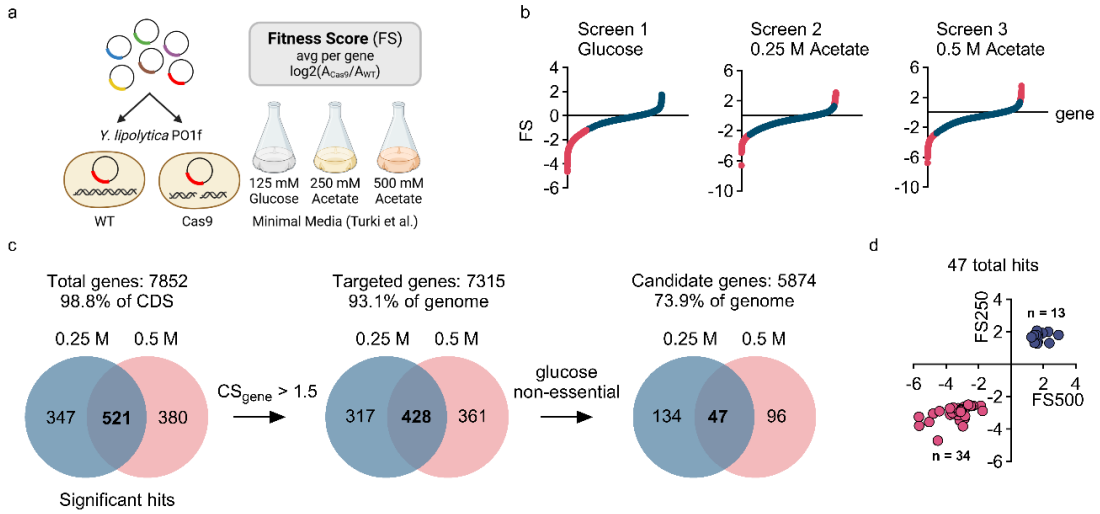


Figure 3.2. Growth-based functional genomic CRISPR screening with acetate as the sole carbon source. (a) Growth screens were conducted using *Y. lipolytica* PO1f and a strain containing Cas9

as control and treatment strains respectively in three media conditions - 0.125 M glucose, 0.25 M 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.

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 (**Figure 3.2c**). 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 (**Figure 3.2d**).

3.4.3. 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 (**Figure 3.3a**). 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 (**Supplemental Figure S3.2**). 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.

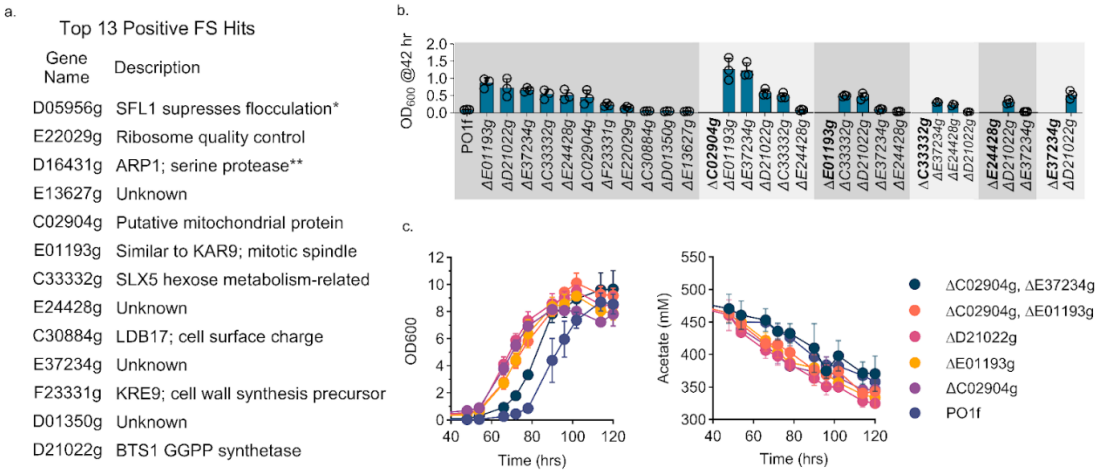


Figure 3.3. Validating acetate metabolism hits. (a) Thirteen positive FS value hits in descending order of high acetate FS. Asterisk (*) indicates knockout was not included due to flocculation, double asterisk (**) indicates knockout was not achieved. (b) Single (left dark gray box) and double knockout (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 hrs, 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.

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 hours, *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 ($p<0.05$) better than the control after correcting for multiple comparisons, including E01193g, D21022g, E37234g, C33332g, E24428g, C02904g, F23331g, and E22029g (**Figure 3.3b**). 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 (**Supplemental Figure S3.3**) 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 (**Figure 3.3b**).

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 (**Figure 3.3c**). 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 (**Supplemental Figure S3a**). Roughly 25% of acetate was converted to dry cell weight (DCW). All hits have a shorter lag phase (reduced by 24 hours in the most potent knockout) and consume more acetate than the PO1f control strain (**Figure 3.3c**) and ΔC02904g, ΔE01193g, ΔE37234g, and ΔC02904g/ΔE01193g strains had significantly higher growth rates (**Supplemental Figure S3.4**). 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 et al., 2000)), C02904g (FMP42, putative mitochondrial protein in *S. cerevisiae* (Imamura et al., 2015)), and D21022g (GGPP synthetase homolog (Saikia et al., 2009, Morton et al., 2011)). GGPP synthetase converts farnesyl pyrophosphate (FPP), the precursor to steroids and N-glycans, to GGPP, a precursor to terpenoids and other metabolites (Chen et al., 2024, 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.

3.4.4. 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 (**Figure 3.4a**).

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 (**Figure 3.4b**).

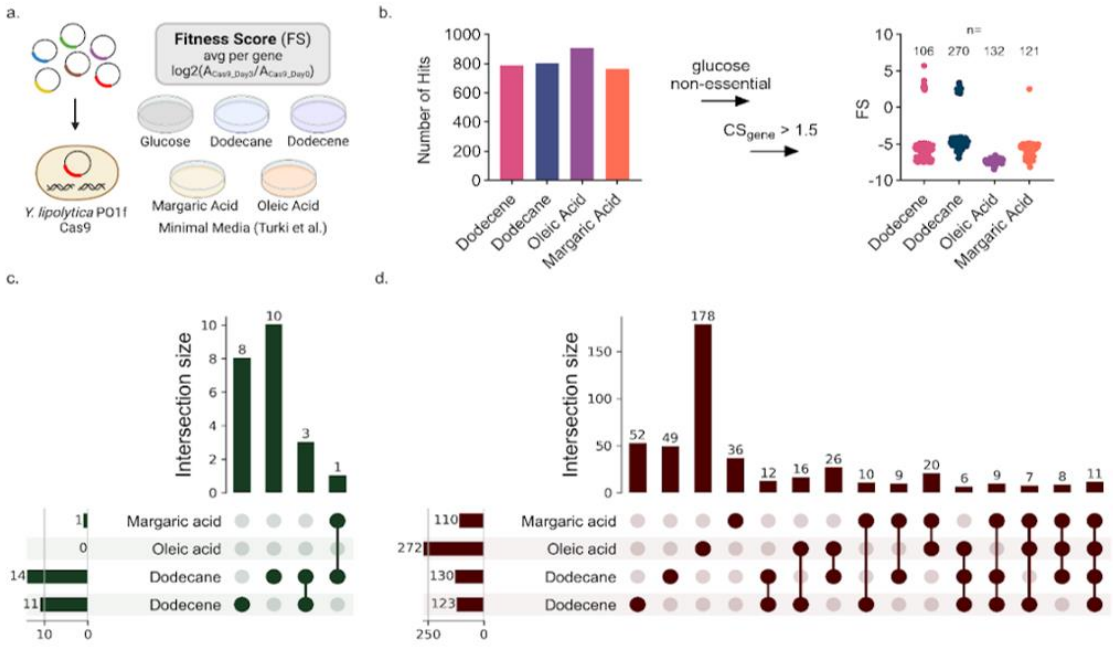
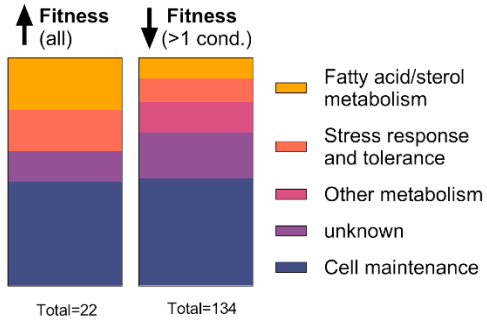


Figure 3.4. Growth-based functional genomic CRISPR screening with hydrocarbon substrates. (a) Growth screen, carbon sources screened and fitness score (FS). Genes with a higher FS indicate that the knockout gave a fitness advantage. (b) Filtering of gene hits obtained for each hydrocarbon condition; 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.

Upset plot analysis is a visual means of identifying hits that are common across the various conditions (**Figure 3.4c,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 (**Figure 3.4c**). 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 and eight for dodecene (see **Supplemental File S3.6** for a list of all genes in the upset plots). Additional analysis of screen hits with positive FS is shown in **Figure 3.5**. With respect to the essential genes (*i.e.*, those with low FS values), the overlap between carbon sources was more comprehensive (**Figure 3.4d**). 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. Functional analysis of these genes, as well as hits with a fitness advantage, are presented in **Figure 3.5**.

a.



b.

Gene	Homology	Description
E36308g	*S.c. HOB1/2*	Lipid transport protein
E03584g	*S.c. PDI*	Triglyceride transfer
A18344g	*S.c. ERG5*	C22 sterol desaturase
D20202g	*S.c. UBP15*	Growth defects on oleic acid

Figure 3.5. 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) Fatty acid metabolism-related gene hits that improve fitness in at least one hydrocarbon condition.

Our screen revealed 134 total genes that reduce fitness in at least two hydrocarbon/fatty acid conditions (**Figure 3.5a**). Twelve hits that reduced fitness were directly related to hydrocarbon metabolism. One identified was C03415g, a hit in all four hydrocarbon conditions, and is known to be essential for fatty acid metabolism in *Aspergillus nidulans* (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.c VPS29*,a membrane-bound vesicle reverse-transport protein (Seaman et al. 1998)), and D14097g, found to be essential for both hydrocarbons and margaric acid, which is homologous to *S.c VPS53*, responsible for retrograde sorting in the Golgi apparatus (Conibear and Stevens 2000). We expected hits such as these two, as there are believed to be genes involved with endomembrane trafficking that affect alkane metabolism (Fukuda 2013). Other hits include the previously mentioned *ADN1* homologue, and E07121g found in our dodecane hits (homologous to *PMT4* in *C. albicans*, plays a role in invasive hyphal growth (Lengeler et al. 2008)). We expected fitness advantages 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 fourteen hits related to stress response/tolerance genes, a class of genes that are expected to be essential for growth under these conditions. There were sixty-three 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 greatly reduce cellular fitness on the carbon source (**Supplemental Figure S5**). Taken together, these hits reveal metabolic pathways and cellular functions that are essential for survival in hydrocarbons and fatty acids.

In addition to the hydrocarbon 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 (Sacristán-Reviriego et al. 2014; Klein et al. 2002; Jungmann, Rayner, and Munro 1999). 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, Yang, and Lee 2021)) and F01346g a homolog of *MSG5* in *S. cerevisiae* involved in cell signaling (Sacristán-Reviriego et al. 2014; Novodvorska et al. 2016) were found in the dodecane and 1-dodecene screens, respectively. We uncovered four hits directly related to improved tolerance to hydrocarbons (**Figure 3.5b**). 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) (**Figure 3.5b**). 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) (**Figure 3.5b**). 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) **(Figure 3.5b**). 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 (Debelyy et al. 2011) (**Figure 3.5b**). It helps regulate transport from the ER to the peroxisome and its disruption causes peroxisomal clustering (Debelyy 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). 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.

3.5 Chapter 3 Conclusion

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 not obvious gene targets. 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.

Conclusions and Future Work: Further development of optimized *Yarrowia lipolytica* bioprocessing strains

In the preceding works, we successfully carried out three functional applications of our CRISPR genome wide knockout screens, demonstrating how they can be used to identify essential genes and promising strain engineering targets in *Yarrowia lipolytica* for a wide range of industrial bioprocessing purposes. We used our optimized genome-wide CRISPR libraries to generate a strain optimized for a scaled up liquid bioproduction batch, identify genes and cellular functions related to growth on solid media, and find deletion mutants that lead to improved metabolism of alternative carbon sources. Each of these results represents a validation of using CRISPR screens to inform our design of optimized bioprocessing strains of *Yarrowia lipolytica*. Further experiments for each of the traits screened for will provide further understanding of *Yarrowia lipolytica* cellular functions, while enabling us to validate and improve upon our newly engineered knockout strains.

Further validation and development of a null-hyphal bioreactor strain

With the null-hyphal PO1f∆*RAS2* validated in terms of growth and one production scenario, we should subject it to further characterization experiments. It would be valuable to knock out *RAS2* in another *Yarrowia lipolytica* production strain to test the null-hyphal strain's viability in a different production scenario. One option could be a limonene production strain (Arnesen et al. 2020). It would also be wise to subject the

strain to additional tests used to validate industrial strains, such as substrate consumption studies (Cossar 2003).

Further studies to quantify performance improvements in null-hyphal strain in a bioreactor setting should also be conducted. As uneven oxygen diffusion was a major rationale for engineering the null-hyphal phenotype in the first place, it is important to quantify any resulting improvements to dissolved oxygen concentration (DOC). Therefore, both the wild type and Δ*RAS2* strain should be subjected to multi-day bioreactor growth runs, with oxygenation probes measuring DOC for both over the full course of the run (Magdouli et al. 2018; Vandermies and Fickers 2019). Other parameters should also be compared during and after the test runs, including pH fluctuations and fouling of the bioreactor surface (Meyer et al. 2021). Any reduction of oxygen fluctuation, pH fluctuation, or biofouling would further demonstrate the utility of a null-hyphal strain.

Strain validations and cellular mechanism studies to expand understanding of growth on solid medium

The initial results of the solid media screens point to the putative essentiality of certain categories of genes in *Yarrowia lipolytica* for growth on solid media. Further investigations should be made into the advantageous knockouts for growth on solid and liquid media, both for validation of the screen and to elucidate possible engineering targets. First, the identified advantageous knockouts should be generated and grown

individually to validate the screen hits. To compare growth between solid and liquid media, dry cell masses could be compared (Try et al. 2018). Additionally, the solid-liquid screens should be repeated with full nutrient media in lieu of minimal media, to ensure that nutrient deficits didn't skew the screen results. Growth on different moisture concentrations of the agarose solid substrates should also be compared, since the water activity parameter (A_w) has been demonstrated to affect solid media growth of *Yarrowia lipolytica* (Gomes et al. 2018). Further solid substrates of industrial interest (such as olive mill solid waste) could also be screened. Even if the essential gene and advantageous knockout hits diverge from our original results, they will help fill in a more complete picture of which genes are most relevant to growth on solid substrates.

It would also be useful to explore how effective our GO-term derived cell function categories (and their respective differences between the screens) were for identifying trends in essentiality on solid media. To test them, we could generate the indicated essential gene knockouts from the screen, and grow them on solid and liquid media along with stains to test the various cellular function categories associated with essential gene hits. For example, we could use endomembrane fluorescent reporters to compare endomembrane structure and interactions with other organelles (Bourett et al. 2007; Benhamou et al. 2018). Mitochondrial stains (such as MitoTracker Red) may also be used to visualize their morphology and distribution under solid and liquid conditions (Epremyan et al. 2022). Wild type cells should also be analyzed with these stains under both conditions, to serve as a comparison to the knockouts, as well as establish baseline endomembrane and mitochondrial characteristics for both conditions.

General lack of knowledge of the differences in gene expression of *Yarrowia lipolytica* between solid and liquid media are still a problem in trying to optimize strain designs for different phase substrates. Therefore, it would be wise to carry out parallel screening approaches to better characterize gene expression, mRNA translation, and protein levels during growth under both solid and liquid conditions. Techniques such as messenger RNA sequencing (Wang et al. 2018), ribosome profiling (Alva et al. 2021) and mass spectrometry (Pomraning et al. 2016) could all be used to identify further differences in expressed genes. The differing levels of gene expression and corresponding protein levels could be used to validate the potential importance of the cellular function categories identified from the solid liquid genome wide knockout screen.

Further strain construction, validation, and testing for growth with hydrocarbons and fatty acids

While we identified gene knockouts that provide putative benefits for *Yarrowia lipolytica* cells grown with acetate, fatty acids, or hydrocarbons as their sole energy source, further characterization of these mutants should be carried out. First, the advantageous knockouts identified for the hydrocarbon and fatty acid screens should be generated in the same manner as for the acetate advantageous knockout hits. Then, their growth rates in the presence of various hydrocarbons can be measured, along with their metabolization of the hydrocarbon via gas chromatography (Zhang et al. 2016). Growth capabilities may be estimated on solid media (Try et al. 2018), but it would be advantageous to test growth rates of the knockouts in liquid conditions. Stirred tank bioreactors may be used to grow

Yarrowia lypolytica with hydrophobic hydrocarbon energy sources, so changes to cell density over time could be used to estimate growth rates (Miranda et al. 2024). It might also be prudent to re-run the whole genome knockout screen for the hydrocarbons and fatty acids in liquid cultures, to both validate the original hit results, and to determine whether growth on solid media led to different identified essential genes and advantageous knockouts.

Overall, the approaches presented in the preceding experiments demonstrate the utility and importance of genome wide functional screens in the development of non-conventional bioprocessing hosts. As we continue to develop new screening and validation tools, we will be able to fill in more of the knowledge gaps on non-conventional industrial organisms such as *Yarrowia lipolytica*, enabling us to design and utilize the next generation of bioprocessing strains for an ever expanding number of roles, from bioproduction to environmental remediation.

Bibliography

Abdel-Mawgoud, A. M., Markham, K. A., Palmer, C. M., Liu, N., Stephanopoulos, G., & Alper, H. S. (2018). Metabolic engineering in the host Yarrowia lipolytica. *Metabolic Engineering*, *50*, 192–208.

Abdullah, R., Nisar, K., Aslam, A., Iqtedar, M., & Naz, S. (2015). Enhanced production of xylanase from locally isolated fungal strain using agro-industrial residues under solid-state fermentation. *Natural Product Research*, *29*(11), 1006–1011.

Adames, N. R., Gallegos, J. E., & Peccoud, J. (2019). Yeast genetic interaction screens in the age of CRISPR/Cas. *Current Genetics*, *65*(2), 307–327.

Afgan, 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., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, *46*(W1), W537–W544.

Ahamed, A., & Vermette, P. (2010). Effect of mechanical agitation on the production of cellulases by Trichoderma reesei RUT-C30 in a draft-tube airlift bioreactor. *Biochemical Engineering Journal*, *49*(3), 379–387.

Alkasrawi, M., Nandakumar, R., Margesin, R., Schinner, F., & Mattiasson, B. (1999). A microbial biosensor based on Yarrowia lipolytica for the off-line determination of middle-chain alkanes. *Biosensors & Bioelectronics*, *14*(8–9), 723–727.

Alva, T. R., Riera, M., & Chartron, J. W. (2021). Translational landscape and protein biogenesis demands of the early secretory pathway in Komagataella phaffii. *Microbial Cell Factories*, *20*(1), 19.

Arnesen, J. A., Kildegaard, K. R., Cernuda Pastor, M., Jayachandran, S., Kristensen, M., & Borodina, I. (2020). Yarrowia lipolytica Strains Engineered for the Production of Terpenoids. *Frontiers in Bioengineering and Biotechnology*, *8*, 945.

Asemoloye, M. D., & Marchisio, M. A. (2022). Synthetic Saccharomyces cerevisiae tolerate and degrade highly pollutant complex hydrocarbon mixture. *Ecotoxicology and Environmental Safety, 241*, 113768.

Baile, M. G., Guiney, E. L., Sanford, E. J., MacGurn, J. A., Smolka, M. B., & Emr, S. D. (2019). Activity of a ubiquitin ligase adaptor is regulated by disordered insertions in its arrestin domain. *Molecular Biology of the Cell*, *30*(25), 3057–3072.

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. *Nature Communications*, *13*(1), 922.

Banat, I. M., Carboué, Q., Saucedo-Castañeda, G., & de Jesús Cázares-Marinero, J. (2021). Biosurfactants: The green generation of speciality chemicals and potential production using Solid-State fermentation (SSF) technology. *Bioresource Technology*, *320*(Pt A), 124222.

Bartel, B., Wünning, I., & Varshavsky, A. (1990). The recognition component of the N-end rule pathway. *The EMBO Journal*, *9*(10), 3179–3189.

Bellou, S., Makri, A., Triantaphyllidou, I.-E., Papanikolaou, S., & Aggelis, G. (2014). Morphological and metabolic shifts of Yarrowia lipolytica induced by alteration of the dissolved oxygen concentration in the growth environment. *Microbiology*, *160*(Pt 4), 807–817.

Benli, M., Döring, F., Robinson, D. G., Yang, X., & Gallwitz, D. (1996). Two GTPase isoforms, Ypt31p and Ypt32p, are essential for Golgi function in yeast. *The EMBO Journal*, *15*(23), 6460–6475.

Berman, J., & Sudbery, P. E. (2002). Candida Albicans: a molecular revolution built on lessons from budding yeast. *Nature Reviews. Genetics*, *3*(12), 918–930.

Bernstein, M., Kepes, F., & Schekman, R. (1989). Sec59 encodes a membrane protein required for core glycosylation in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *9*(3), 1191–1199.

Blazeck, J., Hill, A., Liu, L., Knight, R., Miller, J., Pan, A., Otoupal, P., & Alper, H. S. (2014). Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production. *Nature Communications*, *5*, 3131.

Blazeck John, Liu Leqian, Redden Heidi, & Alper Hal. (2011). Tuning Gene Expression in Yarrowia lipolytica by a Hybrid Promoter Approach. *Applied and Environmental Microbiology*, *77*(22), 7905–7914.

Børsting, C., Hummel, R., Schultz, E. R., Rose, T. M., Pedersen, M. B., Knudsen, J., & Kristiansen, K. (1997). Saccharomyces carlsbergensis contains two functional genes encoding the acyl-CoA binding protein, one similar to the ACB1 gene from S. cerevisiae and one identical to the ACB1 gene from S. monacensis. *Yeast* , *13*(15), 1409–1421.

Botet, J., Mateos, L., Revuelta, J. L., & Santos, M. A. (2007). A chemogenomic screening of sulfanilamide-hypersensitive Saccharomyces cerevisiae mutants uncovers ABZ2, the gene encoding a fungal aminodeoxychorismate lyase. *Eukaryotic Cell*, *6*(11), 2102–2111.

Bowman, E. K., Deaner, 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. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(31), 18424–18430.

Breter, H. J., Ferguson, J., Peterson, T. A., & Reed, S. I. (1983). Isolation and transcriptional characterization of three genes which function at start, the controlling event of the Saccharomyces cerevisiae cell division cycle: CDC36, CDC37, and CDC39. *Molecular and Cellular Biology*, *3*(5), 881–891.

Cachera, 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. *Metabolic Engineering*, *83*, 75–85.

Cai, M., Zhang, Y., Hu, W., Shen, W., Yu, Z., Zhou, W., Jiang, T., Zhou, X., & Zhang, Y. (2014). Genetically shaping morphology of the filamentous fungus Aspergillus glaucus for production of antitumor polyketide aspergiolide A. *Microbial Cell Factories*, *13*, 73.

Castro, I. G., Shortill, S. P., Dziurdzik, S. K., Cadou, A., Ganesan, S., Valenti, 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., Conibear, E., & Schuldiner, 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

Cerda, A., Artola, A., Barrena, R., Font, X., Gea, T., & Sánchez, A. (2019). Innovative Production of Bioproducts From Organic Waste Through Solid-State Fermentation. *Frontiers in Sustainable Food Systems*, *3*. https://doi.org/10.3389/fsufs.2019.00063

Chang, J., Fagarasanu, A., & Rachubinski, R. A. (2007). Peroxisomal peripheral membrane protein YlInp1p is required for peroxisome inheritance and influences the dimorphic transition in the yeast Yarrowia lipolytica. *Eukaryotic Cell*, *6*(9), 1528–1537.

Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*(2), 232–235.

Chen, R., Wang, M., Keasling, J. D., Hu, T., & Yin, X. (2024). Expanding the structural diversity of terpenes by synthetic biology approaches. *Trends in Biotechnology*, *42*(6), 699–713.

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. *Microbial Cell Factories*, *15*(1), 113.

Chilakamarry, C. R., Mimi Sakinah, A. M., Zularisam, A. W., Sirohi, R., Khilji, I. A., Ahmad, N., & Pandey, A. (2022). Advances in solid-state fermentation for bioconversion of agricultural wastes to value-added products: Opportunities and challenges. *Bioresource Technology*, *343*, 126065.

Cho, J. H., Noda, Y., & Yoda, K. (2000). Proteins in the early golgi compartment of Saccharomyces cerevisiae immunoisolated by Sed5p. *FEBS Letters*, *469*(2–3), 151–154.

Chow, J., Dionne, H. M., Prabhakar, A., Mehrotra, A., Somboonthum, J., Gonzalez, B., Edgerton, M., & Cullen, P. J. (2019). Aggregate Filamentous Growth Responses in Yeast. *mSphere*, *4*(2). https://doi.org/10.1128/mSphere.00702-18

Coelho, M. A. Z., Amaral, P. F. F., & Belo, I. (2010). *Yarrowia lipolytica : an industrial workhorse*. https://repositorium.sdum.uminho.pt/handle/1822/16867

Conibear, E., & Stevens, T. H. (2000). Vps52p, Vps53p, and Vps54p form a novel multisubunit complex required for protein sorting at the yeast late Golgi. *Molecular Biology of the Cell*, *11*(1), 305–323.

Cossar, D. (2003). Fermentation Process Validation. In *Fermentation Biotechnology* (Vol. 862, pp. 257–271). American Chemical Society.

Costa, C., Henriques, 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 CAGL0J09944g) in antifungal drug and acetic acid resistance. *Frontiers in Microbiology*, *4*, 170.

Crampin, H., Finley, K., Gerami-Nejad, M., Court, H., Gale, C., Berman, J., & Sudbery, P. (2005). Candida albicans hyphae have a Spitzenkörper that is distinct from the polarisome found in yeast and pseudohyphae. *Journal of Cell Science*, *118*(Pt 13), 2935–2947.

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 …. *Applied Microbiology*. https://link.springer.com/article/10.1007/s00253-016-7791-2

Csutak, O., Corbu, V., Stoica, I., Ionescu, R., & Vassu, T. (2015). Biotechnological Applications of Yarrowia lipolytica CMGB32. *Agriculture and Agricultural Science Procedia*, *6*, 545–553.

Cullen, P. J., & Sprague, G. F., Jr. (2000). Glucose depletion causes haploid invasive growth in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13619–13624.

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. *Microbial Cell Factories*, *17*(1), 136.

Da Silva, N. A., & Srikrishnan, S. (2012). Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. *FEMS Yeast Research*, *12*(2), 197–214.

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 Synthetic Biology*, *9*(1), 26–35.

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. *Molecular Systems Biology*, *14*(1), e8064.

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. *Archives of Microbiology*, *171*(6), 386–396.

de Souza, C. E. C., Ribeiro, B. D., & Coelho, M. A. Z. (2019). Characterization and application of Yarrowia lipolytica lipase obtained by solid-state fermentation in the synthesis of different esters used in the food industry. *Applied Biochemistry and Biotechnology*, *189*(3), 933–959.

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. *The Journal of Biological Chemistry*, *286*(32), 28223–28234.

Dodgson, J., Avula, H., Hoe, K.-L., Kim, D.-U., Park, H.-O., Hayles, J., & Armstrong, J. (2009). Functional genomics of adhesion, invasion, and mycelial formation in Schizosaccharomyces pombe. *Eukaryotic Cell*, *8*(8), 1298–1306.

Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology*, *32*(12), 1262–1267.

Dong, Z., Yang, S., & Lee, B. H. (2021). Bioinformatic mapping of a more precise Aspergillus niger degradome. *Scientific Reports*, *11*(1), 693.

Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R., & Bartel, D. P. (2009). RNAi in budding yeast. *Science*, *326*(5952), 544–550.

El Sheikha, A. F., & Ray, R. C. (2023). Bioprocessing of Horticultural Wastes by Solid-State Fermentation into Value-Added/Innovative Bioproducts: A Review. *Food Reviews International*, *39*(6), 3009–3065.

El-Naggar, N. E.-A., Haroun, S. A., Owis, E. A., & Sherief, A. A. (2015). Optimization of β-glucosidase production by Aspergillus terreus strain EMOO 6-4 using response surface methodology under solid-state fermentation. *Preparative Biochemistry & Biotechnology*, *45*(6), 568–587.

Epp, E., Walther, A., Lépine, G., Leon, Z., Mullick, A., Raymond, M., Wendland, J., & Whiteway, M. (2010). Forward genetics in Candida albicans that reveals the Arp2/3 complex is required for hyphal formation, but not endocytosis. *Molecular Microbiology*, *75*(5), 1182–1198.

Epremyan, K. K., Goleva, T. N., Rogov, A. G., Lavrushkina, S. V., Zinovkin, R. A., & Zvyagilskaya, R. A. (2022). The First Yarrowia lipolytica Yeast Models Expressing Hepatitis B Virus X Protein: Changes in Mitochondrial Morphology and Functions. *Microorganisms*, *10*(9). https://doi.org/10.3390/microorganisms10091817

Fang, M., Ren, H., Liu, J., Cadigan, K. M., Patel, S. R., & Dressler, G. R. (2009). Drosophila ptip is essential for anterior/posterior patterning in development and interacts with the PcG and trxG pathways. *Development* , *136*(11), 1929–1938.

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 Research*, *5*(6–7), 527–543.

Fickers, P., Fudalej, F., Le Dall, M. T., Casaregola, S., Gaillardin, C., Thonart, P., & Nicaud, J. M. (2005). Identification and characterisation of LIP7 and LIP8 genes encoding two extracellular triacylglycerol lipases in the yeast Yarrowia lipolytica. *Fungal Genetics and Biology: FG & B*, *42*(3), 264–274.

Fickers, Patrick, Destain, J., & Thonart, P. (2009). Improvement of Yarrowia lipolytica lipase production by fed-batch fermentation. *Journal of Basic Microbiology*, *49*(2), 212–215.

Fontanille, P., Kumar, V., Christophe, G., Nouaille, R., & Larroche, C. (2012). Bioconversion of volatile fatty acids into lipids by the oleaginous yeast Yarrowia lipolytica. *Bioresource Technology*, *114*, 443–449.

Friedman, A., & Perrimon, N. (2004). Genome-wide high-throughput screens in functional genomics. *Current Opinion in Genetics & Development*, *14*(5), 470–476.

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 Research*, *6*(5), 744–750.

Fukuda, R. (2013). Metabolism of hydrophobic carbon sources and regulation of it in n-alkane-assimilating yeast Yarrowia lipolytica. *Bioscience, Biotechnology, and Biochemistry*, *77*(6), 1149–1154.

Fukuda, R. (2023). Utilization of n-alkane and roles of lipid transfer proteins in Yarrowia lipolytica. *World Journal of Microbiology & Biotechnology*, *39*(4), 97.

Gajdoš, P., Ledesma-Amaro, R., Nicaud, J.-M., Čertík, M., & Rossignol, T. (2016). Overexpression of diacylglycerol acyltransferase in Yarrowia lipolytica affects lipid body size, number and distribution. *FEMS Yeast Research*, *16*(6). https://doi.org/10.1093/femsyr/fow062

Gancedo, J. M. (2001). Control of pseudohyphae formation in Saccharomyces cerevisiae. *FEMS Microbiology Reviews*, *25*(1), 107–123.

Ganesan, V., Spagnuolo, M., Agrawal, A., Smith, S., Gao, D., & Blenner, M. (2019). Advances and opportunities in gene editing and gene regulation technology for Yarrowia lipolytica. *Microbial Cell Factories*, *18*(1), 208.

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*(10). https://doi.org/10.1093/g3journal/jkab233

Garvey, M. (2022). Non-Mammalian Eukaryotic Expression Systems Yeast and Fungi in the Production of Biologics. *Journal of Fungi (Basel, Switzerland)*, *8*(11). https://doi.org/10.3390/jof8111179

Gauss, R., Kanehara, K., Carvalho, P., Ng, D. T. W., & Aebi, M. (2011). A complex of Pdi1p and the mannosidase Htm1p initiates clearance of unfolded glycoproteins from the endoplasmic reticulum. *Molecular Cell*, *42*(6), 782–793.

Gervais, N. C., Halder, V., & Shapiro, R. S. (2021). A data library of Candida albicans functional genomic screens. *FEMS Yeast Research*, *21*(7), foab060.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., … Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, *418*(6896), 387–391.

Gimeno, C. J., Ljungdahl, P. O., Styles, C. A., & Fink, G. R. (1992). Unipolar cell divisions in the yeast S. cerevisiae lead to filamentous growth: regulation by starvation and RAS. *Cell*, *68*(6), 1077–1090.

Gonçalves, F. A. G., Colen, G., & Takahashi, J. A. (2014). Yarrowia lipolytica and its multiple applications in the biotechnological industry. *TheScientificWorldJournal*, *2014*, 476207.

Guha, T. K., & Edgell, D. R. (2017). Applications of Alternative Nucleases in the Age of CRISPR/Cas9. *International Journal of Molecular Sciences*, *18*(12). https://doi.org/10.3390/ijms18122565

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. *Biotechnology for Biofuels*, *14*(1), 41.

Hartwell, L. H., Mortimer, R. K., Culotti, J., & Culotti, M. (1973). Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of cdc Mutants. *Genetics*, *74*(2), 267–286.

Harvey, L. M., & McNeil, B. (1994). Liquid Fermentation Systems and Product Recovery of Aspergillus. In J. E. Smith (Ed.), *Aspergillus* (pp. 141–176). Springer US.

Heier, C., Taschler, U., Radulovic, M., Aschauer, P., Eichmann, T. O., Grond, S., Wolinski, H., Oberer, M., Zechner, R., Kohlwein, S. D., & Zimmermann, R. (2016). Monoacylglycerol Lipases Act as Evolutionarily Conserved Regulators of Non-oxidative Ethanol Metabolism. *The Journal of Biological Chemistry*, *291*(22), 11865–11875.

Hirschhorn, J. N., Brown, S. A., Clark, C. D., & Winston, F. (1992). Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes & Development*, *6*(12A), 2288–2298.

Hölker, U., Höfer, M., & Lenz, J. (2004). Biotechnological advantages of laboratory-scale solid-state fermentation with fungi. *Applied Microbiology and Biotechnology*, *64*(2), 175–186.

Horak, C. E., Luscombe, N. M., Qian, J., Bertone, P., Piccirrillo, S., Gerstein, M., & Snyder, M. (2002). Complex transcriptional circuitry at the G1/S transition in Saccharomyces cerevisiae. *Genes & Development*, *16*(23), 3017–3033.

Horwitz, A. A., Walter, J. M., Schubert, M. G., Kung, S. H., Hawkins, K., Platt, D. M., Hernday, A. D., Mahatdejkul-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 Systems*, *1*(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. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(34), E4689-96.

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*(6), e0130397.

Imandi, S. B., Karanam, S. K., & Garapati, H. R. (2013). Use of Plackett-Burman design for rapid screening of nitrogen and carbon sources for the production of lipase in solid state fermentation by Yarrowia lipolytica from mustard oil cake (Brassica napus). *Brazilian Journal of Microbiology: [Publication of the Brazilian Society for Microbiology]*, *44*(3), 915–921.

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. *Biotechnology for Biofuels*, *14*(1), 188.

Jakočiūnas, 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. *Metabolic Engineering*, *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. *The Journal of Biological Chemistry*, *274*(10), 6579–6585.

Kar, T., Destain, J., & Thonart, P. (2012). Scale-down assessment of the sensitivity of Yarrowia lipolytica to oxygen transfer and foam management in bioreactors: investigation of the underlying physiological …. *Journal of Industrial*. https://academic.oup.com/jimb/article-abstract/39/2/337/5994575

Karimi, F., Mazaheri, D., Saei Moghaddam, M., Mataei Moghaddam, A., Sanati, A. L., & Orooji, Y. (2021). Solid-state fermentation as an alternative technology for cost-effective production of bioethanol as useful renewable energy: a review. *Biomass Conversion and Biorefinery*. https://doi.org/10.1007/s13399-021-01875-2

Kawasse, F. M., Amaral, P. F., Rocha-Leão, M. H. M., Amaral, A. L., Ferreira, E. C., & Coelho, M. A. Z. (2003). Morphological analysis of Yarrowia lipolytica under stress conditions through image processing. *Bioprocess and Biosystems Engineering*, *25*(6), 371–375.

Khanahmadi, M., Mitchell, D. A., Beheshti, M., Roostaazad, R., & Sánchez, L. R. (2006). Continuous solid-state fermentation as affected by substrate flow pattern. *Chemical Engineering Science*, *61*(8), 2675–2687.

Kiss, E., Hegedüs, B., Virágh, M., Varga, T., Merényi, Z., Kószó, T., Bálint, B., Prasanna, A. N., Krizsán, K., Kocsubé, S., Riquelme, M., Takeshita, N., & Nagy, L. G. (2019). Comparative genomics reveals the origin of fungal hyphae and multicellularity. *Nature Communications*, *10*(1), 4080.

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. *The Journal of Biological Chemistry*, *277*(28), 25011–25019.

Konzock, O., & Norbeck, J. (2020). Deletion of MHY1 abolishes hyphae formation in Yarrowia lipolytica without negative effects on stress tolerance. *PloS One*, *15*(4), e0231161.

Kwak, Y.-S., Han, S., Thomashow, L. S., Rice, J. T., Paulitz, 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. *Applied and Environmental Microbiology*, *77*(5), 1770–1776.

Lalo, D., Stettler, S., Mariotte, S., Gendreau, E., & Thuriaux, P. (1994). Organization of the centromeric region of chromosome XIV in Saccharomyces cerevisiae. *Yeast* , *10*(4), 523–533.

Lee, B. N., & Elion, E. A. (1999). The MAPKKK Ste11 regulates vegetative growth through a kinase cascade of shared signaling components. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(22), 12679–12684.

Leidich, S. D., Kostova, Z., Latek, R. R., Costello, L. C., Drapp, D. A., Gray, W., Fassler, J. S., & Orlean, P. (1995). Temperature-sensitive yeast GPI anchoring mutants gpi2 and gpi3 are defective in the synthesis of N-acetylglucosaminyl phosphatidylinositol. Cloning of the GPI2 gene. *The Journal of Biological Chemistry*, *270*(22), 13029–13035.

Leite, P., Sousa, D., Fernandes, H., Ferreira, M., Costa, A. R., Filipe, D., Gonçalves, M., Peres, H., Belo, I., & Salgado, J. M. (2021). Recent advances in production of lignocellulolytic enzymes by solid-state fermentation of agro-industrial wastes. *Current Opinion in Green and Sustainable Chemistry*, *27*, 100407.

Lengeler, K. B., Tielker, D., & Ernst, J. F. (2008). Protein-O-mannosyltransferases in virulence and development. *Cellular and Molecular Life Sciences: CMLS*, *65*(4), 528–544.

Leplat, C., Nicaud, J.-M., & Rossignol, T. (2015). High-throughput transformation method for Yarrowia lipolytica mutant library screening. *FEMS Yeast Research*, *15*(6). https://doi.org/10.1093/femsyr/fov052

Leplat, C., Nicaud, J.-M., & Rossignol, T. (2018). Overexpression screen reveals transcription factors involved in lipid accumulation in Yarrowia lipolytica. *FEMS Yeast Research*, *18*(5). https://doi.org/10.1093/femsyr/foy037

Li, M., Li, Y.-Q., Zhao, X.-F., & Gao, X.-D. (2014). Roles of the three Ras proteins in the regulation of dimorphic transition in the yeast Yarrowia lipolytica. *FEMS Yeast Research*, *14*(3), 451–463.

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 Synthetic Biology*, *9*(5), 1030–1040.

Li, Z. J., Shukla, V., Wenger, K., Fordyce, A., Pedersen, A. G., & Marten, M. (2002). Estimation of hyphal tensile strength in production-scale Aspergillus oryzae fungal fermentations. *Biotechnology and Bioengineering*, *77*(6), 601–613.

Lian, J., Schultz, C., Cao, M., HamediRad, M., & Zhao, H. (2019). Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. *Nature Communications*, *10*(1), 1–10.

Lima-Pérez, J., Rodríguez-Gómez, D., Loera, O., Viniegra-González, G., & López-Pérez, M. (2018). Differences in growth physiology and aggregation of Pichia pastoris cells between solid-state and submerged fermentations under aerobic conditions. *Journal of Chemical Technology and Biotechnology* , *93*(2), 527–532.

Liu, F., Ng, S. K., Lu, Y., Low, W., Lai, J., & Jedd, G. (2008). Making two organelles from one: Woronin body biogenesis by peroxisomal protein sorting. *The Journal of Cell Biology*, *180*(2), 325–339.

Liu, H., Marsafari, M., Wang, F., Deng, L., & Xu, P. (2019). Engineering acetyl-CoA metabolic shortcut for eco-friendly production of polyketides triacetic acid lactone in Yarrowia lipolytica. *Metabolic Engineering*, *56*, 60–68.

Liu, J., Liu, J., & Stråby, K. B. (1998). Point and deletion mutations eliminate one or both methyl group transfers catalysed by the yeast TRM1 encoded tRNA (m22G26)dimethyltransferase. *Nucleic Acids Research*, *26*(22), 5102–5108.

Liu, Jiao, 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. *Nature Communications*, *13*(1), 891.

Liu, R., Bassalo, M. C., Zeitoun, R. I., & Gill, R. T. (2015). Genome scale engineering techniques for metabolic engineering. *Metabolic Engineering*, *32*, 143–154.

Liu, X., Liu, M., Zhang, J., Chang, Y., Cui, Z., Ji, B., Nielsen, J., Qi, Q., & Hou, J. (2022). Mapping of Nonhomologous End Joining-Mediated Integration Facilitates Genome-Scale Trackable Mutagenesis in Yarrowia lipolytica. *ACS Synthetic Biology*, *11*(1), 216–227.

Lizardi-Jiménez, M. A., & Hernández-Martínez, R. (2017). Solid state fermentation (SSF): diversity of applications to valorize waste and biomass. *3 Biotech*, *7*(1), 44.

Löbs, A.-K., Schwartz, C., & Wheeldon, I. (2017). Genome and metabolic engineering in non-conventional yeasts: Current advances and applications. *Synthetic and Systems Biotechnology*, *2*(3), 198–207.

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. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.28219

Magdouli, S., Brar, S. K., & Blais, J. F. (2018). Morphology and rheological behaviour of Yarrowia lipolytica: Impact of dissolved oxygen level on cell growth and lipid composition. *Process Biochemistry*, *65*, 1–10.

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*(9), e0162363.

Mamaev, D., & Zvyagilskaya, R. (2021). Yarrowia lipolytica: a multitalented yeast species of ecological significance. *FEMS Yeast Research*, *21*(2). https://doi.org/10.1093/femsyr/foab008

Manan, M. A., & Webb, C. (2017). Design aspects of solid state fermentation as applied to microbial bioprocessing. *J Appl Biotechnol Bioeng*, *4*, 1–22.

Manzella, L., Barros, M. H., & Nobrega, F. G. (1998). ARH1 of Saccharomyces cerevisiae: a new essential gene that codes for a protein homologous to the human adrenodoxin reductase. *Yeast* , *14*(9), 839–846.

Mao, Y., Chen, Z., Lu, L., Jin, B., Ma, H., Pan, Y., & Chen, T. (2020). Efficient solid-state fermentation for the production of 5-aminolevulinic acid enriched feed using recombinant Saccharomyces cerevisiae. *Journal of Biotechnology*, *322*, 29–32.

Martin, S. M., & Bushell, M. E. (1996). Effect of hyphal micromorphology on bioreactor performance of antibiotic-producing Saccharopolyspora erythraea cultures. *Microbiology*, *142*(7), 1783–1788.

Martins, V. G., Kalil, S. J., & Costa, J. A. V. (2009). In situ bioremediation using biosurfactant produced by solid state fermentation. *World Journal of Microbiology & Biotechnology*, *25*(5), 843–851.

Matthäus, F., Ketelhot, M., Gatter, M., & Barth, G. (2014). Production of lycopene in the non-carotenoid-producing yeast Yarrowia lipolytica. *Applied and Environmental Microbiology*, *80*(5), 1660–1669.

Meyer, V., Cairns, T., Barthel, L., King, R., Kunz, P., Schmideder, S., Müller, H., Briesen, H., Dinius, A., & Krull, R. (2021). Understanding and controlling filamentous growth of fungal cell factories: novel tools and opportunities for targeted morphology engineering. *Fungal Biology and Biotechnology*, *8*(1), 8.

Miranda, S. M., Lopes, M., & Belo, I. (2024). Exploring the use of hexadecane by Yarrowia lipolytica: Effect of dissolved oxygen and medium supplementation. *Journal of Biotechnology*, *380*, 29–37.

Misa, J., & Schwartz, C. (2021). CRISPR Interference and Activation to Modulate Transcription in Yarrowia lipolytica. *Methods in Molecular Biology , 2307*, 95–109.

Mitchell, D. A., Krieger, N., Stuart, D. M., & Pandey, A. (2000). New developments in solid-state fermentation: II. Rational approaches to the design, operation and scale-up of bioreactors. *Process Biochemistry*, *35*(10), 1211–1225.

Mitchell, D. A., von Meien, O. F., Luz, L. F. L., & Berovič, M. (2006). The Scale-up Challenge for SSF Bioreactors. In D. A. Mitchell, M. Berovič, & N. Krieger (Eds.), *Solid-State Fermentation Bioreactors: Fundamentals of Design and Operation* (pp. 57–64). Springer Berlin Heidelberg.

Momen-Roknabadi, A., Oikonomou, P., Zegans, M., & Tavazoie, S. (2020). An inducible CRISPR interference library for genetic interrogation of Saccharomyces cerevisiae biology. *Communications Biology*, *3*(1), 723.

Morgunov, I. G., Kamzolova, S. V., Perevoznikova, O. A., Shishkanova, N. V., & Finogenova, T. V. (2004). Pyruvic acid production by a thiamine auxotroph of Yarrowia lipolytica. *Process Biochemistry*, *39*(11), 1469–1474.

Morton, C. O., Varga, J. J., Hornbach, A., Mezger, M., Sennefelder, H., Kneitz, S., Kurzai, O., Krappmann, S., Einsele, H., Nierman, W. C., Rogers, T. R., & Loeffler, J. (2011). The temporal dynamics of differential gene expression in Aspergillus fumigatus interacting with human immature dendritic cells in vitro. *PloS One*, *6*(1), e16016.

Mösch, H. U., & Fink, G. R. (1997). Dissection of filamentous growth by transposon mutagenesis in Saccharomyces cerevisiae. *Genetics*, *145*(3), 671–684.

Mösch, H. U., Roberts, R. L., & Fink, G. R. (1996). Ras2 signals via the Cdc42/Ste20/mitogen-activated protein kinase module to induce filamentous growth in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(11), 5352–5356.

Mukaiyama, H., Kajiwara, S., Hosomi, A., Giga-Hama, Y., Tanaka, N., Nakamura, T., & Takegawa, K. (2009). Autophagy-deficient Schizosaccharomyces pombe mutants undergo partial sporulation during nitrogen starvation. *Microbiology*, *155*(Pt 12), 3816–3826.

Müller, C., Hansen, K., Szabo, P., & Nielsen, J. (2003). Effect of deletion of chitin synthase genes on mycelial morphology and culture viscosity in Aspergillus oryzae. *Biotechnology and Bioengineering*, *81*(5), 525–534.

Müller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M. P., & Boutros, M. (2005). Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature*, *436*(7052), 871–875.

Nagiec, M. M., Wells, G. B., Lester, R. L., & Dickson, R. C. (1993). A suppressor gene that enables Saccharomyces cerevisiae to grow without making sphingolipids encodes a protein that resembles an Escherichia coli fatty acyltransferase. *The Journal of Biological Chemistry*, *268*(29), 22156–22163.

Naqvi, S. N., Zahn, R., Mitchell, D. A., Stevenson, B. J., & Munn, A. L. (1998). The WASp homologue Las17p functions with the WIP homologue End5p/verprolin and is essential for endocytosis in yeast. *Current Biology: CB*, *8*(17), 959–962.

Nascimento, F. V. do, Lemes, A. C., Castro, A. M. de, Secchi, A. R., & Zarur Coelho, M. A. (2022). A Temporal Evolution Perspective of Lipase Production by Yarrowia lipolytica in Solid-State Fermentation. *Processes*, *10*(2), 381.

Neumeier, J., & Meister, G. (2020). siRNA Specificity: RNAi Mechanisms and Strategies to Reduce Off-Target Effects. *Frontiers in Plant Science*, *11*, 526455.

Next-gen Biomanufacturing Market Size to Reach US$39.4 Billion by 2032. (2024, June 11). BioSpace. https://www.biospace.com/next-gen-biomanufacturing-market-size-to-reach-us-39-4-billion-by-2032

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*. https://www.nature.com/articles/s42003-022-03475-w

Nitsche, B. M., Jørgensen, T. R., Akeroyd, M., Meyer, V., & Ram, A. F. J. (2012). The carbon starvation response of Aspergillus niger during submerged cultivation: insights from the transcriptome and secretome. *BMC Genomics*, *13*, 380.

Oostra, J., le Comte, E. P., van den Heuvel, J. C., Tramper, J., & Rinzema, A. (2001). Intra-particle oxygen diffusion limitation in solid-state fermentation. *Biotechnology and Bioengineering*, *75*(1), 13–24.

Orzua, M. C., Mussatto, S. I., Contreras-Esquivel, J. C., Rodriguez, R., de la Garza, H., Teixeira, J. A., & Aguilar, C. N. (2009). Exploitation of agro industrial wastes as immobilization carrier for solid-state fermentation. *Industrial Crops and Products*, *30*(1), 24–27.

Pakula, T. M., Salonen, K., Uusitalo, J., & Penttilä, M. (2005). The effect of specific growth rate on protein synthesis and secretion in the filamentous fungus Trichoderma reesei. *Microbiology*, *151*(Pt 1), 135–143.

Pal, A., & Khanum, F. (2010). Production and extraction optimization of xylanase from Aspergillus niger DFR-5 through solid-state-fermentation. *Bioresource Technology*, *101*(19), 7563–7569.

Palande, A. S., Kulkarni, S. V., León-Ramirez, C., Campos-Góngora, E., Ruiz-Herrera, J., & Deshpande, M. V. (2014). Dimorphism and hydrocarbon metabolism in Yarrowia lipolytica var. indica. *Archives of Microbiology*, *196*(8), 545–556.

Pan, X., & Heitman, J. (2002). Protein kinase A operates a molecular switch that governs yeast pseudohyphal differentiation. *Molecular and Cellular Biology*, *22*(12), 3981–3993.

Papanikolaou, S., Chevalot, I., Komaitis, M., Marc, I., & Aggelis, G. (2002). Single cell oil production by Yarrowia lipolytica growing on an industrial derivative of animal fat in batch cultures. *Applied Microbiology and Biotechnology*, *58*(3), 308–312.

Paredes-López, O., & Harry, G. I. (1988). Food biotechnology review: traditional solid-state fermentations of plant raw materials--application, nutritional significance, and future prospects. *Critical Reviews in Food Science and Nutrition*, *27*(3), 159–187.

Park, Y.-K., & Nicaud, J.-M. (2020). Screening a genomic library for genes involved in propionate tolerance in Yarrowia lipolytica. *Yeast* , *37*(1), 131–140.

Parveen, S., Singh, S., & Komath, S. S. (2019). Saccharomyces cerevisiae Ras2 restores filamentation but cannot activate the first step of GPI anchor biosynthesis in Candida albicans. *Biochemical and Biophysical Research Communications*, *517*(4), 755–761.

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

Pomraning, K. R., Bredeweg, E. L., Kerkhoven, E. J., Barry, K., Haridas, S., Hundley, H., LaButti, K., Lipzen, A., Yan, M., Magnuson, J. K., Simmons, B. A., Grigoriev, I. V., Nielsen, J., & Baker, S. E. (2018). Regulation of Yeast-to-Hyphae Transition in Yarrowia lipolytica. *mSphere*, *3*(6). https://doi.org/10.1128/mSphere.00541-18

Pomraning, K. R., Kim, Y.-M., Nicora, C. D., Chu, R. K., Bredeweg, E. L., Purvine, S. O., Hu, D., Metz, T. O., & Baker, S. E. (2016). Multi-omics analysis reveals regulators of the response to nitrogen limitation in Yarrowia lipolytica. *BMC Genomics*, *17*, 138.

Raghavarao, K. S. M. S., Ranganathan, T. V., & Karanth, N. G. (2003). Some engineering aspects of solid-state fermentation. *Biochemical Engineering Journal*, *13*(2), 127–135.

Rahardjo, Y. S. P., Weber, F. J., le Comte, E. P., Tramper, J., & Rinzema, A. (2002). Contribution of aerial hyphae of Aspergillus oryzae to respiration in a model solid-state fermentation system. *Biotechnology and Bioengineering*, *78*(5), 539–544.

Ramesh, A., Ong, T., Garcia, J. A., Adams, J., & Wheeldon, I. (2020). Guide RNA Engineering Enables Dual Purpose CRISPR-Cpf1 for Simultaneous Gene Editing and Gene Regulation in Yarrowia lipolytica. *ACS Synthetic Biology*, *9*(4), 967–971.

Raschmanová, H., Weninger, A., Glieder, A., Kovar, K., & Vogl, T. (2018). Implementing CRISPR-Cas technologies in conventional and non-conventional yeasts: Current state and future prospects. *Biotechnology Advances*, *36*(3), 641–665.

Reid, R. J. D., González-Barrera, S., Sunjevaric, I., Alvaro, D., Ciccone, S., Wagner, M., & Rothstein, R. (2011). Selective ploidy ablation, a high-throughput plasmid transfer protocol, identifies new genes affecting topoisomerase I–induced DNA damage. *Genome Research*, *21*(3), 477–486.

Reis, V. R., Bassi, A. P. G., da Silva, J. C. G., & Ceccato-Antonini, S. R. (2013). Characteristics of Saccharomyces cerevisiae yeasts exhibiting rough colonies and pseudohyphal morphology with respect to alcoholic fermentation. *Brazilian Journal of Microbiology: [Publication of the Brazilian Society for Microbiology]*, *44*(4), 1121–1131.

Robertson, N. R., Trivedi, V., Lupish, B., Ramesh, A., Aguilar, Y., Arteaga, A., Nguyen, A., Lee, S., Lenert-Mondou, C., Harland-Dunaway, M., Jinkerson, R., & Wheeldon, I. (2024). Optimized genome-wide CRISPR screening enables rapid engineering of growth-based phenotypes in Yarrowia lipolytica. In *bioRxiv* (p. 2024.06.20.599746). https://doi.org/10.1101/2024.06.20.599746

Rodríguez Couto, S., & Sanromán, M. A. (2005). Application of solid-state fermentation to ligninolytic enzyme production. *Biochemical Engineering Journal*, *22*(3), 211–219.

Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., & Mewes, H. W. (2004). The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Research*, *32*(18), 5539–5545.

Ruiz-Herrera, J., & Sentandreu, R. (2002). Different effectors of dimorphism in Yarrowia lipolytica. *Archives of Microbiology*, *178*(6), 477–483.

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. *Microbial Cell Factories*, *16*(1), 78.

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*(1), e85390.

Sadh, P. K., Duhan, S., & Duhan, J. S. (2018). Agro-industrial wastes and their utilization using solid state fermentation: a review. *Bioresources and Bioprocessing*, *5*(1), 1.

Saikia, S., & Scott, B. (2009). Functional analysis and subcellular localization of two geranylgeranyl diphosphate synthases from Penicillium paxilli. *Molecular Genetics and Genomics: MGG*, *282*(3), 257–271.

Salgado-Bautista, D., Volke-Sepúlveda, T., Figueroa-Martínez, F., Carrasco-Navarro, U., Chagolla-López, A., & Favela-Torres, E. (2020). Solid-state fermentation increases secretome complexity in Aspergillus brasiliensis. *Fungal Biology*, *124*(8), 723–734.

Sandoval, N. R., Mills, T. Y., Zhang, M., & Gill, R. T. (2011). Elucidating acetate tolerance in E. coli using a genome-wide approach. *Metabolic Engineering*, *13*(2), 214–224.

Scholes, D. T., Banerjee, M., Bowen, B., & Curcio, M. J. (2001). Multiple regulators of Ty1 transposition in Saccharomyces cerevisiae have conserved roles in genome maintenance. *Genetics*, *159*(4), 1449–1465.

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. *Metabolic Engineering*, *55*, 102–110.

Schwartz, C., Curtis, N., Löbs, A.-K., & Wheeldon, I. (2018a). Multiplexed CRISPR Activation of Cryptic Sugar Metabolism Enables Yarrowia Lipolytica Growth on Cellobiose. *Biotechnology Journal*, *13*(9), e1700584.

Schwartz, C., Curtis, N., Löbs, A.-K., & Wheeldon, I. (2018b). Multiplexed CRISPR Activation of Cryptic Sugar Metabolism Enables Yarrowia Lipolytica Growth on Cellobiose. *Biotechnology Journal*, *13*(9), e1700584.

Schwartz, C., Frogue, K., Misa, J., & Wheeldon, I. (2017a). Host and Pathway Engineering for Enhanced Lycopene Biosynthesis in Yarrowia lipolytica. *Frontiers in Microbiology*, *8*, 2233.

Schwartz, C., Frogue, K., Misa, J., & Wheeldon, I. (2017b). Host and Pathway Engineering for Enhanced Lycopene Biosynthesis in Yarrowia lipolytica. *Frontiers in Microbiology*, *8*, 2233.

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. *Biotechnology and Bioengineering*, *114*(12), 2896–2906.

Schwartz, C. M., Hussain, M. S., Blenner, M., & Wheeldon, I. (2016a). Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR–Cas9-Mediated Genome Editing in Yarrowia lipolytica. *ACS Synthetic Biology*, *5*(4), 356–359.

Schwartz, C. M., Hussain, M. S., Blenner, M., & Wheeldon, I. (2016b). Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica. *ACS Synthetic Biology*, *5*(4), 356–359.

Schwartz, C., Shabbir-Hussain, M., Frogue, K., Blenner, M., & Wheeldon, I. (2017a). Standardized Markerless Gene Integration for Pathway Engineering in Yarrowia lipolytica. *ACS Synthetic Biology*, *6*(3), 402–409.

Schwartz, C., Shabbir-Hussain, M., Frogue, K., Blenner, M., & Wheeldon, I. (2017b). Standardized Markerless Gene Integration for Pathway Engineering in Yarrowia lipolytica. *ACS Synthetic Biology*, *6*(3), 402–409.

Seaman, M. N., McCaffery, J. M., & Emr, S. D. (1998). A membrane coat complex essential for endosome-to-Golgi retrograde transport in yeast. *The Journal of Cell Biology*, *142*(3), 665–681.

Seppälä, S., Wilken, S. E., Knop, D., Solomon, K. V., & O'Malley, M. A. (2017). The importance of sourcing enzymes from non-conventional fungi for metabolic engineering and biomass breakdown. *Metabolic Engineering*, *44*, 45–59.

Shama, S., Kirchman, P. A., Jiang, J. C., & Jazwinski, S. M. (1998). Role ofRAS2in Recovery from Chronic Stress: Effect on Yeast Life Span. *Experimental Cell Research*, *245*(2), 368–378.

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*(1), 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 Biology*, *17*, 45.

Soccol, C. R., Costa, E. S. F. da, Letti, L. A. J., Karp, S. G., Woiciechowski, A. L., & Vandenberghe, L. P. de S. (2017). Recent developments and innovations in solid state fermentation. *Biotechnology Research and Innovation*, *1*(1), 52–71.

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. *Journal of Industrial Microbiology & Biotechnology*, *42*(11), 1519–1531.

Sudbery, P., Gow, N., & Berman, J. (2004). The distinct morphogenic states of Candida albicans. *Trends in Microbiology*, *12*(7), 317–324.

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. *Frontiers in Microbiology*, *10*, 750.

Swennen, D., & Beckerich, J.-M. (2007). Yarrowia lipolytica vesicle-mediated protein transport pathways. *BMC Evolutionary Biology*, *7*, 219.

Szabo, R. (1999). Dimorphism in Yarrowia lipolytica: filament formation is suppressed by nitrogen starvation and inhibition of respiration. *Folia Microbiologica*, *44*(1), 19–24.

Tenagy, Park, J. S., Iwama, R., Kobayashi, S., Ohta, A., Horiuchi, H., & Fukuda, R. (2015). Involvement of acyl-CoA synthetase genes in n-alkane assimilation and fatty acid utilization in yeast Yarrowia lipolytica. *FEMS Yeast Research*, *15*(4), fov031.

Thomas, L., Larroche, C., & Pandey, A. (2013a). Current developments in solid-state fermentation. *Biochemical Engineering Journal*, *81*, 146–161.

Thomas, L., Larroche, C., & Pandey, A. (2013b). Current developments in solid-state fermentation. *Biochemical Engineering Journal*, *81*, 146–161.

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., … Adams, D. J. (2021). Combinatorial CRISPR screen identifies fitness effects of gene paralogues. *Nature Communications*, *12*(1), 1–11.

Thorwall, S., Schwartz, C., Chartron, J. W., & Wheeldon, I. (2020). Stress-tolerant non-conventional microbes enable next-generation chemical biosynthesis. *Nature Chemical Biology*, *16*(2), 113–121.

Tirnauer, J. S., & Bierer, B. E. (2000). EB1 proteins regulate microtubule dynamics, cell polarity, and chromosome stability. *The Journal of Cell Biology*, *149*(4), 761–766.

Tisi, R., Belotti, F., & Martegani, E. (2014). Yeast as a model for Ras signalling. *Methods in Molecular Biology , 1120*, 359–390.

Tkachenko, A. A., Borshchevskaya, L. N., Sineoky, S. P., & Gordeeva, T. L. (2023). CRISPR/Cas9-mediated genome editing of the Komagataella phaffii to obtain a phytase-producer markerless strain. *Biochemistry. Biokhimiia*, *88*(9), 1338–1346.

Trivedi, V., Ramesh, A., & Wheeldon, I. (2023). Analyzing CRISPR screens in non-conventional microbes. *Journal of Industrial Microbiology & Biotechnology*, *50*(1). 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 Biochemistry*, *64*, 9–15.

Turki, S., Kraeim, I. B., Weeckers, F., Thonart, P., & Kallel, H. (2009). Isolation of bioactive peptides from tryptone that modulate lipase production in Yarrowia lipolytica. *Bioresource Technology*, *100*(10), 2724–2731.

Uhl, M. A., Biery, M., Craig, N., & Johnson, A. D. (2003). Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen C.albicans. *The EMBO Journal*, *22*(11), 2668–2678.

U.S. Government Accountability Office. (n.d.). *Science & Tech Spotlight: Synthetic Biology*. Retrieved August 11, 2024, from https://www.gao.gov/products/gao-23-106648

Valásek, L., Hasek, J., Trachsel, H., Imre, E. M., & Ruis, H. (1999). The Saccharomyces cerevisiae HCR1 gene encoding a homologue of the p35 subunit of human translation initiation factor 3 (eIF3) is a high copy suppressor of a temperature-sensitive mutation in the Rpg1p subunit of yeast eIF3. *The Journal of Biological Chemistry*, *274*(39), 27567–27572.

Vallejo, J. A., Sánchez-Pérez, A., Martínez, J. P., & Villa, T. G. (2013). Cell aggregations in yeasts and their applications. *Applied Microbiology and Biotechnology*, *97*(6), 2305–2318.

Vandenbosch, D., De Canck, E., Dhondt, I., Rigole, P., Nelis, H. J., & Coenye, T. (2013). Genomewide screening for genes involved in biofilm formation and miconazole susceptibility in Saccharomyces cerevisiae. *FEMS Yeast Research*, *13*(8), 720–730.

Vandermies, M., & Fickers, P. (2019). Bioreactor-Scale Strategies for the Production of Recombinant Protein in the Yeast Yarrowia lipolytica. *Microorganisms*, *7*(2). https://doi.org/10.3390/microorganisms7020040

Vastrad, B., & Neelagund, S. (2012). *Optimization of process parameters for rifamycin b production under solid state fermentation from Amycolatopsis Mediterranean mtcc 14*. https://www.academia.edu/download/27338718/vastrad_4_th_paper.pdf

Verma, N., & Kumar, V. (2020). Impact of process parameters and plant polysaccharide hydrolysates in cellulase production by Trichoderma reesei and Neurospora crassa under wheat bran based solid state fermentation. *Biotechnology Reports*, *25*, e00416.

Viniegra-González, G., & Favela-Torres, E. (2006). *Why solid-state fermentation seems to be resistant to catabolite repression?* https://hrcak.srce.hr/clanak/162057

Viniegra-González, Gustavo, Favela-Torres, E., Aguilar, C. N., Rómero-Gomez, S. de J., Díaz-Godínez, G., & Augur, C. (2003). Advantages of fungal enzyme production in solid state over liquid fermentation systems. *Biochemical Engineering Journal*, *13*(2), 157–167.

Wang, F., Xu, L., Zhao, L., Ding, Z., Ma, H., & Terry, N. (2019). Fungal Laccase Production from Lignocellulosic Agricultural Wastes by Solid-State Fermentation: A Review. *Microorganisms*, *7*(12). https://doi.org/10.3390/microorganisms7120665

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 recombineering for improved protein production in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(19), 9324–9332.

Wang, K., Shi, T.-Q., Wang, J., Wei, P., Ledesma-Amaro, R., & Ji, X.-J. (2022). Engineering the Lipid and Fatty Acid Metabolism in Yarrowia lipolytica for Sustainable Production of High Oleic Oils. *ACS Synthetic Biology*, *11*(4), 1542–1554.

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. *Applied Microbiology and Biotechnology*, *104*(24), 10481–10491.

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., El Bakkoury, M., Foury, F., Friend, S. H., Gentalen, E., … Davis, R. W. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science*, *285*(5429), 901–906.

Worland, A. M., Czajka, J. J., Li, Y., Wang, Y., Tang, Y. J., & Su, W. W. (2020). Biosynthesis of terpene compounds using the non-model yeast Yarrowia lipolytica: grand challenges and a few perspectives. *Current Opinion in Biotechnology*, *64*, 134–140.

Xiaoyan, L., Yu, X., Lv, J., Xu, J., Xia, J., Wu, Z., Zhang, T., & Deng, Y. (2017). A cost-effective process for the coproduction of erythritol and lipase with Yarrowia lipolytica M53 from waste cooking oil. *Food and Bioproducts Processing*, *103*, 86–94.

Xu, D., Jiang, B., Ketela, T., Lemieux, S., Veillette, K., Martel, N., Davison, J., Sillaots, S., Trosok, S., Bachewich, C., Bussey, H., Youngman, P., & Roemer, T. (2007). Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in Candida albicans. *PLoS Pathogens*, *3*(6), e92.

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. *Biotechnology for Biofuels*, *12*, 61.

Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., Damude, H. G., Rupert, R. A., Seip, J. E., Wang, J., Pollak, D. W., Bostick, M. W., Bosak, M. D., Macool, D. J., Hollerbach, D. H., Zhang, H., Arcilla, D. M., Bledsoe, S. A., Croker, K., … Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–740.

Yu, M., Qin, S., & Tan, T. (2007). Purification and characterization of the extracellular lipase Lip2 from Yarrowia lipolytica. *Process Biochemistry*, *42*(3), 384–391.

Zacharioudakis, I., Papagiannidis, D., Gounalaki, N., Stratidaki, I., Kafetzopoulos, D., & Tzamarias, D. (2017). Ras mutants enhance the ability of cells to anticipate future lethal stressors. *Biochemical and Biophysical Research Communications*, *482*(4), 1278–1283.

Zhu, Q., & Jackson, E. N. (2015). Metabolic engineering of Yarrowia lipolytica for industrial applications. *Current Opinion in Biotechnology*, *36*, 65–72.

Ziganshin, A. M., Naumova, R. P., Pannier, A. J., & Gerlach, R. (2010). Influence of pH on 2,4,6-trinitrotoluene degradation by Yarrowia lipolytica. *Chemosphere*, *79*(4), 426–433.

Żogała, B., Robak, M., Rymowicz, W., Wzientek, K., Rusin, M., & Maruszczak, J. (2005). Geoelectrical Observation of Yarrowia lipolytica Bioremediation of Petrol-Contaminated Soil. *Polish Journal of Environmental Studies*, *14*. https://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=12301485&AN=18604882&h=cPTCA44BT3wiR95vxdI2uFONsFRgV8Xa541Dc2suUAFNjBGYM4E%2FeQfaWs7qvioRb216nuQqSsanXamVqyHBcg%3D%3D&crl=c

Zöller, E., Laborenz, J., Krämer, L., Boos, F., Räschle, M., Alexander, R. T., & Herrmann, J. M. (2020). The intermembrane space protein Mix23 is a novel stress-induced mitochondrial import factor. *The Journal of Biological Chemistry*, *295*(43), 14686–14697.

Supplemental Materials

Tables and Files

Supplemental Table S1: Strains used in all chapters

Strains	Description	Reference
PO1f	*MatA, leu2-270, ura3-302, xpr2-322, axp1-2*	Madzak et al 2000
PO1f HMEBI	PO1f *CrtI*::*XPR2*, *CrtB*::*AXP*, *CrtE*::*A08*, *HMG1*::*D17*, *HMG1*::*XDH*, *MVD1*::*LEU2*	Schwartz et al 2017
PO1f *CAS9*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*	Schwartz et al 2019
PO1f *CAS9 ΔKU70*	PO1f UAS1B8-TEF(136)-Cas9::A08, ku70 KO	Schwartz et al 2019
PO1f *CAS9 ΔRAS2*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*, ΔYALI1_E35305g	Lupish et al 2022
PO1f *CAS9 ΔSFL1*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*, ΔYALI1_D05956g	Lupish et al 2022
PO1f *CAS9 ΔRHO5*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*, ΔYALI1_E30639g	Lupish et al 2022
PO1f *CAS9 ΔPAXIP1*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*, ΔYALI1_F04690g	Lupish et al 2022

PO1f *CAS9 ΔSNF2*	PO1f UAS1B8-TEF(136)- *CAS9*::*A08*, ΔYALI1_D30097g	Lupish et al 2022
PO1f HMEBI *ΔRAS2*	PO1f *CrtI*::*XPR2*, *CrtB*::*AXP*, *CrtE*::*A08*, *HMG1*::*D17*, *HMG1*::*XDH*, *MVD1*::*LEU2* ΔYALI1_E35305g	Lupish et al 2022
PO1f HMEBI *ΔSFL1*	PO1f *CrtI*::*XPR2*, *CrtB*::*AXP*, *CrtE*::*A08*, *HMG1*::*D17*, *HMG1*::*XDH*, *MVD1*::*LEU2* ΔYALI1_D05956g	Lupish et al 2022
PO1f HMEBI *ΔRHO5*	PO1f *CrtI*::*XPR2*, *CrtB*::*AXP*, *CrtE*::*A08*, *HMG1*::*D17*, *HMG1*::*XDH*, *MVD1*::*LEU2* ΔYALI1_E30639g	Lupish et al 2022
PO1f HMEBI *ΔMHY1*	PO1f *CrtI*::*XPR2*, *CrtB*::*AXP*, *CrtE*::*A08*, *HMG1*::*D17*, *HMG1*::*XDH*, *MVD1*::*LEU2* ΔYALI1_B28150g	Lupish et al 2022
sNR004	PO1f ΔD05956g	Robertson et al 2024
sNR005	PO1f ΔE22029g	Robertson et al 2024
sNR006	PO1f ΔE13627g	Robertson et al 2024
sNR007	PO1f ΔC02904g	Robertson et al 2024
sNR008	PO1f ΔE01193g	Robertson et al 2024
sNR009	PO1f ΔC33332g	Robertson et al 2024

sNR010	PO1f ΔE24428g	Robertson et al 2024
sNR011	PO1f ΔC30884g	Robertson et al 2024
sNR012	PO1f ΔE37234g	Robertson et al 2024
sNR013	PO1f ΔF23331g	Robertson et al 2024
sNR014	PO1f ΔD01350g	Robertson et al 2024
sNR015	PO1f ΔD21022g	Robertson et al 2024
sNR016	PO1f ΔC23902g	Robertson et al 2024
sNR017	PO1f ΔD12493g	Robertson et al 2024
sNR018	PO1f ΔF11871g	Robertson et al 2024
sNR019	PO1f ΔF19899g	Robertson et al 2024
sNR022	PO1f ΔC02904g, ΔE01193g	Robertson et al 2024

sNR023	PO1f ΔC02904g, ΔC33332g	Robertson et al 2024
sNR024	PO1f ΔC02904g, ΔE24428g	Robertson et al 2024
sNR025	PO1f ΔC02904g, ΔE37234g	Robertson et al 2024
sNR026	PO1f ΔC02904g, ΔD21022g	Robertson et al 2024
sNR027	PO1f ΔE01193g, ΔC33332g	Robertson et al 2024
sNR028	PO1f ΔE01193g, ΔE24428g	Robertson et al 2024
sNR029	PO1f ΔE01193g, ΔE37234g	Robertson et al 2024
sNR030	PO1f ΔE01193g, ΔD21022g	Robertson et al 2024
sNR031	PO1f ΔC33332g, ΔE24428g	Robertson et al 2024
sNR032	PO1f ΔC33332g, ΔE37234g	Robertson et al 2024
sNR033	PO1f ΔE24428g, ΔE37234g	Robertson et al 2024

sNR034	PO1f ΔE24428g, ΔD21022g	Robertson et al 2024
sNR035	PO1f ΔE37234g, ΔD21022g	Robertson et al 2024

Supplemental Table S2: Plasmids used in all chapters

Plasmids	Description	Reference (Addgene #)
pCRISPRyl	UAS1B8-TEF(136)-*CAS9*-CycT and SRC1'-tRNAgly-sgRNA	Schwartz et al 2016 (70007)
pCRISPRyl_*RAS2*	pCRISPRyl with YALI1_E35305g targeting sgRNA	This Study
pCRISPRyl_*SFL1*	pCRISPRyl with YALI1_D05956g targeting sgRNA	This Study
pCRISPRyl_*RHO5*	pCRISPRyl with YALI1_E30639g targeting sgRNA	This Study
pCRISPRyl_*MHY1*	pCRISPRyl with YALI1_B28150g targeting sgRNA	This Study
pYL*hrGFP*	UAS1B8-TEF(136)-*hrGFP*	Schwartz et al 2016
p*RAS2*	UAS1B8-TEF(136)-YALI1_E35305g	This Study
p*LEU2-270*	pLEU2-*LEU2*-tLEU2	Blazeck et al 2011
pIW715	Homology donor plasmid for GFP integration at A08 (URA)	Addgene #84615
pIW1009	Homology donor plasmid for Cas9 integration at A08 (URA)	Robertson et al 2024

pIW524	Cas9 and sgRNA to cut at A08 (LEU)	Robertson et al 2024
pIW386	Easyclone Y. lipolytica Cas9 cutter without gRNA (LEU)	Robertson et al 2024
pIW363	pIW386 with gRNA for ku70 knockdown (LEU)	Robertson et al 2024
pSC012	pIW386 with URA instead of LEU marker	Robertson et al 2024
pNR035	D05956g KO gRNAin pSC012	Robertson et al 2024
pNR036	E22029g KO gRNAin pSC012	Robertson et al 2024
pNR037	E13627g KO gRNAin pSC012	Robertson et al 2024
pNR038	C02904g KO gRNAin pSC012	Robertson et al 2024
pNR039	E01193g KO gRNAin pSC012	Robertson et al 2024
pNR040	C33332g KO gRNAin pSC012	Robertson et al 2024
pNR041	E24428g KO gRNAin pSC012	Robertson et al 2024
pNR042	C30884g KO gRNAin pSC012	Robertson et al 2024
pNR043	E37234g KO gRNAin pSC012	Robertson et al 2024
pNR044	F23331g KO gRNAin pSC012	Robertson et al 2024

pNR045	D01350g KO gRNAin pSC012	Robertson et al 2024
pNR046	D21022g KO gRNAin pSC012	Robertson et al 2024
pNR047	C23902g KO gRNAin pSC012	Robertson et al 2024
pNR048	D12493g KO gRNAin pSC012	Robertson et al 2024
pNR049	F11871g KO gRNAin pSC012	Robertson et al 2024
pNR050	F19899g KO gRNAin pSC012	Robertson et al 2024

Supplemental Table S3: Primers and sgRNAs used in all Chapters

Primer	Sequence	Use
*RAS2*_sgRNA.fwd	GGG TCG GCG CAG GTT GAC GTT TGA AGA CTC ATA CTC CAC CGT TTT AGA GCT AGA AAT AGC	sgRNA insert of pCRISPRyl_*RAS2*
*RAS2*_sgRNA.rev	GCT ATT TCT AGC TCT AAA ACG GTG GAG TAT GAG TCT TCA AAC GTC AAC CTG CGC CGA CCC	sgRNA insert of pCRISPRyl_*RAS2*
*SFL1*_sgRNA.fwd	GGG TCG GCG CAG GTT GAC GTG TGT GTG TCT GGT TAT GTG TGT TTT AGA GCT AGA AAT AGC	sgRNA insert of pCRISPRyl_*SFL1*
*SFL1*_sgRNA.rev	GCT ATT TCT AGC TCT AAA ACA CAC ATA ACC AGA CAC ACA CAC GTC AAC CTG CGC CGA CCC	sgRNA insert of pCRISPRyl_*SFL1*
*RHO5*_sgRNA.fwd	GGG TCG GCG CAG GTT GAC GTT TGA CGT TCT CGA ACG ACG GGT TTT AGA GCT AGA AAT AGC	sgRNA insert of pCRISPRyl_*RHO5*
*RHO5*_sgRNA.rev	GCT ATT TCT AGC TCT AAA ACC CGT CGT TCG AGA ACG TCA AAC GTC AAC CTG CGC CGA CCC	sgRNA insert of pCRISPRyl_*RHO5*
*MHY1*_sgRNA.fwd	GGG TCG GCG CAG GTT GAC GTG GAG ATG CGC AGA TAA CGG AGT TTT AGA GCT AGA AAT AGC	sgRNA insert of pCRISPRyl_*MHY1*
*MHY1*_sgRNA.rev	GCT ATT TCT AGC TCT AAA ACT CCG TTA TCT CGC CAT CTC CAC GTC AAC CTG CGC CGA CCC	sgRNA insert of pCRISPRyl_*MHY1*
*RAS2*_ORF.fwd	GCG CGC GGC GCG CCA TGA GTG AAC AAC CCC AGC AAA AGG TTT CCA TTG TC	Restriction Assembly of p*RAS2*

*RAS2*_OR F.rev	GCG CGC GGC TAG CCT AGC AGA TAA CAC AGC ACT TGG ATC CAC CCT TGA TG	Restriction Assembly of p*RAS2*
A_FOR	CCACCGAGCCGTTCATG	D05956g forward primer
B_FOR	GTCCTGGAACGCCATCAG	E22029g forward primer
C_FOR	GCGTCTGAAGCTCGCTAATATCG	D16431g forward primer
D_FOR	CCGAGTGTAGGCCACTTG	E13627g forward primer
E_FOR	GTTAGACAGCACCAGGGTG	C02904g forward primer
F_FOR	GTTCTTGTCGTTGCAGACTCG	E01193g forward primer
G_FOR	CACTCACTTGCCACTGCAG	C33332g forward primer
H_FOR	CGTGATGGAGACTGGGGAG	E24428g forward primer
I_FOR	CCCGACTCTTCGTCTTCATCG	C30884g forward primer
J_FOR	GGTATGAATTCTGGCCCAAACTG	E37234g forward primer
K_FOR	GGCCTTCTCAGACAAGTCGG	F23331g forward primer
L_FOR	GTGATTGGGGTGTTAGGTCG	D01350g forward primer

M_FOR	GCATAAGTCTCAGAGCCAGC	D21022g forward primer
A_REV	CCAGACACACACACAGCC	D05956g reverse primer
B_REV	GCAACAACTTCCGAACTGCTG	E22029g reverse primer
C_REV	CCGTGTGGGACAATCTCTTTTAC	D16431g reverse primer
D_REV	GGAGAGAAAGACAGCGCTTTGC	E13627g reverse primer
E_REV	CTTTCCTCCAGACTTTTCTCCTTCC	C02904g reverse primer
F_REV	GGTTGGAGGGAATCGCG	E01193g reverse primer
G_REV	GTGGTAGTGGGCGAACTG	C33332g reverse primer
H_REV	CCTCCTACATTGCGCATGG	E24428g reverse primer
I_REV	CGAGTCGGCACTGAAGG	C30884g reverse primer
J_REV	GCATCTTGTGTCTGTAGAACCG	E37234g reverse primer
K_REV	GCTCACAGACACCTCTTGTG	F23331g reverse primer
L_REV	CGTTCGTCTGCACACACC	D01350g reverse primer
M_REV	CAACGACGTTGGGTGGC	D21022g reverse primer

neg1FOR	GATCAGAGTCAGGATGGGTGAC	C23902g forward primer
neg2FOR	GATAACGCCGTTCCACGC	D17181g forward primer
neg3FOR	CCTCGTTGCGATCCATTAACC	D12493g forward primer
neg4FOR	GTTGAGCTCATCCAGCTCATG	F21690g forward primer
neg5FOR	CTAATGCTGTCAAAACGGATAGCG	B06015g forward primer
neg6FOR	GCACCATGGTGGTAGGTG	B26350g forward primer
neg7FOR	GCTCAATCCACCACAAGATCAAATC	F11871g forward primer
neg8FOR	CGAGGCATTACCTTTGGAGG	D11769g forward primer
neg9FOR	CAGCAGTAGCCCCAACAC	A22093g forward primer
neg10FOR	CTCCACACTTGGCAGTGG	A21156g forward primer
neg11FOR	GCACTCGTCTTTGAGTCTCAC	F07819g forward primer
neg12FOR	CCAGAACCGTTATTGGTGCC	B05364g forward primer
neg13FOR	CGTCACAGGTATCGGCAAC	D14893g forward primer
neg14FOR	GATCTTCTTGCCGTCGGC	D15424g forward primer

neg15FOR	GTGGAAGAGGTTGTGGCC	A21797g forward primer
neg16FOR	GTCTTCCTTGTCCACCCG	F34749g forward primer
neg17FOR	CCCAACCTGTATTTCGGTGTC	F19899g forward primer
neg18FOR	GGGTATGAGAGGAATTTCGACG	A03183g forward primer
neg19FOR	GTACTGGCAGAAGAGCGC	D25330g forward primer
neg20FOR	CATCTTACCACCCTGAGTAAGTCC	C21487g forward primer
neg21FOR	GAGGCATCCGGTTCGTTC	F26726g forward primer
neg1REV	GCAAATCAATTCTTCGTGCACG	C23902g reverse primer
neg2REV	GTTGTAGATTGTCAGGACCATAATGG	D17181g reverse primer
neg3REV	CACCTGTTCCATGAGCGC	D12493g reverse primer
neg4REV	GAAGGATGTCTAGATGAACCACG	F21690g reverse primer
neg5REV	GATATCTCCGAGACGAGTTTTCCTC	B06015g reverse primer
neg6REV	CAACACTGACGTGACGTCC	B26350g reverse primer
neg7REV	CCACCAAGCTGCTTGGATAAG	F11871g reverse primer

neg8REV	GCTGAGTGTTAGAACAACTCGC	D11769g reverse primer
neg9REV	CACCAAAGAACTGTTTGGAGGG	A22093g reverse primer
neg10REV	CCGAGACCGAATCCTCGAG	A21156g reverse primer
neg11REV	GCGAAGAAACCAGTGCCTG	F07819g reverse primer
neg12REV	GGAACTTAGTAGCCTCGATAGGTC	B05364g reverse primer
neg13REV	CGTTGTCGACGAGGAATCC	D14893g reverse primer
neg14REV	GTCTGGTGATGTTGTCTTGGTG	D15424g reverse primer
neg15REV	GTTGGTCTTAGTGCTGACGC	A21797g reverse primer
neg16REV	CAGTCATTCGCTTTCGAGATGAAG	F34749g reverse primer
neg17REV	CAATTCGCTCCTGCAGAAGG	F19899g reverse primer
neg18REV	GACGTTGAACACGGAAATGCC	A03183g reverse primer
neg19REV	CAGCTCCAACGGGATATTGC	D25330g reverse primer
neg20REV	CGTAACCTCCAGGAGGTTCTC	C21487g reverse primer
neg21REV	GGTTGAACAGCGCGTCC	F26726g reverse primer

qPCR-GW-F	TTATGAACTGAAAGTTGATGGC	qPCR forward primer
qPCR-GW-R	TCACACAGGAAACAGCTATG	qPCR reverse primer
Cr_1665	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTCCGGTTCGATTCCGGGTC	Forward primer Illumina
Cr_1666	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAGTCCGGTTCGATTCCGGGTC	Forward primer Illumina
Cr_1667	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCAGTAGTCCGGTTCGATTCCGGGTC	Forward primer Illumina
Cr_1668	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCAGTAGTCCGGTTCGATTCCGGGTC	Forward primer Illumina
Cr_1669	CAAGCAGAAGACGGCATACGAGATTCGCCTTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1670	CAAGCAGAAGACGGCATACGAGATATAGCGTCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1671	CAAGCAGAAGACGGCATACGAGATGAAGAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1672	CAAGCAGAAGACGGCATACGAGATATTCTAGGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1673	CAAGCAGAAGACGGCATACGAGATCGTTACCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1709	CAAGCAGAAGACGGCATACGAGATGTCTGATGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1710	CAAGCAGAAGACGGCATACGAGATTTACGCACGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina
Cr_1711	CAAGCAGAAGACGGCATACGAGATTTGAATAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCGACTCGGTGCCACTTTTTCAAG	Reverse primer Illumina

sgRNA	Sequence	Strain Generated
RAS2	T TGA AGA CTC ATA CTC CAC C	PO1f HMEBI *ΔRAS2*
SFL1	G TGT GTG TCT GGT TAT GTG T	PO1f HMEBI *ΔSFL1*
RHO5	T TGA CGT TCT CGA ACG ACG G	PO1f HMEBI *ΔRHO5*
MHY1	G GAG ATG CGC AGA TAA CGG A	PO1f HMEBI *ΔMHY1*
YALI1_D05956g_2	CAACATGTACGGCTTCCATA	PO1f ΔD05956g
YALI1_E22029g_3	CACTCCTCTGTTGAGTGCGG	PO1f ΔE22029g
YALI1_D16431g_3	AACCATGACAAATCTGCTCA	PO1f ΔD16431g
YALI1_E13627g_3	TGGATGTGTGTGACGGAAAG	PO1f ΔE13627g
YALI1_C02904g_3	AAAGCCGAAGATGGGGCCCG	PO1f ΔC02904g
YALI1_E01193g_3	CTCTCGGTTTCATTCCCGAG	PO1f ΔE01193g

YALI1_C33332g_2	CGGAAGATATGACGATAAAG	PO1fΔC33332g
YALI1_E24428g_3	GTGTGTCTTGAGGTCCCGCT	PO1fΔE24428g
YALI1_C30884g_2	GCAAGAGCCGAGTCAGCGCA	PO1fΔC30884g
YALI1_E37234g_1	ATTGCAGGTTAGAAATGGGG	PO1fΔE37234g
YALI1_F23331g_2	CCTTGAGCTTGAGACCCTGC	PO1fΔF23331g
YALI1_D01350g_1	AGTTGTCAGTGGCAAGGTAG	PO1fΔD01350g
YALI1_D21022g_1	GATGCTCCGACGAGGCCTGC	PO1fΔD21022g
YALI1_C23902g_1	CGATGACTCTGGGCACGAGG	PO1fΔC23902g
YALI1_D17181g_1	CTACTCGGTCTACAGACGAG	PO1fΔD17181g
YALI1_D12493g_3	GGAGGGCGAAGAGAGGGTCG	PO1fΔD12493g
YALI1_F21690g_2	GTACTCGGTGGACGACAGTT	PO1fΔF21690g

YALI1_B06015g_1	TCAGCTAAACCCATGTCAAA	PO1fΔB06015g
YALI1_B26350g_3	ATGATGGCAATGATCACGGG	PO1fΔB26350g
YALI1_F11871g_2	CCACCGACTTGAGAATGCCT	PO1fΔF11871g
YALI1_D11769g_2	GGTGTCATCAACCGACACTC	PO1fΔD11769g
YALI1_A22093g_3	TCTCCCAGACAGGTCCAGAT	PO1fΔA22093g
YALI1_A21156g_1	GTATCGTGCCTTAGGCCAGG	PO1fΔA21156g
YALI1_F07819g_3	GAAACCCGCAAAAACCTCCA	PO1fΔF07819g
YALI1_B05364g_1	TCAGAGCGTTAAGAATAGCG	PO1fΔB05364g
YALI1_D14893g_3	CGCCTCCGTGTGCTACATCT	PO1fΔD14893g
YALI1_D15424g_3	ATCCTTGATGGAGAGCTTGT	PO1fΔD15424g
YALI1_A21797g_1	GCAGAGCCTGGAAGAACCCA	PO1fΔA21797g

YALI1_F34749g_3	CCTGGGCGAGGAGAGCGATG	PO1fΔF34749g
YALI1_F19899g_1	AAGACCAGAGTAGAACACCA	PO1fΔF19899g
YALI1_A03183g_1	GAGCGACTAGGCACTCTCGA	PO1fΔA03183g
YALI1_D25330g_1	GCAGATACGACAGCTCTGAG	PO1fΔD25330g
YALI1_C21487g_3	TGTACTCGTAGTACTGCACT	PO1fΔC21487g
YALI1_F26726g_2	ATATTGCACAAAGTGGACCA	PO1fΔF26726g

Supplemental Table S4: Galaxy Parameters for NGS reads

Tool	Version	Parameters*
FastQC	v0.11.8	Default settings
Cutadapt	Galaxy Version 1.16.6 [13]	Cutadapt was used to demultiplex samples containing the same Illumina barcode, but different pseudobarcodes at the 5' end of the read. Samples were amplified with reverse primers Cr1669-1673;Cr1709-1711 and forward primers Cr1665-1668 each containing a different pseudo barcode as mentioned in Table ▪ 5' (Front) anchored 6 bp pseudo-barcodes to be demultiplexed (-g): ^NNNNNN (refer to previous table for pseudo-barcode-forward primer association). ▪ Maximum error rate (--error-rate): 0.2 ▪ Match times (--times): 1 ▪ Minimum overlap length (--overlap): 4 ▪ Multiple output: Yes (Each demultiplexed readset is written to a separate file)
Trimmomatic	v0.38	▪ HEADCROP: 30 (if amplified by Cr1665); or 32 (if amplified by Cr1666); or 34 (if amplified by Cr1667); or 36 (if amplified by Cr1668) ▪ CROP: 20
Bowtie2**	v2.4.2	▪ Number of allowed mismatches in seed alignment (-N): 1 ▪ Length of the seed substring (-L): 20 ▪ Function governing interval between seed substrings in multiseed alignment (-i): S,1,0.50 ▪ Function governing maximum number of ambiguous characters (--n-ceil): L,0,0.15 ▪ Alignment mode: end-to-end ▪ Number of attempts of consecutive seed extension events (-D): 20 ▪ Number of times re-seeding occurs for repetitive reads: 3 ▪ Save mapping statistics: Yes

* All parameters other than those mentioned here are kept at default values.
** Bowtie2 usage needs a genome fasta file for alignment. Nontargeting sgRNA and any other sgRNA that Bowtie2 could not find within the original CLIB89 genome file were appended as an extra chromosome so that Bowtie could align all sgRNA for the purposes of generating counts.

Supplemental Table S5: Demultiplexing Barcode Primer Assignments

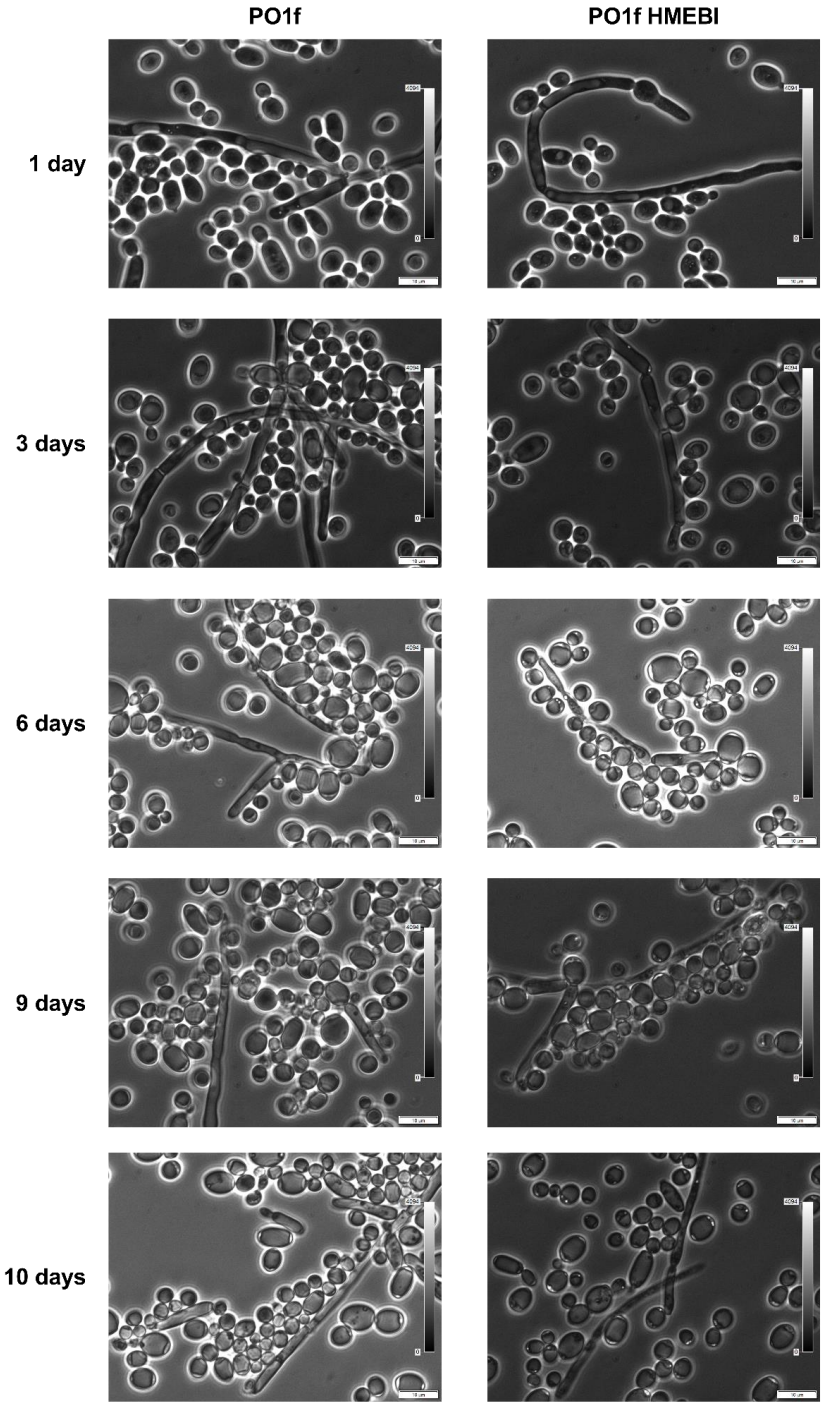
Sample Name	Forward Primer	Forward Barcode	Reverse Primer	Reverse Barcode
Library A	1665	AGTCCG	1710	GTGCGTAA
Library B	1666	GTAGTC	1710	GTGCGTAA
Control Glucose A	1665	AGTCCG	1669	CAAGGCGA
Control Glucose B	1666	GTAGTC	1669	CAAGGCGA
Control 0.25 M Acetate A	1667	CAGTAG	1670	GACGCTAT
Control 0.25 M Acetate B	1668	TCCAGT	1670	GACGCTAT
Control 0.5 M AcetateA	1665	AGTCCG	1671	ACTTCTTC
Control 0.5 M AcetateB	1666	GTAGTC	1671	ACTTCTTC
Cas9 Glucose A	1667	CAGTAG	1672	CCTAGAAT
Cas9 Glucose B	1668	TCCAGT	1672	CCTAGAAT
Cas9 0.25 M Acetate A	1665	AGTCCG	1673	TGGTAACG
Cas9 0.25 M Acetate B	1666	GTAGTC	1673	TGGTAACG

Cas9 0.5 M Acetate A	1667	CAGTAG	1709	CATCAGAC
Cas9 0.5 M Acetate B	1668	TCCAGT	1709	CATCAGAC
Cas9 Time 0 Control A	1667	CAGTAG	1669	CAAGGCGA
Cas9 Time 0 Control B	1668	TCCAGT	1669	CAAGGCGA
Solid-Liquid Screen Liquid Glucose A	1667	CAGTAG	1670	GACGCTAT
Solid-Liquid Screen Liquid Glucose A	1668	TCCAGT	1670	GACGCTAT
Solid-Liquid Screen Solid Glucose A	1665	AGTCCG	1672	CCTAGAAT
Solid-Liquid Screen Solid Glucose A	1666	GTAGTC	1672	CCTAGAAT
Dodecane A	1667	CAGTAG	1670	GACGCTAT
Dodecane B	1668	TCCAGT	1670	GACGCTAT
1-Dodecene A	1667	CAGTAG	1671	ACTTCTTC
1-Dodecene B	1668	TCCAGT	1671	ACTTCTTC
Oleic Acid A	1667	CAGTAG	1672	CCTAGAAT
Oleic Acid B	1668	TCCAGT	1672	CCTAGAAT

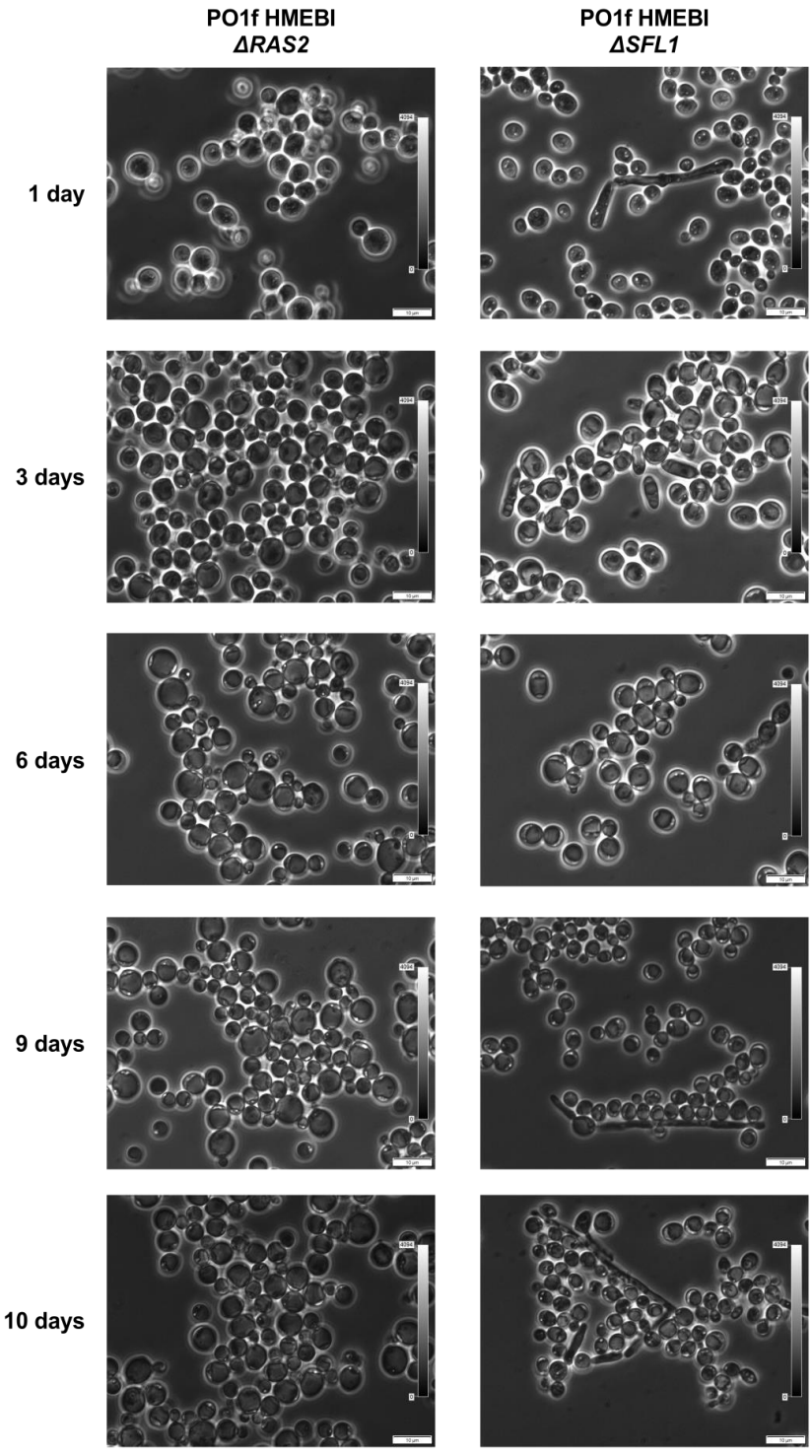
Margaric A	1667	CAGTAG	1673	TGGTAACG
Margaric B	1668	TCCAGT	1673	TGGTAACG

Supplemental File S6: Original sgRNA Library (Chapter 1)
See file: Supp.File 0.4_Original_sgRNA_Library_Chapter_1.xlsx
Note
Supplemental File S7: Refined sgRNA Library (Chapters 2 and 3)
See file: Supp.File 0.5_Refined_sgRNA_Library_Chapter_1.xlsx

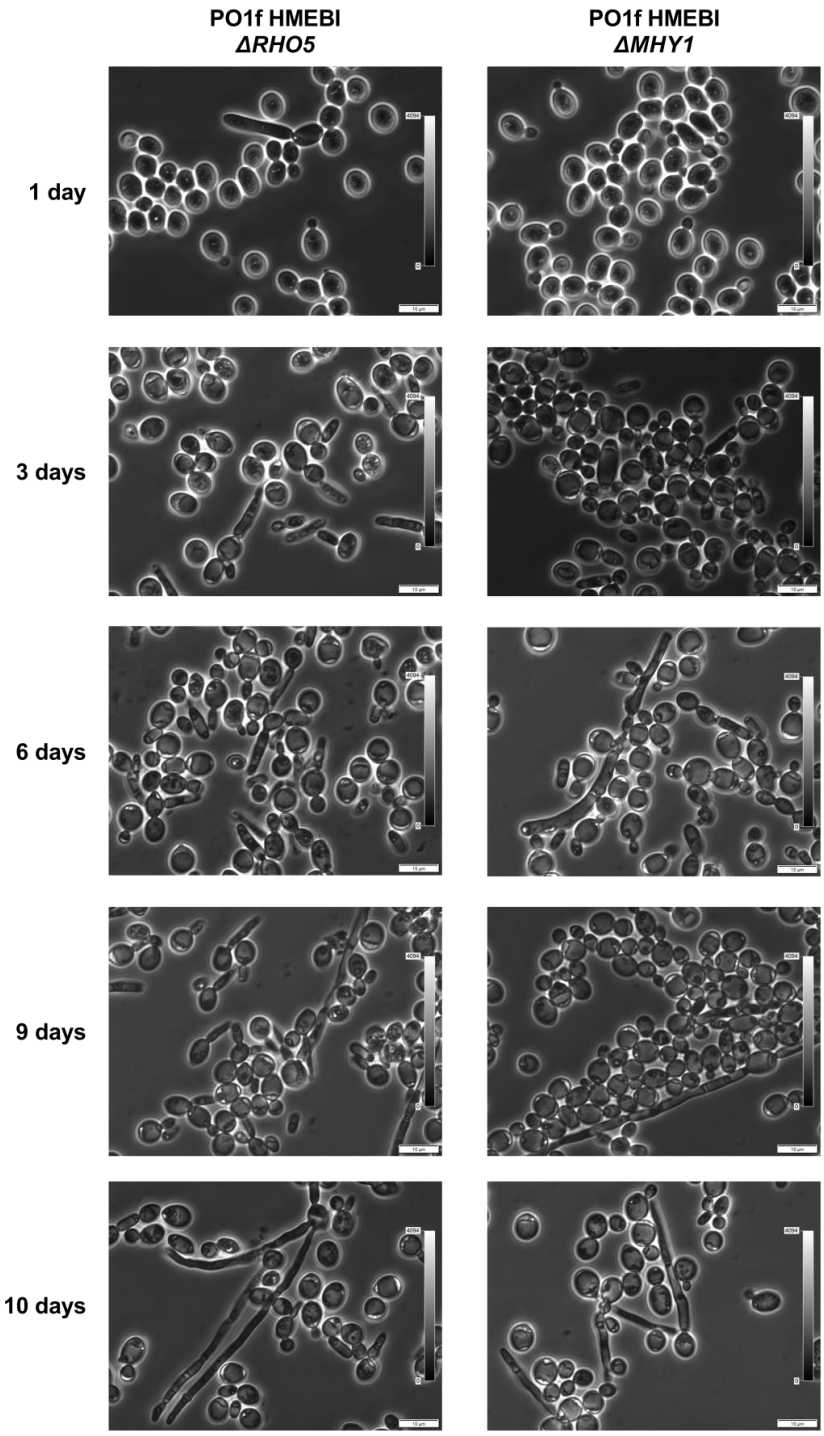
Chapter 1 Supplemental Figures



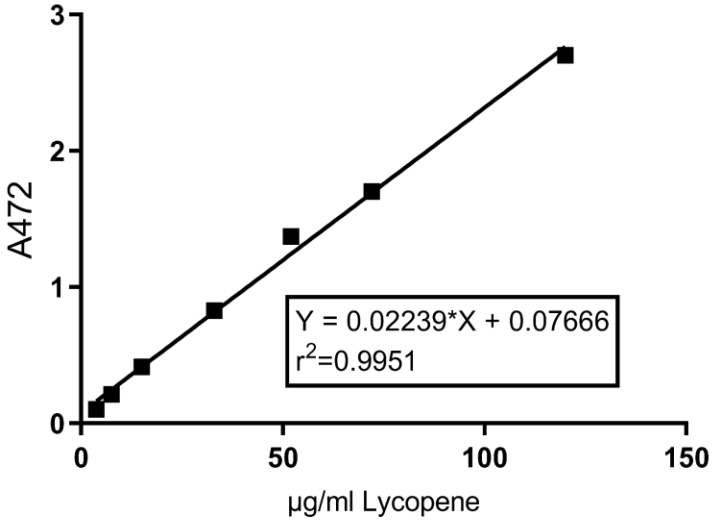
Supplemental Figure S1.1: Representative images of PO1f and PO1f HMEBI cells.



Supplemental Figure S1.2: Representative images PO1f *ΔRAS2* and PO1f *ΔSFL1* cells.

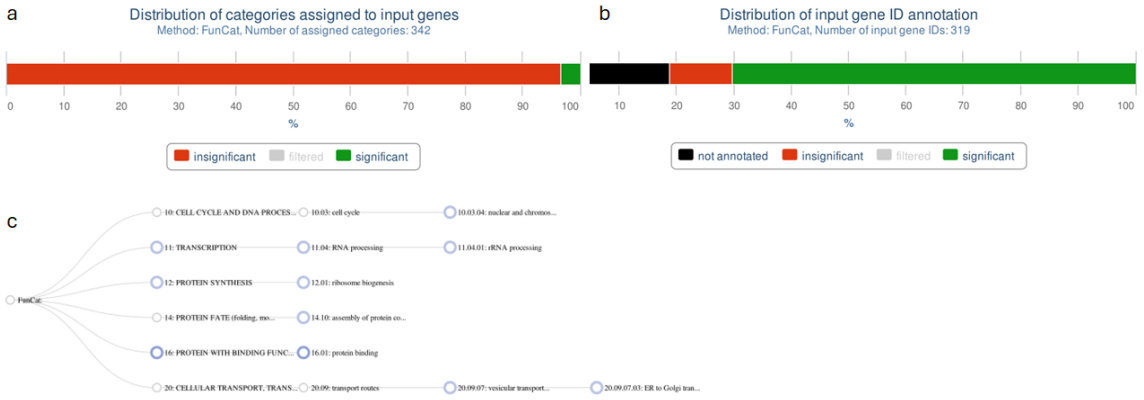


Supplemental Figure S1.3: Representative images of PO1f *ΔRHO5* and PO1f *ΔMHY1*.

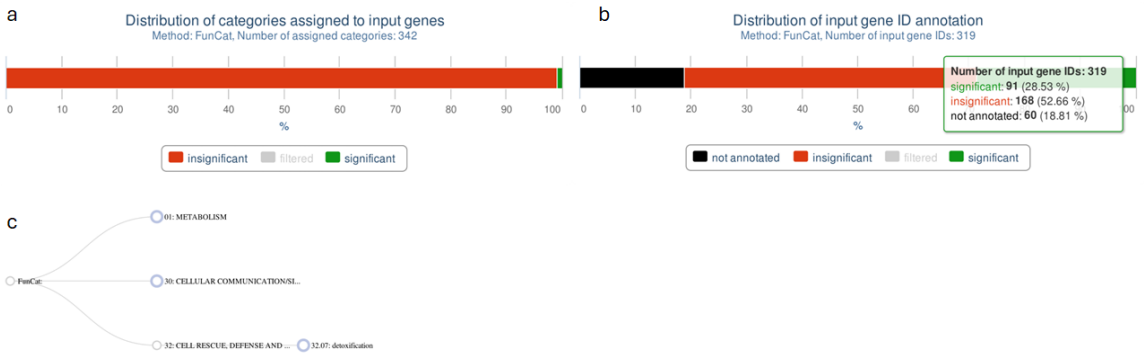


Supplemental Figure S1.4: Lycopene Standard Curve. A series of lycopene solutions (made with pre-purchased lycopene (Sigma-Aldrich)) was prepared at different predetermined concentrations (3.75, 7.50, 15.00, 33.00, 52.00, 72.00, and 120.00 μg/mL). Absorbance was then measured at 472 nm absorbance measurements. Linear regression was applied to calculate a standard curve formula, which was used to convert lycopene absorbance measurements into concentration values.

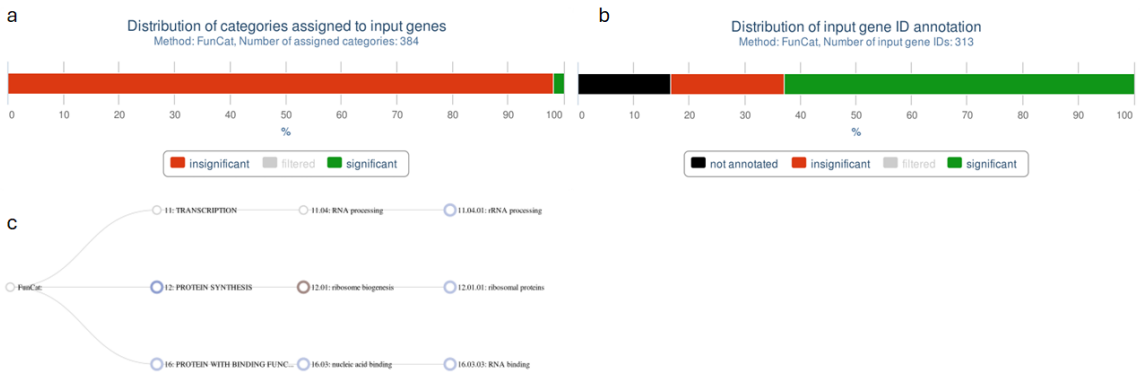
Chapter 2 Supplemental Figures



Supplemental Figure S2.1: FunCat Solid Enriched Screen Parameters – The output category summary from the FungiFun2 FunCat solid media essential gene enriched annotation analysis. **A.** The significance percentage of the distribution of the enriched FunCat categories assigned to the solid media essential genes. **B.** The significance percentage of the distribution of the solid media essential genes with enriched FunCat annotations. **C.** A FunCat hierarchy tree summarizing the hierarchy and relationships of the enriched FunCat annotations across the solid media essential genes. White rings indicate no significance, and blue and brown rings indicating significance ($p < 0.05$).



Supplemental Figure S2.2: FunCat Solid Depleted Screen Parameters – The output category summary from the FungiFun2 FunCat solid media essential gene depleted annotation analysis. **A.** The significance percentage of the distribution of the depleted FunCat categories assigned to the solid media essential genes. **B.** The significance percentage of the distribution of the solid media essential genes with depleted FunCat annotations. **C.** A FunCat hierarchy tree summarizing the hierarchy and relationships of the depleted FunCat annotations across the solid media essential genes. White rings indicate no significance, and blue and brown rings indicating significance ($p < 0.05$).



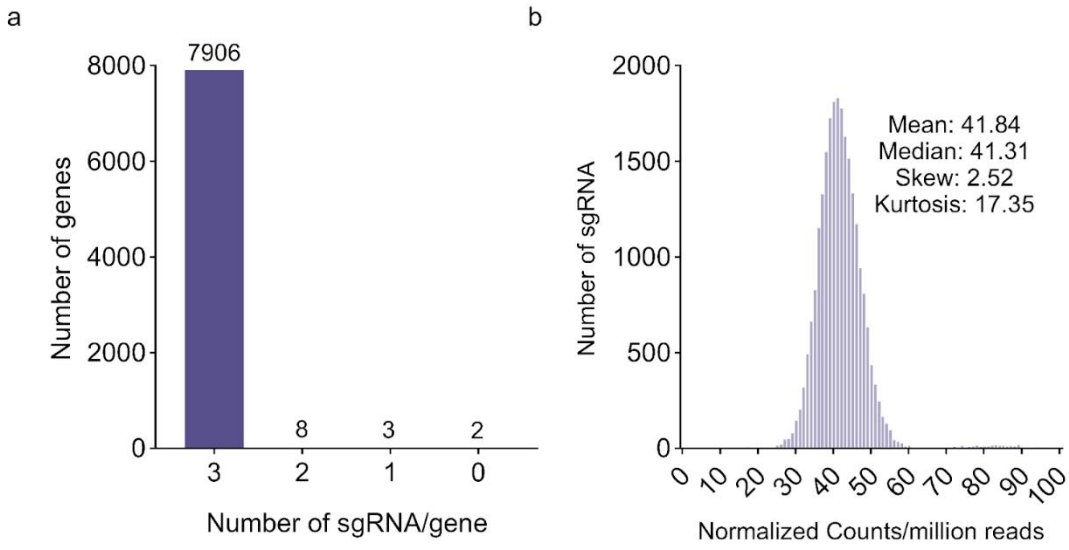
Supplemental Figure S2.3: FunCat Liquid Enriched Screen Parameters – The output category summary from the FungiFun2 FunCat liquid media essential gene enriched annotation analysis. **A.** The significance percentage of the distribution of the enriched FunCat categories assigned to the liquid media essential genes. **B.** The significance percentage of the distribution of the liquid media essential genes with enriched FunCat annotations. **C.** A FunCat hierarchy tree summarizing the hierarchy and relationships of the enriched FunCat annotations across the liquid media essential genes. White rings indicate no significance, and blue and brown rings indicating significance ($p < 0.05$).

Supplemental File S2.4: Solid Essential Hits
See file: Supp.File 2.2_Solid_Essential_Hit_Chart.xlsx

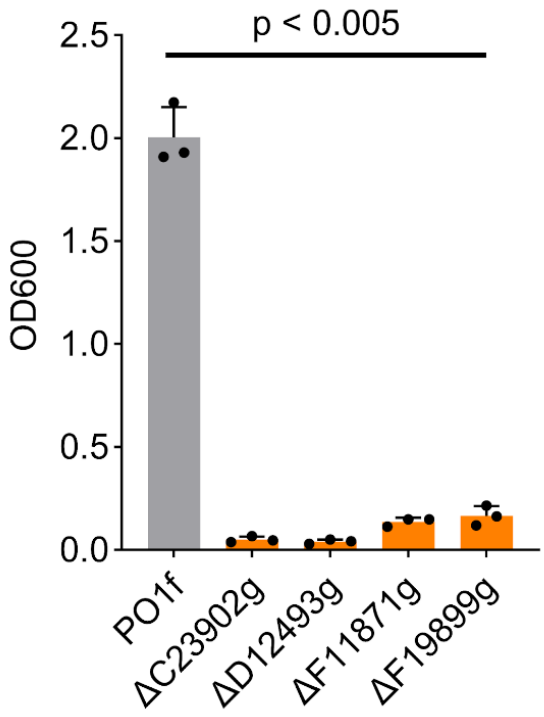
Supplemental File S2.5: Liquid Essential Hits
See file: Supp.File 2.3_Liquid_Essential_Hit_Chart.xlsx

Supplemental File S2.6: FunCat Essential Hit Enriched and Depleted Annotations
See file: Supp.File 2.4_GO_Term_Solid_Liquid_Differences_Chart.xlsx

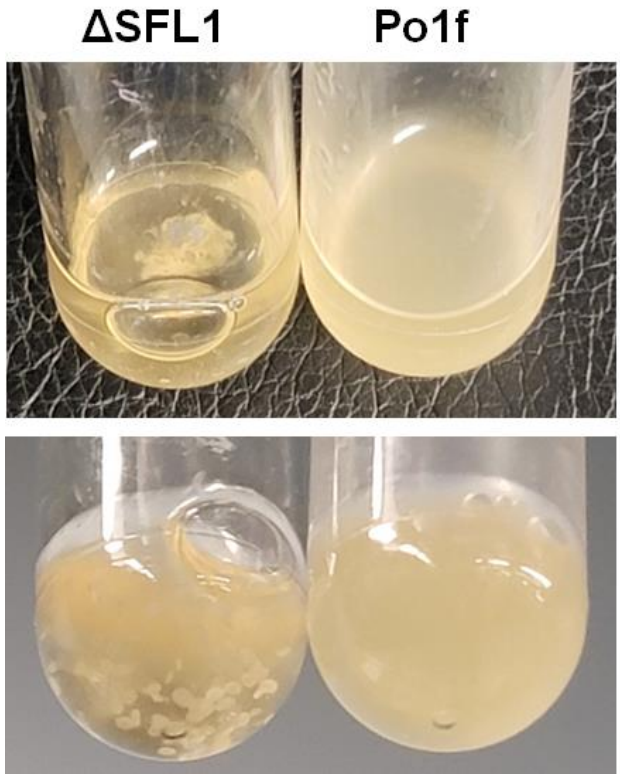
Chapter 3 Supplemental Figures



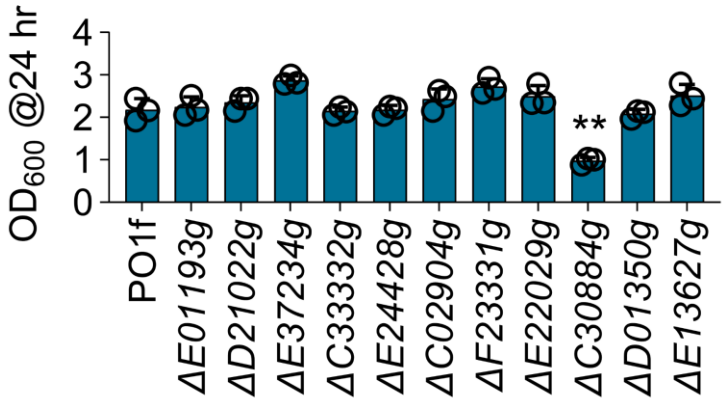
Supplemental Figure S3.1: Library v2 Characterization – Characteristics of lib. v2. (a) 99.8% of genes have three targeting guides. (b) Untransformed lib. v2 contains a tight distribution of guides.



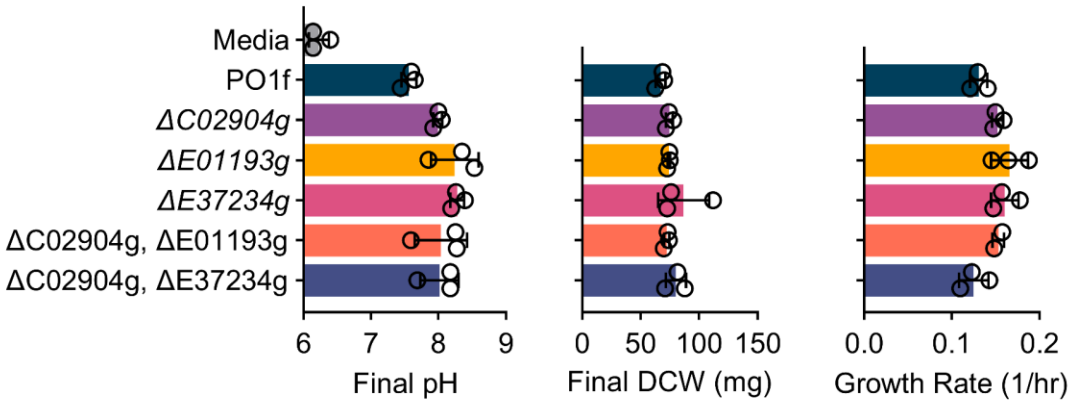
Supplemental Figure S3.2: Essential Hit Characterization – The four generated LOF hits were grown in 96-well plates with 1 mL of wells of 500 mM acetate in minimal media. All grew to significantly lower final culture densities than the wild type



Supplemental Figure S3.3: Top Acetate GOF Hit Characterization – Top acetate GOF hit encodes flocculation suppressing gene SFL1. The ΔSFL1 knockout causes *Y. lipolytica* to flocculate.



Supplemental Figure S3.4: Top Nine FS Acetate Hits in Glucose – Nine significant positive FS hits from the acetate screens grow similar to wild-type, except one. Corrected for multiple comparisons, only ΔC30884g grows worse than wild-type (** indicates $p<0.005$). Minimal media cultures with 0.125 M glucose were inoculated with overnight cultures to an OD600 of 0.05 and allowed to shake 24 hours at 1,000 RPM in 96-well plates containing 1 mL of media at 30 °C.



Supplemental Figure S3.5: Characterization of Top Validated Acetate Hits – Top acetate positive FS hits were characterized at the end of a growth curve performed in 500 mM acetate minimal media in 250 mL shake flasks 30 ℃, 220 RPM shaking, 5 days. Higher levels of growth cause the pH of the media to drop more as acetate is consumed. Ten mL of cell culture was collected at the end of 5 days, pelleted, lyophilized, and weighed to calculate DCW. Top positive FS hit knockouts grow more quickly and produce more biomass than the wild type. Growth rates were calculated from growth curve data from **Figure 3**. For each replicate, the growth rate for each time point was calculated for time points during the exponential phase, then averaged.

Supplemental File S3.6: Hit Significance and Details
See file: Supp.File 3.6_Hit_Significants_and_Details.xlsx

Supplemental File S3.7: All Guides CS and FS
See file: Supp. File 3.7_All_Guides_CS_and_FS