Graduate Research Statement

Introduction/Intellectual Merit: Advanced metabolic engineering allows scientists to use genetic engineering techniques in lower organisms, such as *Escherichia coli* and *Saccharomyces cerevisiae*, to produce molecules of interest through recombinant pathways. Metabolic pathways are groups of genes that encode enzymes that work together to produce the molecule of interest. Scientists have identified the most efficient genetic modification methodologies to create optimal production strains, including promoter libraries. A promoter library consists of a variation in the DNA sequence which varies the transcription initiation rate of the associated gene. This variation can come from promoters in the native organism or in non-native organisms, which are identified using RNAseq data. These libraries have been extensively developed for model organisms such as *E. coli* and *S. cerevisiae*. However, these tools are currently limited for non-conventional organisms. Leaders in the field of biochemical engineering have identified the development of genetic tools for use with non-conventional organisms as a foremost goal because model organisms lack the complexity that non-conventional ones can provide[1]. Leveraging unique properties of non-conventional organisms allows scientists to build upon promising results that push the boundaries of the pharmaceutical, environmental, and cosmetic industries.

Using the knowledge that I've gained from working in two metabolic engineering labs over the last three years, I intend to further explore the non-conventional oleaginous yeast, *Yarrowia lipolytica*. This yeast is particularly interesting because of its ability to naturally prevent bacterial contamination[3], its utilization of various hydrophobic and hydrophilic carbon sources[4], and its ability to efficiently produce large amounts of lipid-based products[5]. *Y. lipolytica* was originally identified in environments containing hydrophobic substrates and studied for its ability to biosynthetically produce enzymes and citric acids[2]. Recently, metabolic engineers have become more interested in this yeast as they explore non-conventional organisms for production of high-value compounds.

Research Plan: Although genetic engineers have made significant strides in the toolkits available for *Y. lipolytica*, they currently lack the diversity necessary to fully exploit its potential. One of the newest toolkits was developed by a group in France and is called the Golden Gate toolkit for *Y. lipolytica*[6]. The Golden Gate toolkit allows researchers to efficiently transform and integrate heterologous pathways in the organism. This toolkit includes a validated promoter library and has been successfully used to express a functional xylose utilization pathway. A major goal of my research plan is utilizing this toolkit and others to integrate several metabolic pathways into the organism. Additionally, scientists have discovered more promoters and have begun to look into computational models for the organism. These promoters include TATA box promoters[7] and one repressible and one bidirectional promoter[2]. Repressible promoters provide negative feedback to down regulate transcription pathways, which is useful for products that are inhibitory to growth. A bidirectional promoter can be initiated in both directions for transcription; this becomes important for efficient gene co-expression. However, in order to further enable the metabolic engineering goals for this non-conventional yeast, more native and non-native promoters need to be identified so that researchers can better optimize the genetic conditions for efficient pathway expression.

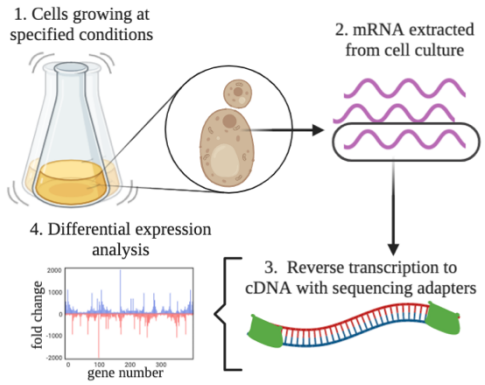

Figure 1. Simplified steps of RNAseq process for identification of promoters.

One way I plan to find new native promoters is through the use of RNAseq, described in **Figure 1**. RNAseq is a technique that can identify transcriptionally active regions of the genome and provide data on the expression levels of genes under various growth conditions or metabolic states. This leads to the identification of relevant promoter regions and data that can then be compared between specified growth conditions and normal growth conditions in which the promoter region should not be active. This technique will also be applied to other organisms in order to create non-native promoters, allowing for the identification of exciting new promoters that behave differently than native ones in the host organism. The identified promoter sequences will then be

obtained for both sets of promoters and will be expressed recombinantly and characterized in the host organism. The identification of both native and non-native promoters will allow for a more robust toolkit than what is currently available.

These new promoters will then be applied to pathways of high interest in the scientific community. One such pathway is the pathway for production of plant oils. Jojoba oil was identified as a leading ingredient for anti-aging formulas for skincare in the 2000s and is still widely used today with increasing demand[8]. Harnessing the ability of *Y. lipolytica*'s high metabolic flux towards fatty acids enables a new route of production for the long chain fatty acids that comprise Jojoba oil. The enzymes in this pathway have been identified[9] and can be recombinantly expressed to produce Jojoba oil. This pathway will be a testing ground and motivator for identifying and characterizing novel promoters. After successful integration of the recombinant pathway, scale-up production studies can begin.

The experiments for genetic optimization will take place in small volumes within 48-well plates (2mL). After creating multiple genetic libraries through use of promoters discovered and discussed previously, fermentation growth conditions can be studied in shake flasks (250mL). Studying the organism in shake flasks informs decisions about growth conditions in large scale studies, such as those done in a bioreactor. After determining high quality growth characteristics, studies will be moved into a bioreactor (typically 1.5L+) to study the industrial feasibility of the process. Few studies have been done using *Y. lipolytica* in bioreactors, so this process will require permutations of multiple parameters to determine the best operating conditions for growth and oil production. This objective will be assessed by the ability of the bioreactor process to be scaled and replicated and for *Y. lipolytica* to produce high oil titers at scale under the optimal conditions.

Throughout the completion of genetic cloning and scale-up studies, enhancement of current computational models will be occurring in parallel. Genome-scale-metabolic models (GEMs) have been created for *Y. lipolytica*. GEMs provide a kinetic model of cellular metabolism, meaning that a GEM can predict the metabolic activity of an organism based on user-defined parameters. The *Y. lipolytica* models need to be improved so that a wider range of researchers can use them. The best models achieve around 80% accuracy to experimental findings, however for people outside the field of metabolic engineering, the models are hard to understand and use[7]. In order to enhance the model, experimental data will be collected on new promoters, new growth conditions, and scale-up studies. Confirmation of already existing gene editing tools will also be collected. Utilizing GEMs is radically different than metabolic engineering's randomized approach for identification of optimal conditions. The GEMs allow researchers actively working with *Y. lipolytica* to first test their hypotheses *in silico* so that they do not spend excess time and resources attempting to screen every genetic and/or fermentation parameter.

Broader Impacts: Currently, Jojoba oil is produced via extraction of the oil from the seeds of the Jojoba plant. This time-intensive and costly process could be made easier through the creation of a heterologous production host. An efficient recombinant host organism allows for scientists to produce the same quality oil with a significant decrease in the environmental stresses associated with harvesting from the natural Jojoba plant. Metabolic engineering enables *Y. lipolytica* to sustainably convert widely available carbon sources into high-value natural products. However, until further promoters are developed, and scale-up studies are done on the organism, its full potential cannot be harnessed. Furthermore, the *in silico* model of *Y. lipolytica* will allow a large community of scientists to work together to better understand the growth and production capabilities of the organism. Taken together, the work proposed above leads to a continued advancement of metabolic engineering technologies to benefit society through development of advanced methodologies for sustainable chemical production.

References:

1. Whitehead, T. *et al. Biotechnol. Bioeng.* (2020).
2. Hussain, M. *TigerPrints* (2017).
3. Michely, S. *et al. PLoS One* (2013).
4. Ledesma, R. *et al. Trends Biotechnol.* (2016).
5. Gonçalves, F. *et al. The Scientific World Journal* (2014).
6. Larroude, M. *et al. Microb. Biotechnol.* (2019).
7. Ma, J., *et al. J. Ind. Microbiol. Biotechnol.* (2020)
8. Ahmad, A. *et al. Biomed. Dermatology* (2020).
9. Miklaszewska, M. *et al. Plant Sci.* (2016).