DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Reference herein to any social initiative (including but not limited to Diversity, Equity, and Inclusion (DEI); Community Benefits Plans (CBP); Justice 40; etc.) is made by the Author independent of any current requirement by the United States Government and does not constitute or imply endorsement, recommendation, or support by the United States Government or any agency thereof.

Final Report: Accomplishments

Award DE-SC0019412

Understanding and Harnessing the Robustness of Undomesticated *Yarrowia lipolytica* Strains for Biosynthesis of Designer Bioesters

Cong T. Trinh (PI), University of Tennessee Knoxville
Bruce Dien and Patricia Slininger (Co-PIs), US Department of Agriculture
Richard Giannone (Co-PI), Oak Ridge National Laboratory

1. What are the major goals of the project?

This project seeks to elucidate and harness the exceptional robustness of novel and undomesticated *Y. lipolytica* isolates, which were identified from a genetic diversity screening for compatibility with bioenergy development. Bioenergy-relevant isolates were further developed as microbial platforms for efficient conversion of undetoxified biomass hydrolysates into designer bio-esters continuously recovered by solvent extraction. The project has three major goals. **Goal 1.** Elucidate and enhance the endogenous metabolism of *Y. lipolytica* for superior growth, sugar utilization, and lipid accumulation in undetoxified biomass hydrolysates under hypoxic conditions. **Goal 2.** Understand and enhance the underlying mechanism of exceptional tolerance of *Y. lipolytica* to organic solvents. **Goal 3.** Elucidate and rewire endogenous metabolism of the most robust *Y. lipolytica* strains for effective conversion of accumulating lipids to designer bio-esters.

Significant progress has been made toward completing all research goals. We elucidated and optimized the robustness of *Y. lipolytica* by utilizing mixed C5 and C6 sugars in switchgrass hydrolysates (SGH) for lipid production (**Aim 1**). We conducted extensive omics analysis to investigate how genetic diversity among *Yarrowia* strains, derived from natural isolates or developed through adaptive laboratory evolution, influences lipid production when utilizing SGH. In **Aim 2**, novel mechanisms and underlying genetics were discovered that enabled *Yarrowia* strains to thrive in cultures containing high ionic liquid (IL) concentrations. Amongst other novel findings, it was found that sterols strengthened cell membranes to confer IL toxicity resistance, specifically via increased ergosterol content upon exposure to IL. In **Aim 3**, mechanistic studies elucidated how *Y. lipolytica* utilized intracellular lipids and alkanes/alkenes, leading to our discovery of novel enzymes and pathways for making short-chain esters. Most notably, thermostable chloramphenicol transferases were repurposed to function as alcohol acetyltransferases in *Y. lipolytica,* as well as the Gram-negative and Gram-positive bacteria *Escherichia coli* and thermophile *Clostridium thermocellum,* respectively.

2. What was accomplished under these goals?

Task 1: Elucidate and enhance the endogenous metabolism of *Y. lipolytica* for superior growth, sugar utilization, and lipid accumulation in undetoxified biomass hydrolysates under hypoxic conditions.

The first aim was to understand how natural genetic diversity of undomesticated *Y. lipolytica* strains relates to complex phenotypes like superior growth, sugar utilization, lipid accumulation, and degradation. In collaboration with JGI, five robust *Y. lipolytica* strains (YB392, YB419, YB420, YB567, and CBS7504) were sequenced and annotated, where robustness was investigated about high lipid production, pentose/hexose sugar utilization, and sugar alcohol secretion from undetoxified SGH[1]. The natural genetic diversity of undomesticated *Y. lipolytica* strains, coupled with detailed strain characterization and proteomic analysis, revealed metabolic processes and regulatory elements conferring desirable phenotypes for growth, sugar utilization, and lipid accumulation in SGH by these natural variants[2]. Both singleton genes and gene groups were discovered exclusive to these robust strains. Strain characterizations from computer-controlled bioreactor cultures revealed that the undomesticated strain YB420 used xylose to support cell growth and maintained high lipid content, while the conventional strain CBS7504 degraded cell biomass and intracellular lipid stores after transitioning to xylose as the sole remaining carbon source. Proteomic analysis revealed that carbohydrate transporters, xylose metabolic enzymes, and pentose phosphate pathway proteins increased in abundance during the xylose uptake stage for both strains. We distinguished proteins in lipid metabolism (e.g., lipase, NADPH generation, lipid regulators, β-oxidation) activated by YB420

(lipid maintenance phenotype) or CBS7504 (lipid degradation phenotype) when xylose was the sole remaining carbon source. To further understand the genetic regulation of complex sugar utilization, we identified and characterized a library of endogenous transcription factors (TFs) to better understand their functional roles on carbon catabolite repression and activation when *Y. lipolytica* grows in the presence of mixed sugars. Overall, results from this study provide an improved understanding of the robust metabolism of *Y. lipolytica* and suggest metabolic engineering strategies to enhance performance.

Among the undomesticated *Y. lipolytica* strains, YB392 and YB420 were the most robust for growth and lipid production on non-detoxified dilute-acid pretreated SGH. To understand the underlying mechanisms and further enhance their robustness, these two strains were engineered to improve xylose assimilation and lipid accumulation by overexpressing the native *Y. lipolytica* xylitol dehydrogenase, xylulokinase, acetyl-CoA carboxylase (ACC1), and diacylglycerol acyltransferase (DGA1), leading to strains YB392-XL and YB419-XL. The engineered strains were improved for growth, C5 and C6 sugar utilization, and lipid accumulation on concentrated non-detoxified SGH using adaptive laboratory evolution (ALE). The best-performing ALE mutants were isolated after approximately 120 generations in up to 45% (v/v) non-detoxified SGH. The adapted strains demonstrated significant improvements in growth and xylose assimilation when challenged with 70% (v/v) non-detoxified SGH compared to the non-adapted strains, which demonstrated poor growth. Notably, the evolved YB419-XL and YB392-XL strains produced approximately 1.6-fold and 2.2-fold more lipids than the wildtype YB419 and YB392, respectively. Encouragingly, the evolved YB392-XL strain produced a 3.2-fold improvement over the wild type when grown on xylose. To gain comprehensive insights into the biological systems of these ALE-evolved YB392-XL and YB419-XL strains, the strains were characterized for growth and lipid production on 50%v/v concentrated non-detoxified SGH using computer-controlled bioreactors; cell samples were submitted for genomics, transcriptomics, and proteomics analyses.

To deepen our understanding of the cellular robustness of *Y. lipolytica*, we also investigated several critical parameters for enhanced biomass and lipid production in non-detoxified SGH, including nitrogen sources, oxygen levels, pH, and a two-stage culture scheme. First, we evaluated various nitrogen sources, including synthetic medium, urea, yeast extract, and casein amino acids. Our results indicated that the synthetic medium with urea (SMU) was superior for cell growth compared to the medium containing casein amino acids. Moreover, adding yeast extract to the SMU further improved growth in the 60% (v/v) non-detoxified SGH, albeit marginally. Next, we examined the effect of pH on lipid production across a range of pH values from 2 to 6. We found that pH 6 was optimal for lipid production compared to the lower pH levels. Following this, we demonstrated that a two-stage culture significantly increased lipid titer. This approach consisted of an initial "Growth Phase" with additional nitrogen followed by a "Lipid Production Phase" conducted without a nitrogen source. We produced a comparable amount of lipid from the refined sugars using the 60% (v/v) non-detoxified SGH in this two-stage culture. Since *Y. lipolytica* is an obligatory aerobe, oxygen is critical for its growth and metabolic activity. We investigated how lipid production was affected by oxygen concentration (10-60% of saturation) and aeration rate (150-300 rpm) in the bioreactor. We found that the percentage of oxygen had minimal impact on lipid accumulation. This suggests that production costs can be decreased by ramping aeration down after cell growth and during lipid production. However, we observed significant differences in lipid production at various rpm; the maximum lipid titer was obtained at 250 rpm (~9.5 g/L), followed by 300 rpm (~6.5 g/L).

Overall, the investigation into the genetic diversity of undomesticated *Y. lipolytica* strains, coupled with detailed strain characterization and proteomic analysis, revealed metabolic processes and regulatory elements conferring desirable phenotypes for growth, sugar utilization, and lipid accumulation in undetoxified biomass hydrolysates by these natural variants. The research findings advance our understanding of *Y. lipolytica* genetics and metabolism, which can be harnessed for biomanufacturing biofuels, biochemicals, and biomaterials.

Task 2: Understand and enhance the underlying mechanism of exceptional tolerance of *Y. lipolytica* to organic solvents.

Biotransformation in organic solvents enables a two-phase fermentation beneficial for high-product synthesis and bioprocessing. Implementation of these advanced fermentation systems requires an understanding of and the ability to engineer solvent-tolerant microorganisms. Prior to this project, we determined that wildtype *Y. lipolytica* successfully performed simultaneous saccharification and fermentation for high-yield production of alpha-ketoglutarate when challenged with 10% (v/v) 1-ethyl-3-methylimidazolium acetate [EMIM][OAc], a concentration at which industrial microorganisms *S. cerevisiae* and *E. coli* were severely inhibited. To elucidate and enhance the IL-tolerance of *Y. lipolytica*, we first performed adaptive laboratory evolution (ALE). After 200 generations of ALE in benchmark [EMIM][OAc], we isolated the most IL-tolerant *Y. lipolytica* strain YlCW001 ever reported, which tolerates 18% (v/v) [EMIM][OAc] (**FIG. 1**). This strain (YlCW001) also grew robustly in a broad range of solvents [3,4]. Genome-wide analysis, including scanning electron microscopy (SEM), metabolomics, lipidomics, and transcriptomics elucidated the mechanisms behind its exceptional IL tolerance. SEM revealed that ILs cause membrane disruption, as evidenced by dents and wrinkles in the WT strain, while YlCW001 maintained a healthy morphology resistant to IL-perturbation. Multi-omics characterization pointed to sterols as a critical membrane component in conferring high IL-tolerance in *Y. lipolytica*. For confirmation, overexpression of the novel sterol transcription factor bestowed high solvent tolerance in the parent strain. A model was formulated that explained how *Y. lipolytica* restructured its membrane to cope with IL toxicity.

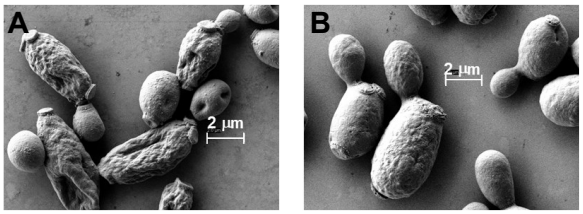

FIGURE 1: (A) Wildtype & **(B)** engineered *Y. lipolytica* in [EMIM][OAc].

Since IL-tolerance is a complex phenotype, we also aimed to elucidate the genotypes that map to this robustness at the genome scale in *Y. lipolytica*. In collaboration with JGI, we performed dynamic transcriptomic analyses for the wildtype and IL-tolerant *Y. lipolytica* strains growing with or without IL. The dynamic transcriptomic data was used for differential gene analyses to identify key cellular processes of *Y. lipolytica* responsive to IL exposure. The top candidates for validation were selected using the in-house developed Gene Co-expression Connectivity (GeCCo) method to identify gene targets conferring desirable phenotypes based on their number of connected genes. We validated GeCCo through the single and dual expression of gene targets conferring IL tolerance. We discovered a non-intuitive strategy that combined overexpression of BRN1 and OYE2, associated with cell division and Golgi vesicle transport, to synergistically confer IL tolerance[5].

Results from the FICUS project expanded our understanding of IL-tolerance in *Y. lipolytica*. We collaborated with the PNNL team to perform cryo-TEM, proteomics, lipidomics, and metabolomics, and the JGI for epigenetics. While the COVID-19 pandemic disrupted this project's progress and completion, we managed to collect cryoTEM, proteomics, lipidomics, and metabolomics data together with in-house microscopy. We discovered many other processes and genotypes conferring IL tolerance. These findings will be disseminated through a future manuscript.

Overall, this research provides a deep understanding of the biochemistry responsible for the cellular robustness of *Y. lipolytica*, enabling survival and growth in the presence of organic solvents. This important phenotype directly benefits industrial biomanufacturing of fuels, chemicals, and materials.

Task 3: Elucidate and rewire endogenous metabolism of most robust *Y. lipolytica* strains for effective conversion of accumulating lipids to designer bio-esters.

While *Y. lipolytica* is known to be a thiamine auxotroph, the basis of this genetic deficiency and its effect on lipid production were unknown. One key highlight of this task was to elucidate and eliminate the detrimental effect of thiamine deficiency for *Y. lipolytica*[6]. Using proteomics and physiological characterization, we identified key metabolic processes influenced and regulated by thiamine availability and identified the genetic basis of thiamine auxotrophy in *Y. lipolytica*. Thiamine availability influenced lipid biosynthesis, with supplementation enhancing lipid production in the engineered thiamine-

prototrophic *Y. lipolytica*. In addition, thiamine-regulated promoters were discovered that can be harnessed in the future as biosensors to control metabolic processes.

Short-chain esters, including butyl acetate and hexyl acetate, are synthesized using alcohol acetyltransferases (AATs) originating from fruits or plants. However, these AATs exhibit low solubility, limiting their activities[7]. We discovered that thermostable chloramphenicol acetyltransferase (CAT) could be repurposed to function as a thermostable AAT through genome mining and model-guided protein engineering[8]. A single mutation CATec3 Y20F can improve its activities in a broad range of alcohols (e.g. 21) and eight acyl CoAs, allowing for the synthesis of 168 unique esters. This CATec3 Y20F has excellent expression and solubility and is compatible with mesophilic and thermophilic bacteria and yeast such as *Y. lipolytica*. Due to the diversity of AATs, we also developed a high-throughput colorimetric screening method to probe the specificities of these AATs[9]. We built a dynamic regulator-actuator system that enables *Y. lipolytica* to sense and convert hydrophobic substrates into designer short-chain esters (manuscript in preparation).

Studies were conducted to better understand how *Y. lipolytica* rewires its endogenous metabolism to convert hydrophobic substrates, such as accumulating lipids or alkanes. The first study was to grow various non-model *Y. lipolytica* strains on biomass hydrolysates with mixed sugars, as described in Task 1[2]. The second study involved growing *Y. lipolytica* in defined, nitrogen-limited, glucose-containing media for lipid accumulation, followed by an endogenous lipid degradation phase with an external methanol (or ethanol) supply for biodiesel synthesis. Proteomics was performed at PPNL through the FICUS project to identify endogenous enzymes (lipases, to be published) responsible for biodiesel production. The third study characterized *Y. lipolytica* strains able to grow on alkanes and depolymerized plastic oils derived from polyolefin. In using these hydrophobic substrates, two populations of *Y. lipolytica*, including oil-adhering cells and planktonic cells, were observed with distinct proteomes, revealing how *Y. lipolytica* fine tuned their metabolisms. *Y. lipolytica* exhibited more robust growth on hexadecane than glucose but was inhibited by plastic oils. We found a reduction in proteome allocation for protein biosynthesis, at the expense of the observed increase towards energy and lipid metabolisms, might have caused the inhibitory effect of DP oil on cell growth[10]. To guide strain designs to make various esters, efforts in this task involved the development of modular cell engineering framework[11] and CRISPR genetic tools[12].

3. What opportunities for training and professional development has the project provided?

The project trained five postdoctoral researchers (UTK), three PhD students (UTK), six Masters students (UTK), four rotation students (UTK), five undergraduates (UTK, ORNL), one research assistant (USDA), and one research intern (USDA). Many of these students were advised by the PI and Co-PIs of the project. The researchers had the opportunity to present their results at the DOE PI annual meeting, the AIChE annual conference, the ACS BIOT conference, the SIMB annual conference, and SFBC conference.

4. How have the results been disseminated to communities of interest

The project generated eleven peer-reviewed research papers[1-6,8-11], three peer-reviewed current opinion papers[13,14], two book chapters[15,16], and four provisional patents. It contributed to developing software, including ModCell (Modular Cell) for rational strain design and GeCCo (Gene Coexpression Connectivity) for performing gene co-expression connectivity to predict gene targets conferring robustness phenotypes. In addition, the research results in developing synthetic biology tools such as CRISPR-GRIT, thiamine biosensor, and sugar transcription factors. The team has disseminated the research results through various activities:

- Teaching the interdisciplinary class “Elements of Metabolic Engineering and Synthetic Biology” to both undergraduate and graduate students with broad educational backgrounds (e.g., Chemical Engineering, Bioengineering, Microbiology, Biochemistry);
- Training undergraduate students to conduct independent research;
- Providing lab tours and performing hands-on demonstrations for high school students along with their parents. For instance, NCAUR has developed a scientific display on the topic of yeast that has been presented to the public and met with visiting groups on a regular basis. The Trinh lab hosts multiple

outreach activities for high schoolers every semester. Presenting research findings at general scientific conferences and the DOE annual PI meetings; and

- Depositing research data (proteomics, DNA sequencing) in publicly available databases.

References

1 Walker, C., Ryu, S., Na, H., Zane, M., LaButti, K., Lipzen, A., Haridas, S., Barry, K., Grigoriev, I. V. & Quarterman, J. Draft Genome Assemblies of Five Robust Yarrowia lipolytica Strains Exhibiting High Lipid Production, Pentose Sugar Utilization, and Sugar Alcohol Secretion from Undetoxified Lignocellulosic Biomass Hydrolysates. *Microbiol Res Announc* **7**, e01040-01018 (2018).

2 Walker, C., Dien, B., Giannone, R. J., Slininger, P., Thompson, S. R., Trinh, C. T. Exploring Proteomes of Robust Yarrowia lipolytica Isolates Cultivated in Biomass Hydrolysate Reveals Key Processes Impacting Mixed Sugar Utilization, Lipid Accumulation, and Degradation. *mSystems* **6**, e00443-00421 (2021).

3 Walker, C., Ryu, S. & Trinh, C. T. Exceptional solvent tolerance in Yarrowia lipolytica is enhanced by sterols. *Metabolic engineering* **54**, 83-95 (2019).

4 Walker, C., Ryu, S., Haridas, S., Na, H., Zane, M., LaButti, K., Barry, K., Grigoriev, I. V. & Trinh, C. T. Draft Genome Assemblies of Ionic Liquid-Resistant Yarrowia lipolytica PO1f and Its Superior Evolved Strain, YlCW001. *Microbiology Resource Announcements* **9** (2020).

5 Walker, C., Ryu, S., Garcia, S., Dooley, D., Mendoza, B. & Trinh, C. T. Gene Coexpression Connectivity Predicts Gene Targets Underlying High Ionic-Liquid Tolerance in Yarrowia lipolytica. *mSystems* **7**, e00348-00322 (2022).

6 Walker, C., Ryu, S., Giannone, R. J., Garcia, S. & Trinh, C. T. Understanding and Eliminating the Detrimental Effect of Thiamine Deficiency on the Oleaginous Yeast *Yarrowia lipolytica*. *Applied and Environmental Microbiology* **86**, e02299-02219 (2020).

7 Lee, J.-W. & Trinh, C. T. Controlling selectivity of modular microbial biosynthesis of butyryl-CoA-derived designer esters. *Metabolic Engineering* **69**, 262-274 (2022).

8 Seo, H., Lee, J.-W., Giannone, R. J., Dunlap, N. J. & Trinh, C. T. Engineering promiscuity of chloramphenicol acetyltransferase for microbial designer ester biosynthesis. *Metabolic Engineering* (2021).

9 Lee, J. W., Seo, H., Young, C. & Trinh, C. T. Probing specificities of alcohol acyltransferases for designer ester biosynthesis with a high-throughput microbial screening platform. *Biotechnology and Bioengineering* **118**, 4655-4667 (2021).

10 Walker, C., Mortensen, M., Poudel, B., Cotter, C., Myers, R., Okekeogbu Ikenna, O., Ryu, S., Khomami, B., Giannone Richard, J., Laursen, S. & Trinh Cong, T. Proteomes reveal metabolic capabilities of Yarrowia lipolytica for biological upcycling of polyethylene into high-value chemicals. *mSystems* **8**, e00741-00723 (2023).

11 Garcia, S. & Trinh, C. T. Computational design and analysis of modular cells for large libraries of exchangeable product synthesis modules. *Metabolic Engineering* **67**, 453-463 (2021).

12 Cotter, C. J. & Trinh, C. T. CRISPR-GRIT: Guide RNAs with Integrated Repair Templates Enable Precise Multiplexed Genome Editing in the Diploid Fungal Pathogen *Candida albicans*. *The CRISPR Journal* (2024).

13 Garcia, S. & Trinh, C. T. Modular design: Implementing proven engineering principles in biotechnology. *Biotechnology Advances* **37**, 107403 (2019).

14 Lee, J. W. & Trinh, C. T. Towards renewable flavors, fragrances, and beyond. *Curr Opin Biotechnol* **61**, 168-180 (2020).

15 Ryu, S. & Trinh, C. T. Methods to activate and elucidate complex endogenous sugar metabolism in Yarrowia lipolytica. *Yarrowia lipolytica: Methods and Protocols*, 175-189 (2021).

16 Carper, D., Trinh, CT., Giannone, RJ. Proteomic Analysis of the Oleaginous Yeast Yarrowia lipolytica. (Springler, 2025).