The Dissertation Committee for Emily Bowman Certifies that this is the approved version of the following Dissertation:

Microdroplet Assay Development for Metabolic Engineering and Synthetic Biology Applications

Committee:

Hal Alper, Supervisor

Jeffrey Barrick

Lydia Contreras

Kenneth Johnson

David Taylor

Microdroplet Assay Development for Metabolic Engineering and Synthetic Biology Applications

by

Emily Katherine Bowman

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2023

Dedication

To my supportive and loving husband.

Acknowledgements

First, I would like to thank my advisor, Professor Hal Alper, for his mentorship and guidance throughout my graduate school experience. His experience guided my projects and helped cultivate a supportive lab culture.

I also would like to thank my committee members (in alphabetical order): Professor Jeffrey Barrick, Professor Lydia Contreras, Professor Kenneth Johnson and Professor David Taylor. Their support and guidance throughout my PhD pursuits have been motivating and helpful.

I finally would like to thank all of the members of the Alper lab, as well as external collaborators who made this work possible. In particular, Phuong Tran Huong and Gina Partipilo, who's help executing and planning experiments culminated in exciting discoveries and publications.

Abstract

Microdroplet Assay Development for Metabolic Engineering and Synthetic Biology Applications

Emily Katherine Bowman, PhD

The University of Texas at Austin, 2023

Supervisor: Hal Alper

Sustainable, efficient production of societally relevant chemicals in microorganisms is the epicenter of the field of metabolic engineering. Modern developments in DNA synthesis and sequencing have enabled high-information gene perturbation and protein engineering libraries that require reliable, reproducible, high throughput screening assays. Detailed herein are examples of utilizing growth-based assays and microfluidic screening to extract pertinent information from different gene perturbation libraries, microbial consortia, and engineered proteins.

The first study details an approach to regulate gene expression in *S. cerevisiae* through utilization of a single-guide RNA library paired with dCas9 fused to either Mxi1 or VPR. Through these fusions and targets of guide RNAs, genes are either up or down regulated at increasing degrees. Using this library it was possible to identify target genes for moderate up and down regulation to improve growth on alternative carbon sources, namely glycerol and galactose. Final analysis of enriched guides via next generation sequencing identified moderate down regulation of essential genes, novel perturbations that would not have been isolated in traditional gene knock-out approaches.

The studies detailed in chapters 3 and 4 delve into technical applications of microfluidics to screen for improved small molecule production. In chapter 2, the library detailed above is utilized to screen for gene targets that improve small molecule production in *S. cerevisiae.* This study also results in the development of a pico-injection microdroplet approach that utilizes cell-based biosensors to transduce small molecule production into a fluorescent signal. The study in chapter three characterizes this application further, ultimately leading identifying gene perturbations that improve early productivity or higher overall production, based on the time at which they were screened.

In the fourth study, a previously developed CuAAC probe for extracellular electron transfer (EET) is ported into droplets to analyze an environmental microbial consortium. This study centers on the development of an oxygen-limited platform for anaerobic microbe cultivation in microdroplets, ultimately enriching environmental microorganisms previously uncharacterized for EET. This work expands the range of microorganisms compatible with this microdroplet system. In the final study, the microorganism array is expanded further through the development of a cytotoxicity assay for Sf9 insect cells is developed for use in microdroplets. This work lays a foundation for future applications to identify novel insecticidal toxins from engineered protein libraries. Collectively, these studies establish generalizable assays for high throughput screening of a wide array of organisms related to the field of synthetic biology.

Table of Contents

List of Tables

List of Figures

to make spent media. Varying amounts of TAL standard were added to the spent media to generate emulsions of different TAL concentrations. *E. coli* TAL biosensors were pico-injected into these emulsions at an OD_{600} of 5.0 (final in-droplet OD_{600} of 1.0) in concentrated LB (5x LB-Kan selective media) to the spent media TAL standard emulsions along with a sodium phosphate buffer at pH 7.2. (*ii*) Here, BL21DE3 spent LB media was used with varying concentrations of L-DOPA. Biosensor response at concentrations flanking that were already made by the base strain was linear. These emulsions were run through the sorting chip to collect fluorescence measurements of 100,000 droplets for each concentration of analyte on three separate occasions (totaling 300,000 droplets). Weighted averages of each concentration are plotted along with their respective SDs (Calculated separately from the weighted average). Note: Triplicate data are shown, but it is difficult to distinguish as the fluorescence measurements were very similar. (*B*) Fermentations of individual isolates from an unsorted control and sorted pool for TAL producing *S. cerevisiae* show population enrichment. The most frequently occurring sgRNAs identified by this screen were evaluated alongside controls to demonstrate enhanced production. Error bars represent SD. (C) Fermentations of individual isolates from an unsorted control and sorted pool for L-DOPA producing *E. coli* likewise shows population enrichment. Individual isolates identified via TAIL PCR

relative abundance of DNA in each sort condition. This assay indicated a clear enrichment for active toxin (txn1) in the sorted sample versus the unsorted control, indicated isolated fluorescence from the screen did pull out those droplets containing active toxin and dead Sf9 cells. ----------146

Chapter 1: Introduction

The central premise of industrial metabolic engineering is to rewire cellular metabolism and create strong overproduction phenotypes[1-3]. While rational design approaches have advanced in recent years to aid this goal, nearly all strain-engineering efforts rely upon some degree of strain screening or selection[4-6]. In recent years, the field has emerged form a classic strain development scheme marked by mutagenesis and selection into a synthetic biology-driven library construction paradigm. Regardless of the exact approach, there is a reliance on high-throughput screening[7]. Additionally, new advances in DNA synthesis have led to the creation of large and complex combinatorial libraries, many leveraging clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 systems for multiplexing gene editing, thus creating a further bottleneck in the test step of the design-build-test paradigm[8,9]. However, these two approaches have limitations, such as the need for strong automation and the incompatibility with many phenotypes, such as secreted products (because genotype and phenotype are physically unassociated), respectively. In this regard, metabolic engineering is limited by the ability to rapidly and accurately screen through large libraries that can be created at the rate of transformation or even cell growth in the case of continuous evolution systems[10-12]. This limitation is especially poignant in the area of natural products engineering, where large libraries of putative genes and varied refactored clusters are routinely tested.

An emerging and alternative approach for rapid screening involves picoliter-sized aqueous-in-oil emulsions containing a single cell library member in each microdroplet. As discussed further later, this technique enables a high-throughput approach that re-establishes the genotype-phenotype linkage desired in microtiter-based screens while maintaining the throughput and miniaturization marked by fluorescence e activated cell

sorting (FACS)-based screening. To this end, several applications have begun to emerge that use microdroplet-based technologies for metabolic engineering. Microfluidic screening is an emerging approach with great potential; however, it has not been fully implemented in the metabolic engineering community compared with other alternatives. This technique has been mainly used in the area of single enzyme evolution and detection but is highly applicable to the area of metabolic engineering. Given the many advantages of this approach and the growing access to instrumentation, it is expected that its use and adoption will greatly increase in the future.

1.1 MICROFLUIDIC MICRODROPLET SYSTEMS AS AN EMERGING TOOL FOR BIOTECHNOLOGY

Microdroplet technologies allow for high-throughput, single cell screening of biology on a scale that has the capacity to cover libraries and natural variation at a rate of 10^6 per hour[13]. Across the biological sciences, this technology has enabled applications that span from single cell sequencing and. Medical diagnostics to metabolic engineering applications on the basis of its inherent reliability and extremely high-throughput capacity[14,15]. As an example, microdroplets have led to the development of diagnostic tools, including droplet digital (dd)PCR, a technique that enables PCR on single cells in aqueous-in-oil emulsions[16]. This technique is sensitive and has led to diagnostic tests for chlamydia as well as having applications in single cell identification[15,17]. In addition, single cell genome sequencing enabled via microdroplets has allowed for the profiling of individual microbe members from a consortia[18]. This technology has also lent itself to single cell RNA-sequencing (RNA-seq) for transcriptomic profiling of individual cells[19,20]. Antibiotic screening and susceptibility have been explored within microdroplet systems by encapsulating pathogenic bacteria and testing strains in a gradient of antibiotic to evaluate

drug potency[21]. Further diagnostic and fundamental applications in microdroplets include drug design assays as well as querying cell-to-cell communication[22,23]. Just as microdroplets have influenced these fields and utilizations, their use in metabolic applications has been to afford a better screening technology compared with microtiter plate parallelization and single-cell FACS sorting.

1.2 ADVANTAGES AND DISADVANTAGES OF MICRODROPLET SYSTEMS IN METABOLIC ENGINEERING

Traditional approaches for high-throughput selection in metabolic engineering have relied on the use of FACS and growth selections. As examples, previous approaches to identify improved industrial yeasts have included screening transposon libraries for gene disruptions leading to higher levo-dopa (L-DOPA) production[24] as well as knockout libraries for improvement of lactic acid production[25], identified via FACS and growth assays, respectively. In this regard, FACS screening provides the high-throughput capabilities necessary to screen large libraries. However, such screening is inherently limited to intracellular (or cellular-membrane associated) fluorescence. As a major limitation here, the secretion of a product inherently disconnects the genotype-phenotype connection required in FACS. By contrast, microtiter plates can maintain the link of secretion phenotype to genotype when combined with automation and/or high-throughput chemical assays. However, screening is somewhat limited to a throughput of up to 10^5 members[26]. The use of picoliter-sized droplets can physically link the genotype and phenotype for secreted products and even serve to concentrate secreted products to enable higher effective concentrations and bypass dilution effects inherent with larger scale culturing approaches[27].

1.2.1 Microdroplets are Better Suited than FACS for Secreted Products

A previous study compared the use of microdroplet-enabled selections with traditional single cell FACS sorting. The results indicated that, despite the same throughput and stringency (because both modes were collected similarly using standard flow cytometry equipment), the different modes of selection yielded different phenotypes. In particular, traditional single cell FACS isolated strains with increased intracellular riboflavin content, whereas the water/oil/water (w/o/w) microdroplet FAFCS isolated strains with increased secreted content. In the latter cases, the resulting phenotype and selection capacity is advantageous for industrially relevant phenotypes because most small-molecule production schemes would favor secreted products for ease of harvest from the culture. Moreover, while both approaches did serve to increase the total riboflavin content (i.e., the sum of intracellular and extracellular content), all cell isolated through microdroplet selection outperformed those from single cell FACS[28]. Fluorescence activated droplet sorting (FADS) is emerging as a promising method for screening microbial libraries for increased small-molecule production. The screens detailed in Table 1-1 show a wide range of applications, with methods ranging from direct product detection to cell-based biosensors.

Organism	Product(s)	Library type	Class	Screening method	Sorting method	Benefits of microdroplet screening	Refs
Escherichia coli	2-ketoisovalerate/ L-tyrosine	Chemical mutagenesis	Biofuel precursor/ natural product	Syntrophic growth with fluorescently labeled sensor strain	FADS	Only retrieved high producer (RFP tagged) cells in sorted droplets; low false positive	[32]
Yarrowia lipolytica	Riboflavin	EMS[a]	Nutraceutical/ pharmaceutical precursor	Inherently fluorescent product	Microdroplet FACS	Increased secretion and production of Riboflavin (54-fold)	[31]
Saccharomyces cerevisiae	*p*-coumaric acid	28 individual strains	Nutraceutical	*E. coli*-based biosensor	FADS	96% enrichment of *p*-coumaric acid-producing yeast strains	[62]
	α-amylase	UV mutagenesis	Recombinant protein	Fluorogenic enzyme substrate	FADS	Identification of many generalizable mutations for increased protein secretion	[57]
	Tyrosine and streptavidin	Plasmid-based ARO4 mutant library	Natural product/ recombinant protein	RNA aptamers	FADS	Threefold increase in streptavidin titer, fivefold improvement in tyrosine titer	[37]
	Ethanol/xylose fermentation	Overexpression plasmid-based library	Biofuel	Oxidase enzyme/horseradish peroxidase/Amplex UltraRed and droplet merging	FADS	Identification of beneficial overexpression of XylA, and increased xylose consumption	[18]
Chlamydomonas reinhardtii	Lipids	EMS mutagenesis	Natural product	Lipid staining in droplets	FADS	1.47-fold increase in growth rate and 2.75-fold increase in lipid production	[58]
Bacillus coagulans	Lactic acid	Sodium nitrite (chemical mutagenesis)	Natural product	Fluorescent dye (Probe 1)	Microdroplet FACS	52% increase in lactic acid titer	[45]
Lactococcus lactis	Riboflavin	UV/EMS	Nutraceutical/ pharmaceutical precursor	Inherently fluorescent product	FADS	30% increase in riboflavin titer	[53]

Table 1.1: Screening, Sorting, and Benefits of Microfluidic Droplet-Based Systems for Microbial Production.

1.2.2 Microdroplets Allow for Higher Throughput Compared with Microtiter Plates

In a similar fashion, recent research compared the use of microfluidic microdroplet systems with those of microtiter plates[29]. In one such study, co-culturing of sensor cells with producer cells was utilized to measure product concentration. As a result, traditional FACS approaches, such as those described earlier for riboflavin, are not feasible because co-culturing is not compatible in single cell FACS applications. Therefore, the authors sought to ultimately evaluate their primary options, including microtiter plates, an agar-plate based assay, and microdroplet sorting. In this study, cells with improved content of the desired products (either 2-ketoisovalerate or L-tyrosine) enabled the sensor cells to

grow and, thus, provide for the subsequent growth of the producer cell through syntrophic growth (Figure 1-1 and 1-2)[29-31]. In this case, co-encapsulation of the sensor strain with the production strain was ideal for the syntrophic growth-sensing strategy. Ultimately, microdroplets allowed for a throughput of up to 10^7 library members, as well as a measurement that relied on fluorescence. This comparative analysis illuminates some particular advantages of microdroplets over the other high-throughput systems.

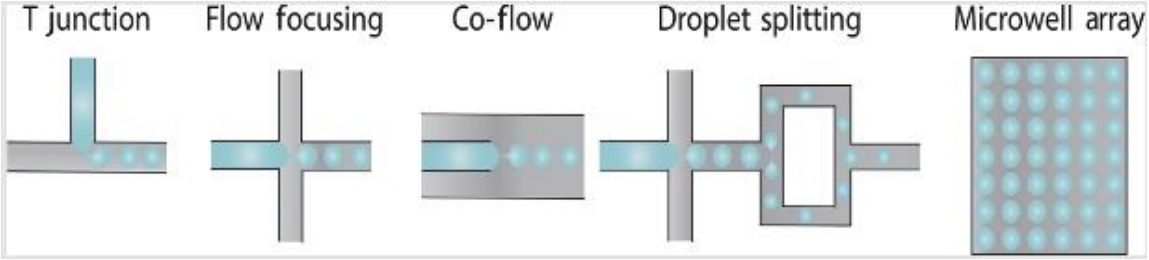

Figure 1-1: Schematic of Common Droplet Generation Systems. There are multiple methods to generate microdroplets, including T junctions, flow focusing, co-flow and microdroplet splitting. In flow focusing, the continuous oil phases flow directly at one another, with the aqueous phase flowing through this junction. Droplet diameter directly corresponds to the size of the junction itself and, therefore, different sized chips can be used to control the microdroplet diameter. In addition, microwell arrays in which microdroplets are printed onto a grid allow for direct indexing similar to the indexing available in microtiter plates, but on a smaller scale. The droplet generation method of choice depends upon the needs of the user, with flow-focusing devices being the most common for high-throughput generation and subsequent sorting[18].

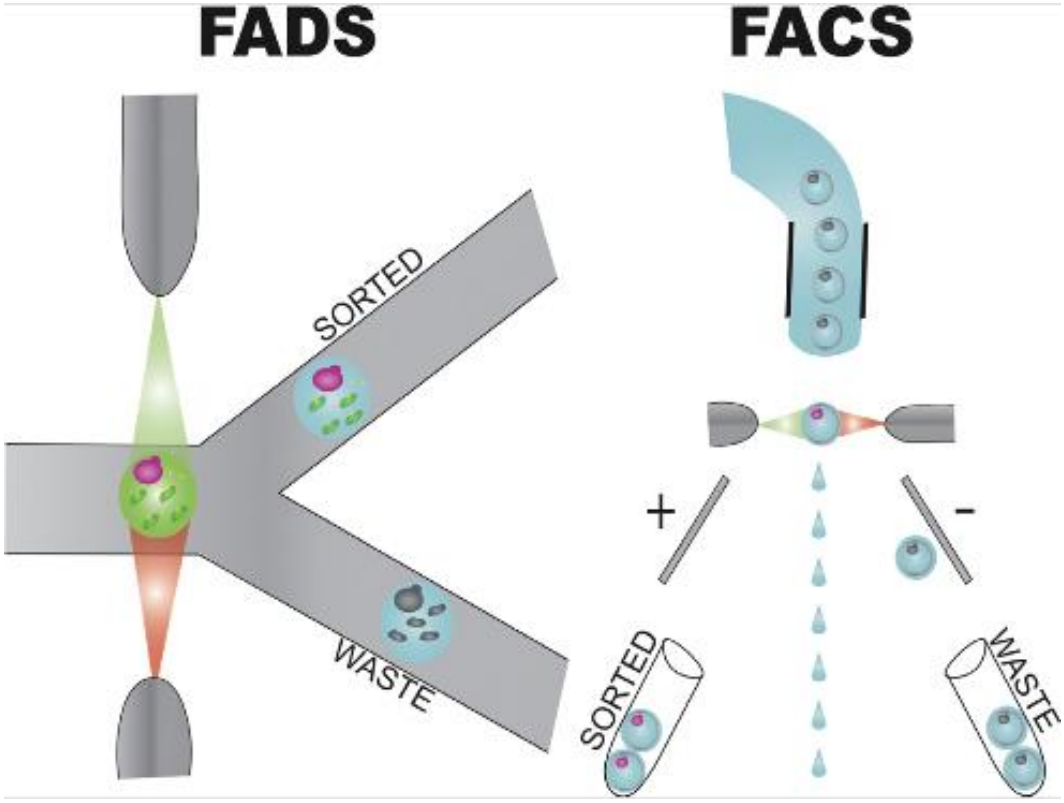

Figure 1-2:Common Droplet Sorting Modalities of Evaluation. Fluorescence activated droplet sorting (FADS) and cell sorting (FACS) are both common methods used to sort droplets. FADS allows for sorting of aqueous-in-oil emulsions, while FACS must sort water-oil-water (w/o/w) emulsions to be compatible with the fluidics of the machine.

1.2.3 Disadvantages of Microdroplets over Traditional Methods

While the studies discussed earlier directly identified improvements and advantages of microdroplet-based selection over traditional methods, they also identified some disadvantages. While the throughput of microdroplets is higher than that of microtiter plates and agar plating, there is also the issue of cell-to-cell variability that comes from single cell-based assays. This is because, similar to fluorescence activated cell sorting (FACS) only a single fluorescent measurement is taken for each microdroplet. However, microtiter plates allow for individual library members to be measured in replicate and averaged, producing a representation of cellular fluorescence that results in a more robust

enrichment system. Limitation to the overall throughput of the microdroplet-based systems is that initial single cell per microdroplet seeding is limited by the *Poisson distribution* when using standard approaches[32]. This results in a high number of empty microdroplets when aiming for one cell per drop[29]. This limitation is being studied with new ways to both establish microdroplets and tag or pre-sort only microdroplets that contain cells[33]. Moreover, these issues of cell-to-cell variability and limitations in droplet seeding can be ameliorated through iterative microdroplet enrichments, an approach typically utilized when isolating high-producing strains[34]. Finally, a technical limitation when using aqueous-in-oil microdroplets is the partition coefficient of hydrophobic molecules into the oil phase[35]. Specifically, the diffusion of hydrophobic molecules out of the microdroplet and it the oil phase limits the use of microdroplets with these substrates and products. While w/o/w emulsions can reduce the impact here, the resulting microdroplets are less stable overall[36]. These limitations notwithstanding, microdroplet selections have the potential to relieve the bottleneck that currently exists in the test modality of the design-build-test cycle.

1.3 MICRODROPLET-SORTING MODALITIES

A variety of microfluidics designs are available for microdroplet-based metabolic engineering applications. These techniques have specific applications for the kind of information that is desired.

1.3.1 Microdroplet Generation and Experimental Design

Aqueous-in-oil microdroplet or microarray generation technologies aim to generate picoliter-sized microdroplets consistently and quickly (Figure 1-1)[37]. These multiple modes of microdroplet generation allow for a flexibility in assay design associated with

microbial library sorting. As an example, encapsulating aqueous-in-oil microdroplets into agar to create microbeads can stabilize the emulsion and potentially secure the microdroplets for static (i.e., nonflow) analysis on microarrays[38,39] microdroplet merging[22] and pico-injection[23,24]. As with all technologies, each of these methods has different advantages and challenges. Nonetheless, they have all either been utilized for improving microbial strains or have the potential for such applications.

1.3.2 Fluorescence and Absorbance-Activated Microdroplet Sorting

At the cornerstone of most microdroplet-sorting applications is FADS. Similar to FACS, this system utilizes a laser to detect fluorescent signals and applies a voltage to the sample when the signal is greater than the user-defined threshold, thus enabling a differential sort via dielectrophoresis[23-26]. At the onset, this basic principle is similar to FACS; however, the detection and sorting takes place in an oil phase using microfluidic chips with electrodes imbedded into the chip itself. As a result, this approach is more simplified than the double emulsion systems that utilize w/o/w microdroplets[28,40] and more stable overall.

Ultimately, FADS technology can be used to sort for phenotypes of interest in a manner analogous to FACS. Recently, Tovar and colleagues expanded this technology to allow for multicolor detection in microfluidics via demodulation of multiple fluorescent signals[41]. Although this study was only an initial technological proof of concept conducted with microdroplets containing four different fluorescent dyes, it has the potential to improve microbial selection and expand the number of applications.

Finally, microdroplet sorting and detection on the basis of absorbance (as opposed to fluorescence) has been demonstrated[42]. This technology requires two optical fibers within the chip that detect a decrease in transmitted light followed by subsequent sorting.

This approach is certainly useful when evaluating enzymatic products that commonly have an absorbance at a specific wavelength. However, the measurement of absorbance tends to be a slower measurement than fluorescence, thus limiting its overall throughput.

1.3.3 Microdroplet Merging and Pico-injection

Microdroplet-merging and pico-injection approaches are manipulations that can enable alternative selection and sorting schemes. As discussed earlier, typical microdroplet seeding is limited by the *Poisson* distribution. As a result, if a given assay requires seeding the microdroplet with two different cells (one each from two different populations), a common co-flow focusing system would be severely limited in the number of microdroplets containing the desired ratio (Figure 1-1)[33]. To overcome this limitation, Chung and colleagues prepared two separate emulsions, each containing particles from two different populations[33]. These microdroplets were subsequently sorted (based on fluorescence) to isolate microdroplets that contained either a particle or a cell of interest. One population of these microdroplets was then immobilized onto a special microarray, and the other was flowed over, allowing for colocalization of each of the microdroplet populations. Upon treatment with an electromagnetic field, these microdroplets then merged and resulted in a collection of microdroplets that almost exclusively contained single particles from each population[33]. Thus, this approach of microdroplet merging is conceivable for applications of merging assay reagents with microdroplets containing cells. Other merging approaches allow for the addition of sensing reagents or dyes to measure product concentrations within droplets. For example, Wang and colleagues utilized a droplet-merging approach to add oxidase enzyme/horseradish peroxidase/Amplex UltraRed assay reagents into individual droplets to sense metabolites of interest (Table 1-1)[14]. In cases where co-flow or co-culturing are not available (and instances when

overcoming the *Poisson* distribution is not necessary), pico-injection of an assay reagent (such as a chemical mixture, enzymatic mixture, or biosensor) can be used to establish a signal[43]. On a fluidics basis, pico-injection appears similar to a T-junction microdroplet generator (Figure 1-1 & 1-2); however, two electrodes are embedded on either side of the incoming aqueous solution, and the already-formed emulsion flows past the aqueous solution rather than the oil in the case of microdroplet generation. Similar to FADS, a voltage is applied at this junction to merge the incoming aqueous solution with the flowing emulsion, injecting each, evenly spaced microdroplet, with the same amount of liquid[43-45]. This approach can be used to inject various reagents into the cell-containing microdroplets and, thus, enable product detection (or provide for any other inducers as necessary).

1.4 METABOLIC ENGINEERING EXAMPLES AIDED BY MICRODROPLET TECHNOLOGIES

We now turn our attention to actual metabolic engineering examples that have been achieved using microdroplet-based sorting systems. As with any high-throughput assay (especially for ones that rely on fluorescence detection and/or sorting), linking the small-molecule product to a detectable signal is essential[33]. With FADS as one of the more high-throughput sensing and/or sorting modalities, it is essential to transduce product concentration into a measurable fluorescent signal. This transduction can be accomplished by detecting the product itself, through the use of fluorescent dyes or RNA aptamers, through improved growth identified via GFP or RFP tagging microbes, with the use of chemical or enzyme-linked reactions, or finally, through transcription-based biosensors (Table 1-1, Figure 1-3). FADS is not the only assay option for microdroplet assays because other applications, including density-based sorting and MS, have been used to identify increased growth and metabolomic profiles, respectively[46,47]. As a general rule of thumb, if an assay has already been developed for screening the molecule of interest in microtiter

plates, it can readily translate into microdroplets. While we focus here primarily on FADS-based selection technologies with increasingly complex detection schemes, we briefly discuss the use of magnetic nanoparticles and MS.

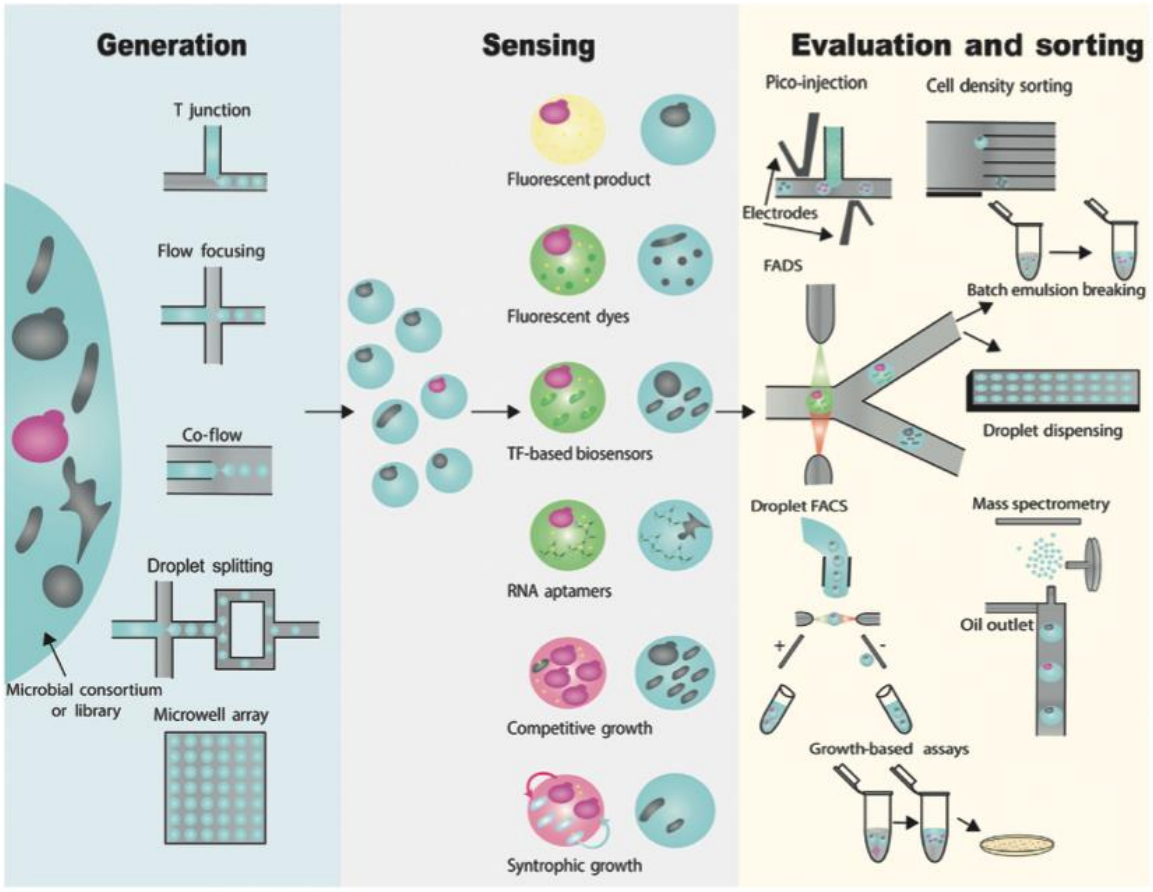

Figure 1-3: *Microdroplet Generation, Sensing, and Sorting Techniques.* Generation of microdroplets can be performed in many ways, with the most popular for metabolic engineering purposes being flow focusing. This is due to the control it allows for making specifically sized emulsions, based on the width of the intersection of the two phases, as well as the flow rate of each phase. Product-sensing techniques are based on the availability of the technique, and whether a secondary way to sense the molecule of interest is needed or if the product is itself fluorescent. Finally, analysis and sorting techniques are used to identify library members that have the phenotype of interest, the most common version of this kind of sorting is fluorescence-activated microdroplet sorting (FADS).

1.4.1 Direct Product Detection

In the most ideal case, the product of interest is itself fluorescent and, thus, can be directly used in FADS. This inherent trait was utilized in the example described earlier for riboflavin production in *Yarrowia lipolytica*, as well as by Chen and colleagues in *Lactococcus lactis* for the same product[48] (Figures 1-1 through 1-3). In particular, the inherent fluorescence of riboflavin was leveraged to directly sort through a mutant library of riboflavin-producing microbial strains. Wagner and colleagues then utilized this test case to compare traditional single cell FACS and microdroplet-based FACS[28]. In the end, strains isolated via microdroplet-based sorting were able to produce 54-fold higher riboflavin than the initial wild type (nearly 100mg/L in flasks on average for the top 15 strains) in a mode that favored secreted product (nearly 90% of riboflavin secreted). While this situation was clearly ideal with respect to innate detection, it is also rare, and, thus, nearly all other applications require some means to transduce chemical concentration into a detectable signal.

1.4.2 Fluorescent Dyes for Product Sensing

Fluorescent dyes that are unquenched upon exposure to the desired condition are a straightforward way to measure improved production or enzymatic activity. As mentioned earlier, the diffusion of molecules across the water-oil barrier is an important consideration during the design phase of microdroplet-based assays[35,36,49] and, thus, it is important to seek fluorescent dyes with desired properties. In a recent example, an assay was developed to detect low pH using a designed dye title Probe 1[40]. With this dye, it was possible to use microdroplets to detect for high levels of lactic acid produced by a *Bacillus coagulans* strain[40]. Sorting of a mutant library from this producer strain was conducted via a w/o/w emulsion in a traditional FACS instrument, similar to that of Wagner and colleagues

described earlier for riboflavin. Ultimately, this approach led to the isolation of a mutant with a 52% increase in lactic acid production compared with the parental strain (Table 1-1)[40]. While successful and suitable for commercially available FACS machines, these w/o/w approaches are less stable and can contribute to microdroplet shrinkage with yeast growth, resulting in nonhomogeneous population being fed through the FACS sorter[32].

Beyond a direct fluorescent dye, it is possible to use microdroplets for enzyme evolution when the substrates can be turned into a fluorescent product. As an example, Ma and colleagues leveraged molecule diffusion across the water-oil interface and identified molecules that became trapped within microdroplets upon enzymatic conversion and, thus, could serve as probes[49]. After screening several synthetic substrates for phosphotriesterase activity from *Geobacillus kaustophilus*, one derivative of coumarin was found to have the dual properties of hydrophobicity to diffuse across the aqueous-oil interface and suitability to be converted into a fluorescent substrate that would be retained in the droplet[49]. Another group also utilized microfluidics screening to improve enzyme activity indicators for future microdroplet assays[50]. Fenneteau and colleagues utilized microfluidic droplet screening to identify fluorophores in a small-molecule library that, when cleaved by peptidase activity, released a hydrophilic molecule that enabled fluorescent signal to remain trapped inside the droplet[50]. This screen resulted in the identification of a novel rhodamine-based substrate that had an approximately sevenfold increase in fluorescence in the presence of *Streptomyces griseus* amino peptidase.

In another, more simplified example, Qiao and colleagues screened for increased lipolytic activity of enzymes across different microbial strains[45]. In this case, fluorescein dibutyrate (FDB) was pico-injected into preincubated microdroplets to enable lipase screening with a reduction in spontaneous FDB hydrolysis to limit false positives[45]. This emulsion was subsequently sorted via FADS and seven individual microbial strains were

identified with significantly higher lipase activity[45]. In a similar approach, Huang and colleagues utilized a fluorogenic substrate of α-amylase to screen UV-mutagenesis-generated libraries of *Saccharomyces cerevisiae* for increased protein secretion[51,52]. This approach identified key mutations leading to increased α-amylase secretion and indicated the possible generalizability of these results by showing that some of the identified mutant targets also increased secretion of 1,4-α-glucosidase. The use of a co-encapsulated fluorogenic substrate takes advantage of the secretion phenotype-genotype linkage in microfluidics in an elegant way.

In addition, by utilizing lipid-staining assays within droplets, Kim and colleagues were able to isolate improved lipid production and cell growth by encapsulating an ethyl methanesulfonate (EMS)-mutagenized library *Chlamydomonas reinhardtii* and utilizing an on-chip fluorescent lipid staining process (BOPIDY)[53]. This screen increased growth rate by 1.35-fold and lipid production by 2.75-fold (Table 1-1).

A final technique leveraging fluorescent dyes in microdroplets utilized RNA aptamers with dual functionality, wherein binding of the desired substrate subsequently allowed for fluorescent dye activation upon a conformational change[34]. Many RNA aptamers capable of binding common proteins and small molecules, including streptavidin and tyrosine, were introduced into droplets and shown to exhibit quantitative product detection. In this work, Abatemarco and colleagues used FADS to sort aqueous-in-oil microdroplets containing plasmid libraries of 10^5 in size along with RNA aptamers for improved production of tyrosine and streptavidin, leading to 28-fold and threefold improvements, respectively (Table 1-1).

1.4.3 Cell-Based Biosensors

More complex sensing systems that involve intercellular biosensors have also been developed. When deployed, some of these assays rely on a fluorescently tagged cell along with cell viability to transmit information about product concentration[29,54]. One recently developed system, termed syntrophic co-culture amplification of production phenotype or SnoCAP, used the fluorescent signal from increased growth of a fluorescently tagged producer strain to identify high producers of 2-ketoisovalerate and L-tryptophan[29]. Specifically, there is a synergy within SnoCAP that uses sensor strains auxotrophic for the product of interest and a production strain that is auxotrophic for a molecule produced by the sensor strain. In this regard, growth of each strain is directly related to the product of interest based on predetermined mathematically modeled transfer functions[29]. By fluorescently tagging the sensor strain in a random mutagenesis library, it was possible to use FADS to select for higher production of 2-ketoisovalerate, ultimately resulting in over 100-fold enrichment of the highest 2-KIV secretor (Table 1-1)[29].

Applications of photoswitchable proteins for metabolic engineering have included the detection of *p*-coumaric acid, via the reconstitution of a photoactive yellow protein domain[55]. Photoswitchable proteins have also been used in microdroplets to screen for increased growth phenotypes of both *Escherichia coli* and *S. cerevisiae*[53]. In this scheme, cells harboring plasmids that encoded photoswitchable proteins were encapsulated in microdroplets and allowed to proliferate. Using a 100x objective, directed illumination of microdroplets with high cell growth allowed for a switch of this green photoconvertible fluorescent protein to red-only in these specific droplets. The emulsion was then broken, and cells were sorted via FACS for RFP signal to select for cells that were a part of the high-growth population. While this technique did not directly sort for production

phenotypes, it has the potential to be used in these applications, especially when combined with the syntrophic growth approach described earlier.

Cell-based biosensors also commonly use transcription factors that respond to a product of interest by altering transcription rates[56]. While biosensors integrated within the production strain itself relay information about intracellular concentration, a separate biosensor cell allows for secreted concentration measurements in both microwell and microdroplet scenarios. A biosensor developed in *E. coli* for *p*-coumaric acid was used with microdroplets to sort for high-producing yeast strains in a library comprising 28 strains with known titers[57]. Using iterative rounds of sorting, it was possible to use FADS to obtain an enriched pool of high-producing library members (Table 1-1). Collectively, the use of biosensors and other responsive proteins can be used in conjunction with microdroplet systems to detect, and subsequently sort for, small molecule production.

1.5 CONTINUING MICRODROPLET APPLICATIONS FOR METABOLIC ENGINEERING

Microdroplet technologies are expanding with respect to both potential applications and capabilities. Some companies have commercialized microfluidic chips for droplet generation and sorting, including Dolomite and Dropletex, to allow for more consistent assays. In addition, several companies, including Missionbio, Sphere Fluidics, and On-chip biotechnologies, have begun to commercialize fully integrated benchtop systems that minimize the microfluidics expertise required to enable this technology. While it is still early, both the capacity and expansion potential of these systems should enable a variety of microdroplet manipulations and sensing modalities over time.

Following the ability to sort microdroplets based on fluorescent signal, the capacity to then separate these microdroplets and index them would allow for further downstream analysis. To this end, Qin and colleagues developed a microdroplet sorter and dispenser

that deposits fluorescently sorted microdroplets into wells of a microtiter plate, allowing for the indexing of microdroplets along with the acquired fluorescent signal[58]. Ultimately, this technology can provide for a seamless microdroplet to microtiter plate workflow that bypasses some of the issues with pooled sorts. In addition, recent advances in microdroplet analysis with MS allow for label-free single cell droplet proteomics using microwells[59,60]. This process is essentially a scaled-down microtiter plate in which a droplet array is made to contain a single cell with a catalogued position[59]. In addition, enzyme activity assays utilizing MS to detect the presence of product have been optimized for use in conjunction with microdroplets[61]. This technology could be paired with microbial production strains to query for molecules or intermediates of interest. This type of approach can be particularly helpful while sequentially building a complex natural product pathway with unknown catalytic steps. However, a fully integrated system in which MS replaces fluorescent readings for product detection has not yet been fully actualized owing to difficulties in maintaining a viable sample indexed to a specific spectrum and limited quantitative measurements. Ideally, MS applications would allow for the sampling and sorting of microbes in droplets, while maintaining a viable sample in the droplet measured, or applying previously mentioned indexing technologies to perform MS on samples of individual droplets.

Ultimately, the future of microdroplet technologies for optimizing microbial product generation is bright. Advances in assays for fluorescent in-droplet product sensing allow for applications using FADS for fast fluorescence-based sorting of large libraries at speeds of up to 30Hz[13]. Other developments in sorting technologies of droplets will allow for indexed dispensing of sorted droplets or MS analysis that can catalog contents while maintaining sample viability. Despite the expanse of future applications, these technologies have already helped alleviate the bottleneck of the test step in the design-build-test cycle

of metabolic engineering for some products, and their future directions are promising for improving the microbial production of an array of desirable small molecules and proteins. It is anticipated that the use of microdroplet systems for the detection of natural products will likewise speed the process development of both known and unknown pathways in a variety of host organisms.

1.5.1 Developing Generalizable Microfluidic Applications for Metabolic Engineering and Synthetic Biology

In my dissertation, I elaborate on the work that I have completed to not only develop generalizable microfluidics protocols for applications in metabolic engineering, crop sciences, microbial consortia phenotyping, and protein engineering, but also apply novel synthetic biology tools for genetic perturbation and expression manipulation. This work expands the scope of cell types compatible with the microfluidic system from standard lab strains to microbial consortia and complex eukaryotic microorganisms. In **Chapter 2,** I apply a guide RNA library to identify gene targets for knock down or increases in expression, specifically for growth on alternative carbon sources. This uses previous work, performed by Dr. Matt Deaner, titled STEPS, or Systematic Testing of Enzyme Perturbation Sensitivities. This identifies rate limiting steps in alternative sugar or glycerol consumption, and importantly, showcases the power of this tool by identifying essential genes as knock down targets, thus illuminating perturbations that would not have been found in a traditional knockout library. In **Chapter 3**, I discuss developing a pico-injection based generalizable microfluidic platform to sort for improved production of secreted small molecules. Due to limitations in fermentation time and co-culture incompatibility, a pico-injection system in which cell-based biosensors are added to droplets after initial incubation allows for a mix-and-match approach. This system was also applied to a

million-member library to showcase the throughput capabilities of fluorescence-based microfluidic droplet sorting. In **Chapter 4**, I detail applications of this generalizable system for identifying distinct growth-dependent and growth-independent production phenotypes in *Yarrowia lipolytica.* This takes advantage of the successful co-encapsulation strategy from my previous work, sorting at different timepoints to achieve early-fermentation production or higher final titer. This approach identified distinct production phenotypes from two different gene perturbations in a PiggyBac library.

In **Chapter 5**, I discuss applications of the microdroplet system for microbial consortia screening to identify strains capable of extracellular electron transfer (EET). Utilizing a fluorescent probe designed by Gina Partipilo from the Keitz lab at UT Austin, we designed an oxygen-limited workflow for the droplet system to allow for anaerobic culturing of microorganisms. This resulted in enrichment of a lakewater sample for organisms capable of performing EET, identified via rDNA sequencing. In **Chapter 6,** I detail applications of this system for cytoxicity assays in collaboration with Bayer crop sciences. I designed high throughput assays for toxin screening utilizing SYTOX dye to identify insect cells that were killed via pico-injection of an active toxin. This work utilized both the lab-strain of insect cells, Sf9, as well as primary cells harvested from insect guts. I utilized pico-injection of increasing toxin concentrations to develop a dose-response curve in droplets for Sf9 cells and a known toxin. In addition, I used this known interaction to sort a mock-library treated with both active and inactive toxin. This sort was reliant on cell death as a readout, so to determine successful sorting, I developed a qPCR assay to identify enrichment of active toxin over inactive in sorted samples compared to an unsorted control.

Finally, in **Chapter 6**, I discuss developing a thermal shift assay for secreted enzymes in droplets. Utilizing a protein secretion system developed by Dr. Hongyuan Lu,

I optimized protein production in droplets via this secretion system, and scoped droplet heating and sorting applications to be able to sort an enzyme mutagenesis library. In **Chapter 7,** I conclude this dissertation by discussing future work utilizing this droplet system, and implications associated with developing these assays for synthetic biology and microfluidic screening.

Chapter 2: Bi-Directional Titration of Gene Expression using a Pooled CRISPR guide RNA Approach[1]

2.1 Chapter Summary

As mentioned in the previous chapter, gene perturbation libraries are the anchor of the Design and Build steps in the Design Built Test Learn cycle canonical to metabolic engineering. Traditional genetic modulation approaches are typically restricted to binary perturbations, single sided titrations (either graded up-or down-regulation), or individual gene expression modulation. In this chapter, I discuss utilizing a library approach that can simultaneously modulate gene expression in a metabolism-wide manner. This library, when coupled with next-generation sequencing, allows for the identification of gene perturbations that would have been missed by classic approaches. For the examples tested, this library identified targets that improved growth on alternative carbon sources as well as improved production of betaxanthins, leading to the identification of unique gene targets. By using this approach, we identified essential genes and other targets that were missed by classic genetic approaches. Some targets identified, when modified by their identified guide RNA for improved growth or production, perform better than their knockout equivalents when knockouts are possible to make.

2.2 Introduction

Classic genetic approaches for identifying gene targets have traditionally been limited to binary modifications consisting of either deletion or strong overexpression[62-64]. This approach is not ideal for identifying all targets, including essential genes that need to

[1] Adapted from: Emily K. Bowman, Matthew Deaner, Jan-Fang Chen, Robert Evans, Ernst Oberortner, Yasuo Yoshikuni and Hal Alper, 2020. *Bidirectional titration of yeast gene expression using a pooled CRISPR guide RNA approach* ***PNAS*** *117 (31)* 18424-18430. DOI: https://doi.org/10.1073/pnas.2007413117. EKB performed the growth enrichment assays, NGS analysis, final target validation, and wrote the manuscript for this research article.

be repressed or any genes whose expression only requires modest modulation. Moreover, even most emerging genetic tools including CRISPR/CRISPRi, TALEN and RNAi still mainly invoke binary gene expression modulation[65-67]. While these approaches have been successful in ascribing functional annotation and identifying dominant gene targets, they are often insufficient when seeking to optimize cellular metabolic function, or when attempting to identify other salient genetic targets. For example, deletion library screens are unable to identify essential genes whose knockdown may in fact be positively correlated with phenotype. Moreover, it has long been recognized that optimal expression of genes resides at intermediate expression levels between the extremes of complete gene deletion and strong overexpression[68]. Traditional approaches fail to identify these targets that are only effective at intermediate expression levels. Thus, new approaches are needed to both enable high-throughput identification of targets and complement the limitations of most genetic screens.

Transcriptional element libraries (such as promoter libraries[69]) have traditionally been used to enable graded gene expression for individual pathway applications for phenotypes such as small-molecule production[70]) and the consumption of alternative carbon sources[71,72]. However, these techniques are often not accessible to genome-wide, high-throughput implementation. Recent advances to establish graded expression-level libraries have utilized RNAi approaches with either micro-RNAs or full-length complementary mRNAs to afford two levels of gene knockdown[14]. More recently, a CRISPR interference (CRISPRi) based single-guide RNA (sgRNA) library implemented in mammalian cells to create a broader spectrum of down-regulation to identify causative genetic targets for fitness phenotypes[73]. In addition, one group utilized dCas9 fused to VP64-p65-Rta (VPR) to identify finely tuned, ideal levels of gene expression for 168 different genes, by tiling each one 21 times[74]. While these approaches are all working

toward an accessible high-throughput library for graded expression ranges, they fall short of achieving both knockdown and over-expression capacity and thus lack the full range of possible gene expression perturbations.

In this chapter, I showcase the implementation and utility of a plasmid-based library methodology that affords bidirectional titration of yeast gene expression in a manner that is complementary to traditional genetic approaches (Figure 2-1A). This worked leveraged the previously reported STEPS approach to design a panel of single-guide RNAs that tile promoter regions for all 969 genes represented in the ito977 genome-scale metabolic model[75]. These sgRNAs were synthesized in a pooled format through collaboration with the US Department of Energy (DOE) Joint Genome Institute (JGI) Synthesis Science Program. By coupling this sgRNA library with a previously established CRISPR-dCas9 system, we can take advantage of dCas9 fusions with Mxi1 for heterochromatin formation or VPR for recruitment of the mediator complex, thus creating graded up and down regulations, respectively[76].

2.3 Library Enrichments on Alternative Carbon Sources Identify Novel Targets

As in the first set of case studies, I utilized growth-based enrichments on alternative carbon sources using the common laboratory strain of *S. cerevisiae,* BY4741. Here, we chose to select enrichments on both glycerol and galactose due to their industrial relevance (galactose comprises a significant portion of marine biomass, and glycerol is produced in large quantities as a byproduct of biodiesel transesterification[77,78]) as well as extensive prior studies on these carbon sources using classic approaches[4]. Prior to conducting a deep-sequencing analysis of enrichment using these pools, we first validated the enrichment capacity of these libraries using a repeated sub-culturing colony isolation, and sequence analysis approach (Figure 2-1B and Figure 2-2). In these trials, individual beneficial guide

RNAs were isolated and confirmed to not only show enrichment over time, but also confer a growth advantage when retransformed.

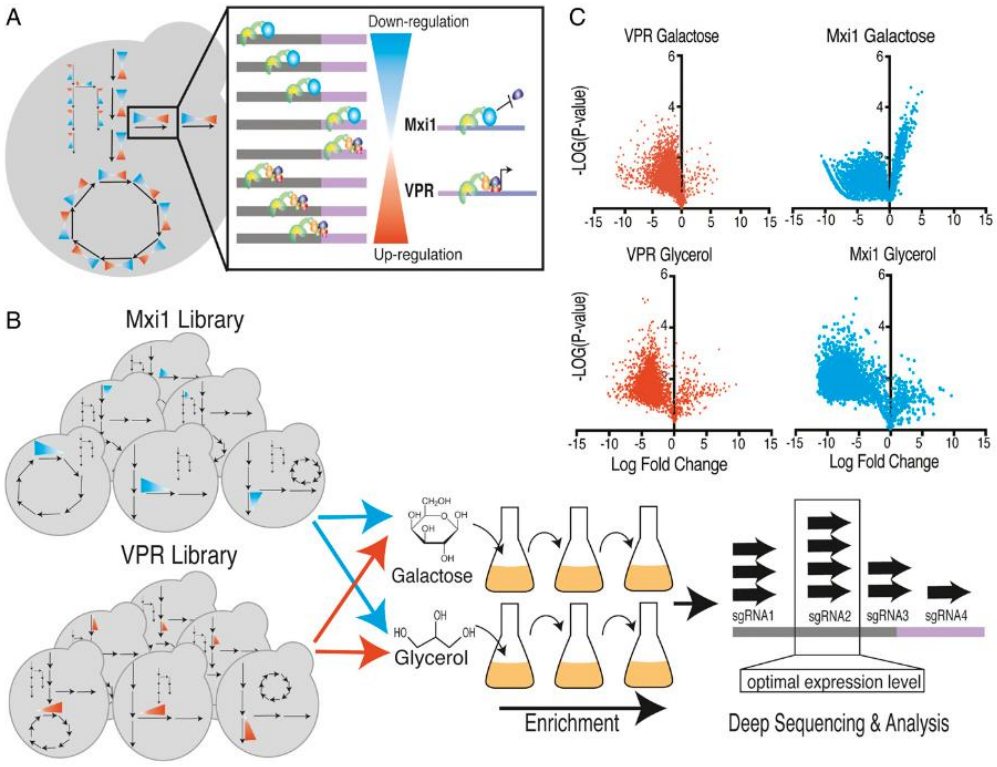

Figure 2-1: *Enrichment of bidirectionally titrated metabolism-wide library.* (A) A metabolism-wide bidirectional titration panel of sgRNAs was synthesized based on the ito977 model of metabolism and enabled via dCas9 fused to either Mxi1 or VPR. These fusions allow for graded knockdown or overexpression of targeted genes of interest, respectively. (B) Enrichments were performed on alternative carbon sources of galactose and glycerol using the Mxi1 and VPR libraries independently to allow for ease of deep sequencing. Final populations from each enrichment were deep sequenced to identify statistically significantly enriched guides. (C) Volcano plots of guide enrichment for each condition. These plots show that most guides in the library were depleted, indicating that most perturbations were outcompeted in growth on alternative carbon sources.

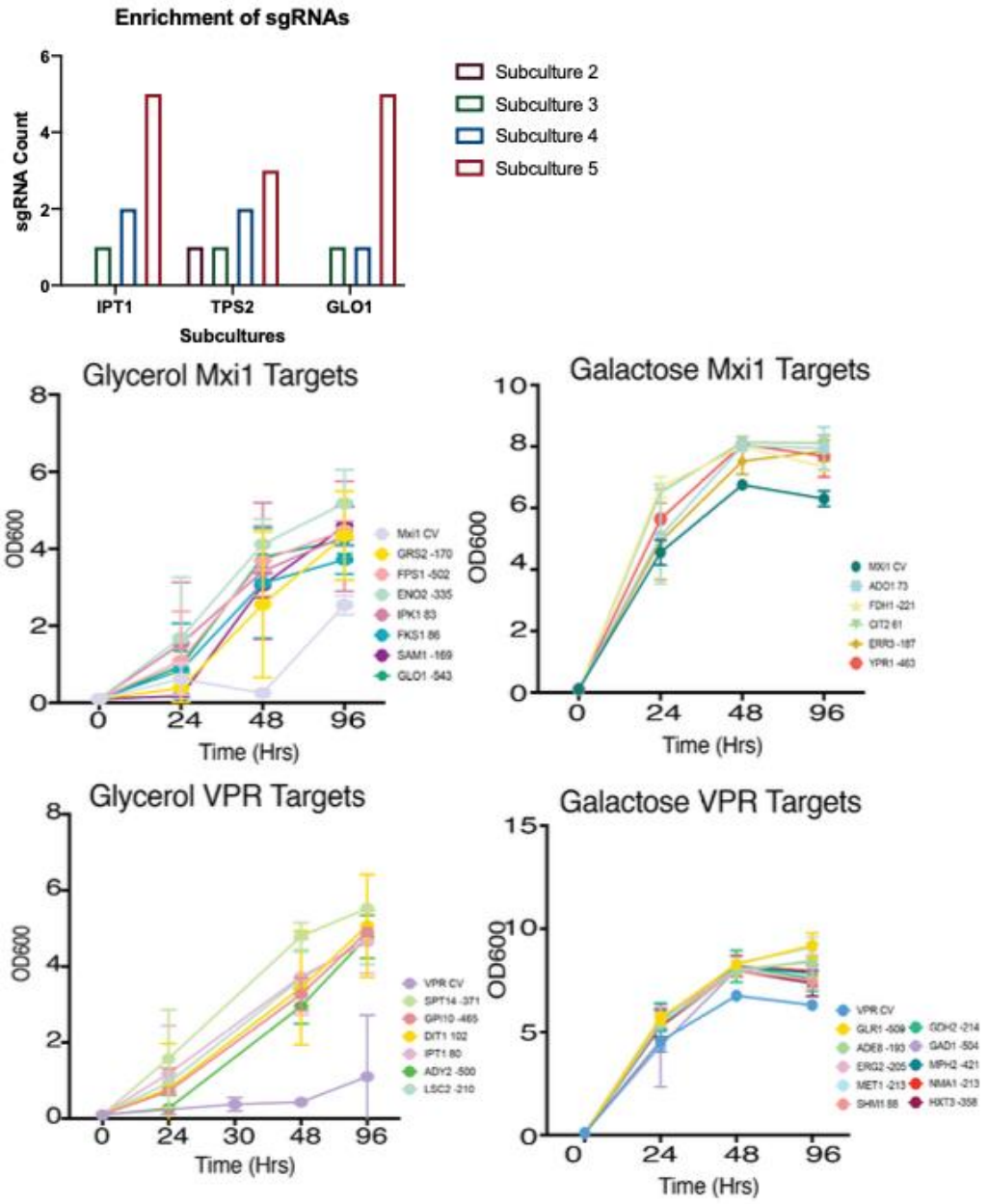

Figure 2-2: *Confirmation of Enrichment over time.* Isolates were sequenced individually and plasmids with target guides re-transformed into BY4741. a) Bar graph showing number of reads of representative individual guides from isolates submitted. This indicated enrichment occurring over time. b) Isolates from bar graph – and other identified, for each library under each condition.

Following these validation tests, a full-scale growth enrichment process accompanied by deep-sequencing analysis was used to globally identify targets along with their optimal expression levels (Figure 2-1B). To do so, we chose a partial subculture condition to detect both enrichment and depletion as well as preventing over-enrichment by a few dominant targets. Macroscopic analysis of statistical enrichment and depletion of guide RNAs within the library illustrates that the majority of guides were depleted in the post-enrichment pools, indicating that most perturbations to gene expression are outcompeted in this assay when grown on glycerol and/or galactose (Figure 2-1C).

Given the high-resolution aspect of this dataset (i.e., having both target identifications along with their optimal expression levels), multiple modes of analysis are possible. For example, a cluster analysis allows for a full mapping of gene expression-level phenotype enhancement for both carbon sources (Figures 2-1C, 2-3A, and 2-4). Each major cluster links together targets whose optimal expression profile and patterns are similar. Initial evaluation of these trends and patterns indicates an overrepresentation of guide enrichment at moderate levels of expression (both for knockdown and over-expression) (Figures 2-1B, 2-3A & B, and 2-4).

At the onset, these data illustrate the complementary nature of this approach to coarse-level, binary modification of gene expression. Several gene targets emerge whereby moderate knockdown greatly enhances growth relative to the complete knockout. As examples, knockdown of *IPK2* and *TPS2* resulted in improved growth on glycerol, but the complete deletion of these targets results in substantially reduced growth (Figure 2-3B). Additional examples include identifying guides targeting *FUN26* whose growth showed improvement far over the counterpart deletion of this gene (Figure 2-5). On the opposite end of the spectrum, glycerol-growth-enhancing targets like GRS2 over-expression are

only optimal/functional when targeted at modest overexpression sgRNA localization regions. This point is especially poignant when comparing growth with both lower and higher levels of expression for this target (Figure 2-3C). Examples such as these illustrate the depth of new targets identifiable with this approach.

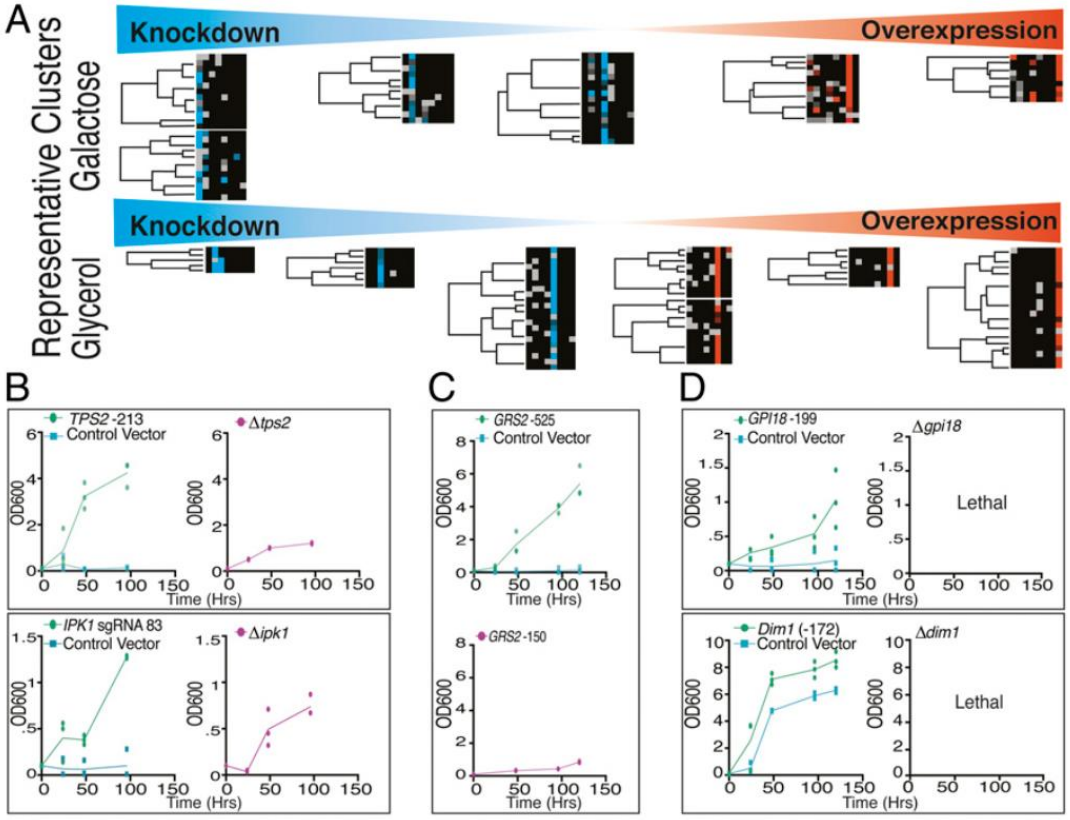

Figure 2-3: *Cluster analysis and subsequent ideal guide confirmation of glycerol and galactose enrichments.* (A) A representative set of clusters for enriched guides are illustrated for both galactose and glycerol. In this representation, gene ID is on the vertical axis, and the bidirectional titration is on the horizontal axis going from strong knockdown to strong overexpression. The predominant level of expression in each cluster is highlighted by blue or orange. Results illustrate a strong enrichment of moderate expression levels across these conditions. (B) An example of intermediate knockdown targets identified from glycerol enrichment is compared with the knockout, demonstrating that moderate regulation was optimal. (C) As an example of a moderate expression target, GRS2 overexpression is highlighted wherein the guide furthest upstream (i.e., slightest up-regulation) was significantly enriched following serial culturing and provided a far improved growth over

the strong expression guide and the wild type. This ability to tune gene expression to specific levels of overexpression is not seen in traditional overexpression libraries. (D) Essential genes were uniquely identifying with this method as down-regulation targets. As examples, *gpi18* was identified from the glycerol enrichment and *dim1* from the galactose enrichment. Both guides were recloned and showed improved growth, whereas the deletion is lethal and thus these represent new targets unseen with traditional approaches.

Beyond visually confirming the premise that different expression levels are required for different subsets of genes, these clusters can be analyzed to determine underlying metabolic trends for these growth phenotypes. For example, through Gene Ontology (GO) analysis, a significant number of phosphate-related metabolic genes were seen to be enriched for the medium knockdown level in galactose selection (utilizing the SGD GO analysis tool) (Figure 2-6). There are obvious differences in associated GO terms between different clusters, with the highest down-regulated genes identified in Glycerol being associated with sulfide metabolism and biosynthesis, as well as sugar catabolism (Glycerol Mxi1 -150). GO terms indicate a medium-level knockdown enrichment in glycerol, of genes associated with amino acid metabolism as well, while medium over expression is primarily associated with co-factor metabolism (Glycerol Mxi1-300, Glycerol VP -300). Medium knockdown from galactose enrichment indicates amino acid metabolism as well as carboxylic acid metabolism overrepresentation (Galactose Mxi1 80). In addition, slightly lower levels of knockdown under this condition were associated with phosphate metabolism (Galactose Mxi1 -300). GO analysis of the cluster of high overexpression resulted in the identification of a single GO-term associated with ATP biosynthesis (Galactose VPR -150). More specifically, the most represented genes within this knockdown level were associated with phosphorous (44.4%) metabolic processes. Gene expression clusters for glycerol consumption indicated significant (P-value < 0.05) enrichment of genes associated with organic acid synthesis (42.9%) in the case of the

highly up-regulated cluster (VPR -150). Phosphorous metabolism is also important for glycerol catabolism, wherein we observed the Mxi1 -500 cluster enriched with genes significantly (P-value < 0.05) associated with phosphorous metabolic processes (50%) and further enrichment of genes associated with carbohydrate phosphorylation (28.6%) in the low-knockdown cluster, VPR +80. Thus, large-scale trends of metabolism can be extracted through this analysis.

The approaches described in this chapter are indeed complementary to more classic genetic approaches such as gene deletion libraries, especially in the capacity to identify essential genes as critical targets for gene knockdowns. The sgRNA library synthesized here contained 86 essential genes. Within this set, 8 of these targets are identified as down-regulation targets in galactose, and 10 of these genes were identified as down-regulation targets in glycerol (with an additional 20 of these genes enriched as over-expression targets) (Figure 2-7). These analyses performed demonstrate the complementary nature of this approach. This analysis illustrates only a 10% overlap between galactose-associated targets identified in this work and deletion targets in the Saccharomyces Genome Database (SGD), highlighting the uniqueness and breadth of targets identifiable when graded expression is considered in a competitive growth environment. These examples in particular are poignant demonstrations of the benefit of an expression titration library, as these targets would not be identified in a full knockout collection. As examples, knock-downs were identified for both *DIM1* and *GPI18*, both of which are essential genes. The knockdown condition substantially improved growth rates in their respective alternative carbon sources, whereas the complete deletion of these genes is lethal (Figure 2-3). Thus, this approach identified a unique set of targets for which galactose and glycerol-improvement phenotypes were not previously ascribed.

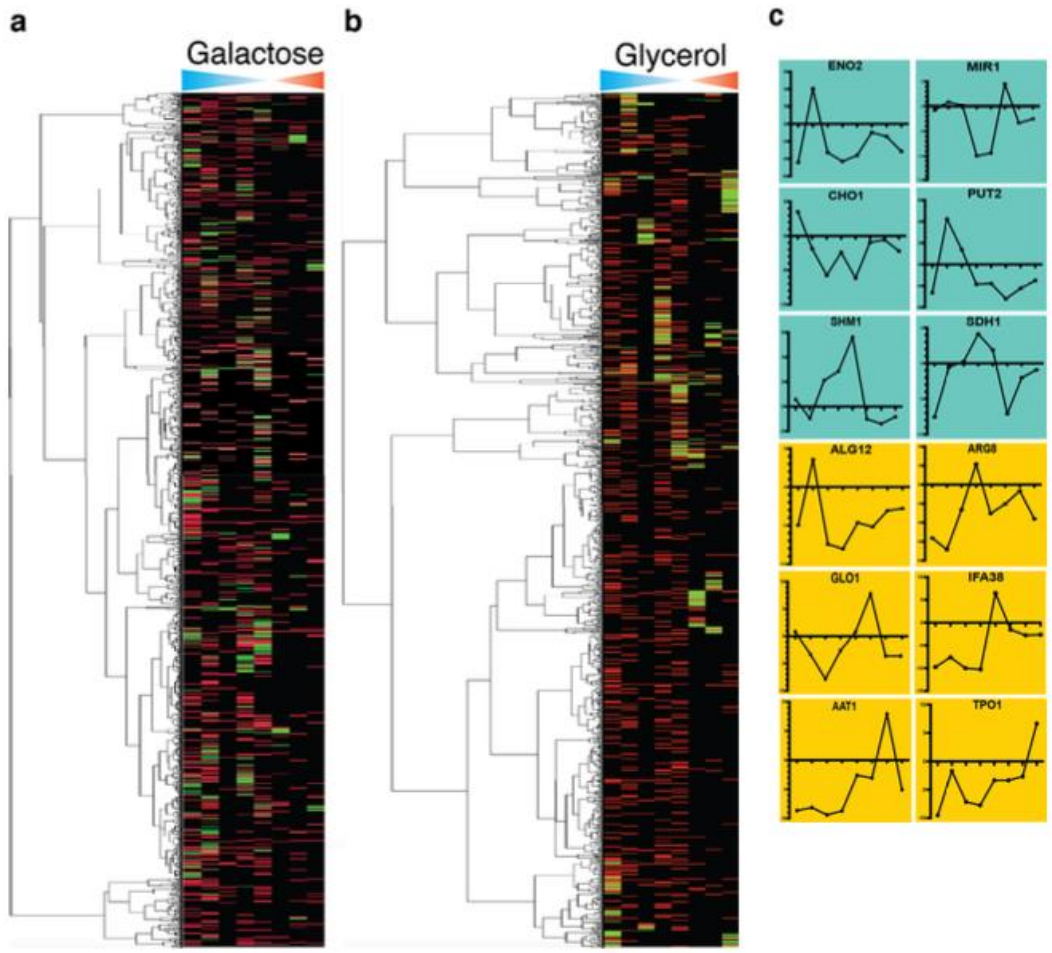

Figure 2-4: (a&b) *Cluster analysis of all sgRNA library members using Euclidean Distance and complete linkage.* Select individual nodes were used for GO analysis (See Figure 2-3). This analysis indicated clustering of optimal expression levels for particular pathways under different conditions. c) Individual gene targets pulled from NGS analysis. Teal indicates guides enriched under galactose consumption conditions, yellow indicates glycerol consumption enrichment. Enrichment of guides was performed in duplicates, with initial growth in raffinose, followed by dilution into alternative carbon source at an OD of 0.1. This enrichment of single guide RNAs was identified via next generation sequencing. These indicate primarily a single optimum level of expression, with some samples showing graded levels leading up to the optimum level (as seen for SHM1, SDH1, and PUT2).

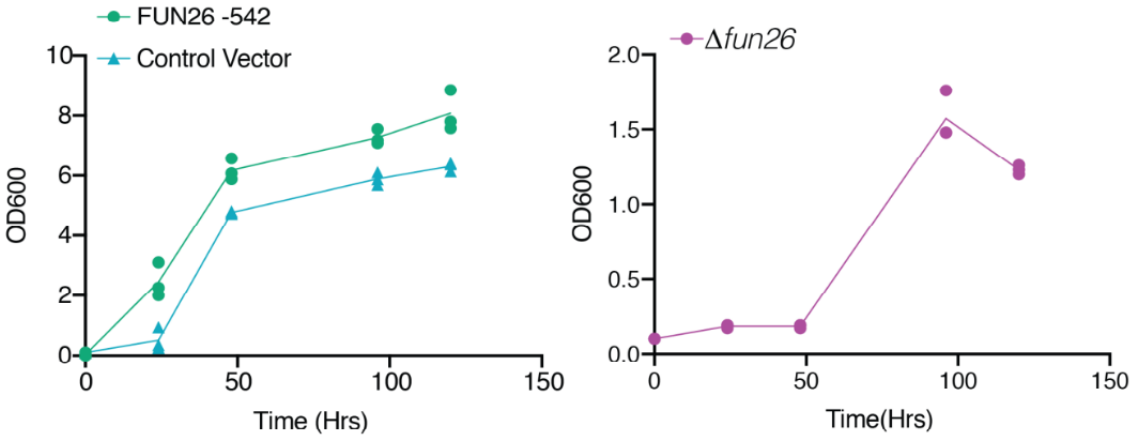

Figure 2-5: Beneficial guide targeting FUN26 at -542 from the open reading frame. When compared to the knockout results in improved growth, versus worse growth from the knockout of *fun26*. Here, the phenotype from the guide RNA creates a beneficial effect on growth over the control.

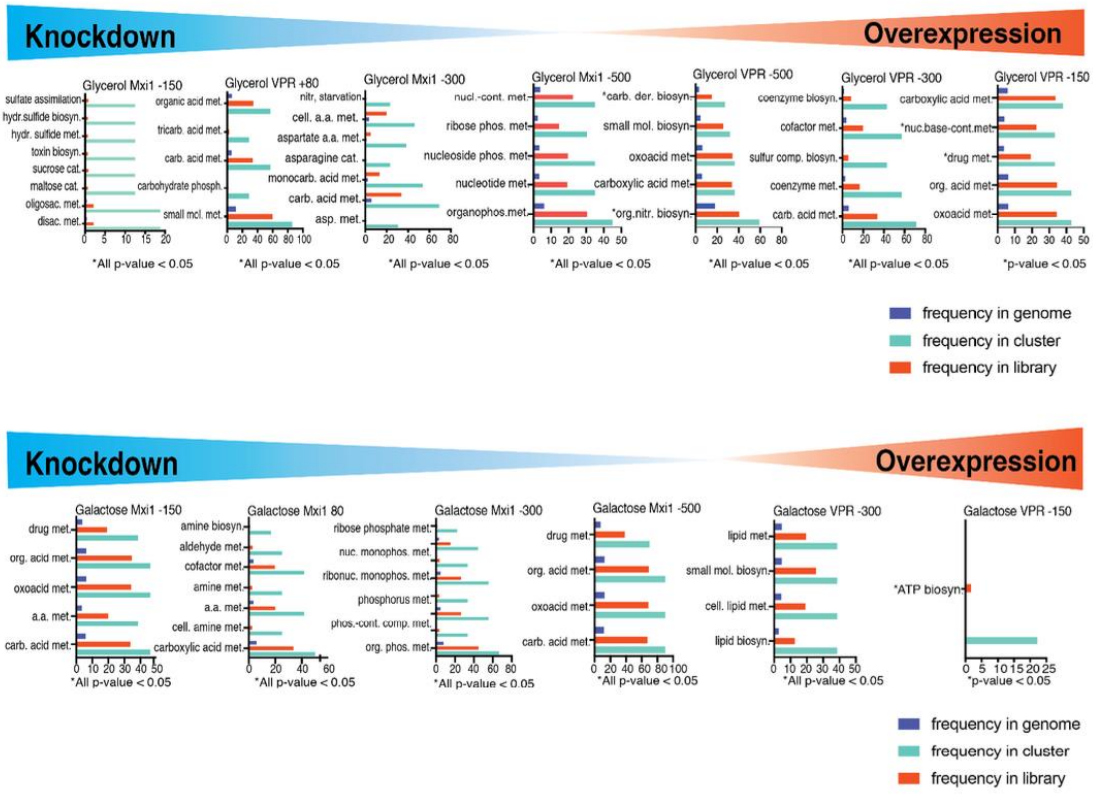

Figure 2-6: Gene ontology (GO) analysis performed utilizing SGD's GO analysis tool. All shown go terms are statistically significant unless otherwise specified

underneath each individual graph. This figure directly corresponds to Figure 2-3, with only a few clusters missing as these lacked significant GO terms.

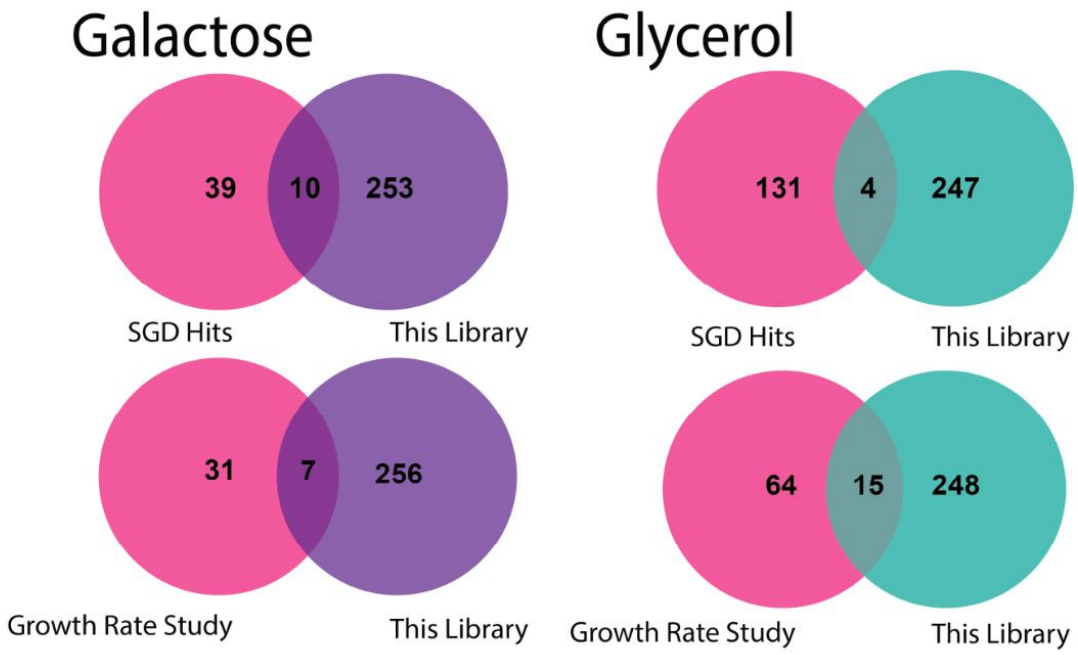

Figure 2-7: Venn diagrams of gene targets identified in this study and those that improved growth rates on alternative carbon sources from Vandersluis et. al, as well as genes associated with improved consumption of glycerol or galactose from the *Saccharomyces cerevisiae* genome database *Yeastminer* application (yeastmine.yeastgenome.org)[79].

Finally, overlap of the targets identified here with previously identified literature targets that serve to improve or hinder growth on alternative carbon sources was evaluated and found to be limited (Figure 2-7). Of particular note, this study utilized a competitive growth enrichment assay, whereas most targets in literature were isolated from knockout collections and utilized individual target analysis. Nevertheless, there is some overlap of targets. For example, *HXT17* is a known over-expression target for improving growth on galactose[80], and this same target was identified through this screen. In some cases, other well-known individual overexpression targets for improved growth on galactose, including *PGM2*, were not identified as statistically enriched, likely due to being outcompeted in this

assay. However, sgRNAs associated with down-regulation of *PGM2* were in fact significantly depleted, thus demonstrating the implication of this enzyme on galactose consumption in general (Table 2-1). Additional known targets such as the detrimental effect of *aim10* and *gal10* deletion growth on galactose[79] are instead complemented by the enrichment for only moderate levels of down-regulation in our library. Collectively, the coupling of our sgRNA library with a competitive growth enrichment adds a dimension of enriching for fitness and optimal expression levels. Thus, while we identified some known targets, we predominantly highlight the knowledge of optimal expression levels. As with any screening process, the mode of library screening is important for target identification, and this library is likewise suitable for individual candidate analysis as has been conducted with the gene deletion libraries.

Name	LFC	Guide Location	Library	Carbon Source
HXT17	2.5	-200	VPR	Galactose
GAL10	2.5	-300	Mxi1	
AIM10	2.5	-300		
HXT10	3.75	-300		
HXT9	2.1	-150		
GAL7	-9	80	VPR	
GAL1	-8.75	-300	Mxi1	
PGM2	-8	-150		
ALD4	-8.5	-150		
HUT1	-8	-300		
GPD1	5.5	-500	Mxi1	Glycerol
TPI1	2.5	-150		
STL1	3	-300		
GUP1	3.5	-300		
GPD2	6	-300		

Table 2-1: Galactose consumption associated genes with corresponding sgRNA library enrichment information. List of genes identified in library enrichment that have been associated with galactose or glycerol consumption as identified by *yeastminer* (yeastime.yeastgenome.org)[79]. Log fold change values have been rounded to the closest quarter for ease of reading.

2.4 Conclusion

This case study highlights the importance of utilizing bidirectional titration of gene expression for identification of novel gene targets as well as their optimal level of expression. This approach is highly complementary to classic genetic approaches that tend to only create binary changes in gene expression. In many cases, we find gene knockdowns of essential genes that greatly impact phenotypes of interest yet have been missed thus far due to a reliance on binary genetic changes. Likewise, we identify targets for which only moderate levels of modulation are optimal. The unique ability to screen for novel targets as well as their optimal level of expression using this library provides a powerful tool for studying genotype-phenotype relationships. As mentioned above, screening applications will be detailed in the next chapter utilizing this guide RNA library in a high throughput screen to showcase depth of analysis can help identify novel targets.

2.5 Materials and Methods

Data Availability

All NGS data can be found at National Center for Biotechnology Information (NCBI) under accession number PRJNA625119. In addition, all code used for analysis can be found at the following link: https:// github.com/emkbowman/Bi-directional-Titration-NGS-Analysis.

Strain Design

Dh10β was used to propogate all yeast expression vectors including those described by Mumberg[81] and pJED103-based[82]. To amplify plasmids, *Escherichia coli* strains were cultivated in Luria broth (LB) or super optimal broth (SOB) media (Teknova) supplemented with 50μg/mL ampicillin or kanamycin (Sigma) with 225 rpm orbital shaking at 37°C. Yeast strain BY4741 (EUROSCARF) was cultured in yeast synthetic complete (YSC) medium containing roughly 6.7g/L yeast nitrogen base (Difco), 20g/L glucose (MP Biomedicals) and 1X CSM-URA, CSM-URA-LEU, or CSM-URA-LEU-HIS (MP biomedicals) depending on the required auxotrophic selection.

Transformations

Transformations were performed as previously described[82]. Briefly, in order to transform Gibson cloning reactions, 3μL of threefold diluted Gibson reaction (NEB, 254ng backbone) was mixed with 50μL of electrocompetent *E. coli* DH10β and electroporated (2mm Electroporation Cuvettes) with a BioRad Genepulser Xcell at 2.5kV. For transformation of ligations, 3μL ligation mix (NEB) was directly added to 50μL competent cells and transformed as above. For ampicillin-marked plasmids, transformants were resuspended in 500μL of SOB, plated on LB agar supplemented with 50μg/mL ampicillin, and incubated at 37°C overnight. Transformants of kanamycin-marked plasmids were recovered in 500μL SOB for 30 min at 37°C and then plated. Individual clones were picked into SOB media containing 50μg/mL antibiotics and incubated at 37°C overnight. Plasmids were then miniprepped (GeneJET Plasmid Miniprep Kit, Thermo Scientific) and sequence-verified via Sanger sequencing. The Frozen EZ Yeast Transformation II Kit (Zymo Research) was used to transform plasmids into yeast. Briefly, between 100ng and 1μg of plasmid was mized with 20μL chemically competent cells prepared by manufacturer's instructions and 200μL EZ Solution III followed by incubation at 30°C for 45 min. Transformations were then plated on YSC+agar plates containing either CSM-URA, CSM-

URA-LEU, or CSM-URA-LEU-HIS and incubated at 30°C for 2 days. Individual colonies were randomly picked in triplicate into 1mL of YSC media and incubated at 30°C for another 2 days. For long-term storage, all yeast strains with the exception of transformed libraries were stocked in 15% glycerol and kept at -80°C in sterile flat-bottomed microtiter plates (Corning) covered with an adhesive aluminum foil seal (Thermo Scientific) and plastic lid.

Cloning Procedures

Oligonucleotides were purchased from Integrated DNA Technologies. Sequences and annotations can be found in Appendix 1. PCR and anneal/extend double-stranding reactions were performed with Q5 DNA Polymerase from New England Biolabs according to the manufacturer specifications. Digestions were performed according to manufacturer's (NEB) instructions. PCR products and digestions were cleaned with a QIA-quick PCR Purification Kit (Qiagen). All vectors were dephosphorylated with Antarctic Phosphatase (NEB) according to the manufacturer's instructions and heat-inactivated for 15 min at 65°C.

All plasmids for expression of dCas9 were derived from the pJED103 vector series acquired from AddGene catalog #46921[82]All RGR plasmids were derived from the dCas9-Mxi1 and dCas9-VPR plasmids previously reported[76]. To construct new RGRs, the dCas9-Mxi1 and dCas9-VPR RGR cloning vectors were linearized at the 5' end of the sgRNA scaffold using the SpeI enzyme, and then a 100bp fragment containing a variable HH-sgRNA sequence was inserted via Gibson assembly to create the full TEF1p-HH-sgRNA-HDV-TKC217 cassette. The 100bp insert fragments were constructed by an anneal/extend PCR using two 60bp oligos (IDT) with 20-bp overlaps at their 3' ends.

Library Enrichment on Alternative Carbon Sources

BY4741 was transformed in a pooled format utilizing a scaled-up Geitz transformation. These libraries were cultivated in 2% Raffinose, CSM-L, YNB to maintain the plasmid and prevent any errant glucose-based enrichment. Libraries were stocked in 1-mL aliquots, mixed 1:1 with 40% glycerol and stored at -80°C. For use, these aliquots were thawed and grown in a 50-mL flask containing 2% Raffinose YSD-L until the OD_{600} was over 1.0. Libraries were then diluted down to an optical density (OD_{600}) of 0.1 and resuspended into 50mL of media comprised of CSM-L, YNB, and either 6% Glycerol or 4% Galactose. Biological duplicates were used for library enrichment (i.e., two separate enrichments were performed for each condition). The 6% Glycerol media was adjusted to a pH of 4.0. These cultures were monitored for growth, and once the OD_{600} reached 1.0, they were serially diluted into a fresh flask containing new media with a starting OD_{600} of 0.1 for three times before conducting a yeast miniprep (Zymo) followed by NGS analysis.

NGS Sample Prep and Analysis

Yeast mini preps (Zymo) were used as a template and amplified utilizing Q5 polymerase and following manufacturer's instructions (NEB) (Primers found in Appendix 1). To avoid sequence bias, only 20 amplification cycles of PCR were used. The product was purified via gel electrophoresis and extracted utilizing a gel purification kit from Qiagen prior to being submitted for library construction and analysis at the UT FBS Sequencing Core. All raw NGS data can be found at NCBI under accession number PRJNA625119. Analysis of sequencing reads was performed by obtaining individual read counts and aligning to the sgRNA library, utilizing code that can be found at the link in the Data Ability segment of this chapter. These were normalized to the total number of reads and then averaged. A two-tailed *t*-test was run on each guide to confer significance of enrichment, determined by a P value of <0.05. Finally, the log fold change was calculated for each guide utilizing Excel and clustered using Cluster 3.0, with city-block distance and

complete linkage settings. This analysis was visualized in Java TreeView, with pixel settings centered at log 1.0.

Growth Analysis of Select Guides

Guides identified via either Sanger sequencing or NGS were either retransformed (if isolated via Sanger sequencing to confirm enrichment of beneficial guides) or recloned utilizing cloning procedures mentioned above and transformed utilizing the EZ yeast transformation according to manufacturer's instructions and plating on synthetic selective media (CSM-L, 2% Raffinose). Three clones were picked from each transformation and grown in 2% Raffinose synthetic defined media and glycerol stocked. Glycerol stocks were taken out and grown overnight in 2% Raffinose synthetic defined media and then diluted down to an OD_{600} of 0.1 in the same 6% Glycerol or 4% Galactose enrichment media from which they were identified. OD measurements were collected roughly every 24 hours.

Chapter 3: Sorting for Secreted Small Molecule Production using a Biosensor-in-Microdroplet Approach[1]

3.1 Chapter Summary

As detailed above, depth of screening is important for identifying beneficial gene targets from large libraries. The design–build–test cycle of metabolic engineering is swiftly becoming test limited owing to advances in DNA synthesis and library design. At the same time, improving extracellular pro- duction requires high-throughput studies to detect extracellular (not intracellular) content. Traditional approaches that utilize chromatography-based analysis or plate-based fluorescence assays are relatively throughput limited and require a high degree of liquid handling. These limitations are avoided here by creating a generalizable approach for small molecule production screening. Namely, this paradigm combines "off-the-shelf" biosensors with large library screening from producing cells in a cell- type agnostic manner.

In this chapter, I discuss the application of a combining producer/secretor libraries with whole-cell biosensors using a microfluidic-based screening workflow. This approach enables a mix-and-match capability using off-the-shelf biosensors through either coencapsulation or pico-injection. I demonstrate the cell type and library agnostic nature of this workflow by showing its application for single-guide RNA, transposon, and ethyl-methyl sulfonate mutagenesis libraries across three distinct microbes (*Escherichia coli, Saccharomyces cerevisiae, and Yarrowia lipolytica)*, biosensors from two organisms (*E.*

[1] Adapted from: Emily K. Bowman, James M. Wagner, Shuo-Fu Yuan, Matthew Deaner, Claire M. Palmer, Simon D'Oelsnitz, Lauren Cordova, Xin Li, Frank F. Craig, and Hal S. Alper, 2021. *Sorting for secreted molecule production using a biosensor-in-microdroplet approach 118(36)* DOI: https://doi.org/10.1073/pnas.2106818118. EKB performed proof of concept pico-injection experiments, all *S. cerevisiae* screens and target validations, *Y. lipolytica* target validation in a clean strain, TAL and Naringenin biosensor validation in droplets and wrote the manuscript for this research article.

coli and S. cerevisiae) and three products (triacetic acid lactone, naringenin and L-DOPA) to identify targets improving production and secretion.

3.2 INTRODUCTION

The canonical design-build-test cycle of synthetic biology and metabolic engineering can enable bio-based production for a diverse range of small molecules[1,83,84]. To this end, recent advances in computational design tools, DNA synthesis capacity, and genome editing technology have certainly expedited the design and build parts of this cycle, while at the same time, they have placed higher demands on testing and screening platforms. However, it is often challenging to develop assays that have a throughput parity matched with library size, especially when utilizing large genome-wide libraries or directed evolution approaches[24,85,86]. Moreover, small molecules are often secreted, and thus the genotype and phenotype in these instances are physically separated. While it can be tempting to use intracellular concentration as a surrogate, it was recently demonstrated that in some cases, only secreted concentration is correlated with total production capacity[28]. As a result, top production strains could be inadvertently lost when using intracellular concentration-based whole cell screens.

As an alternative approach, quickly transducing chemical concentrations into a detectable signal (such as fluorescence) can move rather laborious and throughput-limited chromatography-based screening into a higher-throughput mode that is compatible with technologies such as microfluidics-based fluorescence-activated droplet sorting (FADS)[86]. The simplest applications of this technology have relied on inherently fluorescent molecules (such as riboflavin[28]) or RNA-aptamer based biosensors (such as for select small molecules and a secreted recombinant protein[34]). At the same time, synthetic biology is rapidly identifying and developing suites of biosensors capable of detecting a much larger

breadth of small molecules[57,87-89]. To this end, many biosensors are readily available "off-the-shelf" for a wide range of potential analytes.

In this chapter, I posit that microdroplet technologies offer a unique potential for high-throughput library testing, as this approach will physically encapsulate cells (and their corresponding secreted molecules) with external biosensor cells that can transduce secreted product concentration into a fluorescent signal (Figure 3-1). Supporting this approach, previous literature has characterized the differences in utilizing fluorescence-activated cell sorting (FACS) versus FADS for sorting based on intracellular biosensors and demonstrated that pico-liter droplets were better than FACS at sorting for improved production of 3-dehydroshikimic acid[90]. It has been demonstrated that droplets are better suited for selecting extracellular metabolic concentrations whereas FACS approaches sort for intracellular accumulation—a phenotype not necessarily correlated with total production[28]. In addition, other work in the field supports the premise for an external biosensor to improve product secretion as well. For example, a *p*-coumaric acid cell-based biosensor placed in droplets was used to sort out a predesigned combination of low, medium, and high producers at a throughput of 300droplets/second (30 library members/second), thus suggesting library sorting potential[57]. Likewise, cocultures of engineered cells to establish syntrophic growth have been used to identify improved producers/secretors from 40,000 variants[29]. While these applications raise the potential for high-throughput library sorting, they do not yet demonstrate a true plug-and-play capability for utilizing cell libraries with off-the-shelf biosensors.

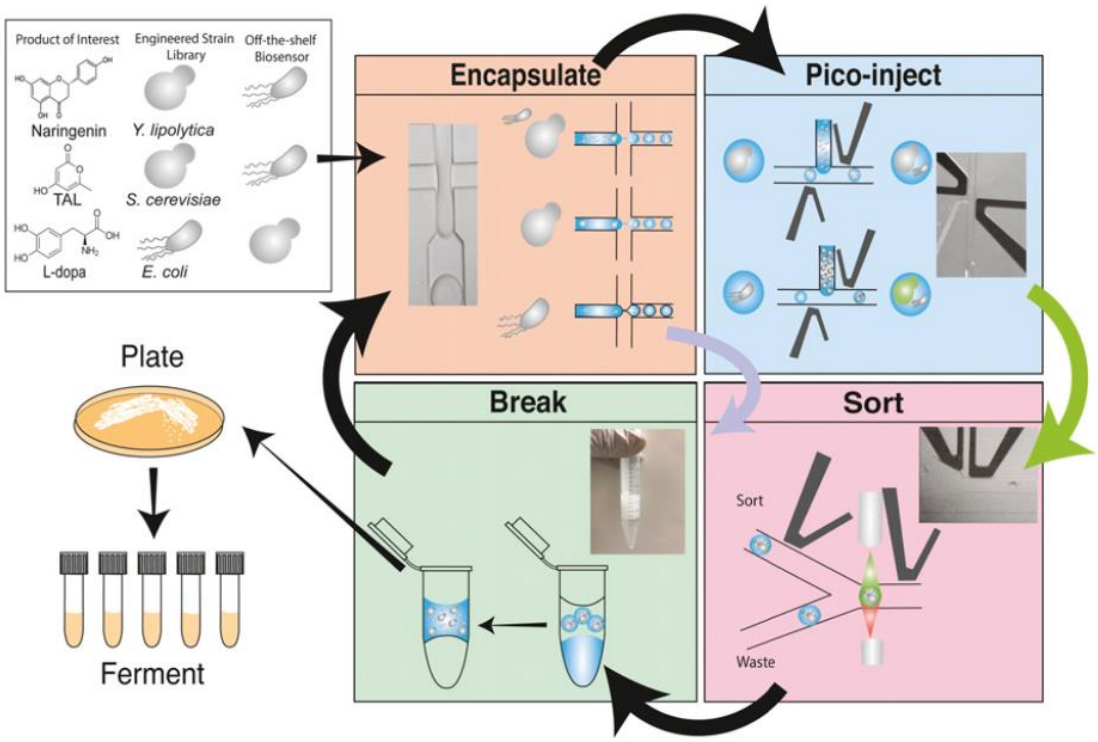

Figure 3-1: Schematic of cell library sorting for improved production/secretion via a biosensor-in-microdroplet approach. The basic workflow here allows for a wide variety of library types, producer cells, and biosensor cells to be combined for screening. The first step of encapsulation is flexible for either coencapsulation of compatible producer and sensor cells or the singular encapsulation of producer cell libraries. Following incubation, pico-injection can be utilized to optimally control the timing of biosensor introduction. Finally, droplets are incubated to allow for signal maturation and then sorted via FADS. The sorted droplet emulsion is then broken, and individual isolates are picked to ferment for production and identify genotype(s) associated with improved titer. Further library enrichments can be performed by iterating this general process.

In this chapter, I detail establishing a generalizable microdroplet-based scheme that introduces off-the-shelf cell-based biosensors via coencapsulation or pico-injection in conjunction with distinct producer microorganism libraries (Figure 3-1). As a demonstration of generalizability, we varied the library of producing cells to include *Yarrowia lipolytica, Saccharomyces cerevisiae,* and *Escherichia coli* as well as varied the

biosensor cells to include *E. coli* and *S. cerevisiae.* Through this work, we demonstrated that the use of microdroplet pico-injection is critical in many cases for achieving optimal biosensor performance. In each case, we were able to identify targets from these diverse libraries (ranging from smaller single-guide RNA (sgRNA) and transposon libraries on the order of 10^3 members, to a large ethyl-methyl sulfonate (EMS) mutagenesis library on the order of 10^6 members) that afford improved production/secretion of various small molecules. Collectively, this established a mix-and-match capacity that leveraged advances in synthetic sensor design in both bacterial and fungal chassis.

3.3 RESULTS

3.3.1 Biocharacterization for Use in Microdroplets

A central tenet of this work is the use of off-the-shelf biosensors that can function in both spent media and microdroplet environments. As test cases for demonstrating this microfluidic sorting approach across three unique production organisms, we focus on the production of the polyketides triacetic acid lactone (TAL) and naringenin from the high-capacity polyketide producer *Y. lipolytica* as well as *S. cerevisiae*[91] and the production of L-DOPA from *E. coli*. To this end, off-the-shelf available biosensor for TAL and naringenin have been constructer using a mutant *E. coli* AraC regulator[87] and *Pseudomonas putida* TtgR transcription factor[88,92], respectively. In both cases, the resulting *E. coli* biosensor strains could import the exogenous polyketide from culture media and transduce concentration to a fluorescent output using a synthetic sensor circuit. These polyketides readily secrete into the culture supernatant of previously engineered *Y. lipolytica* strains as well as engineered *S. cerevisiae* (Figure 3-2A-C)[93,94], thus enabling a coculture droplet-based biosensing approach.

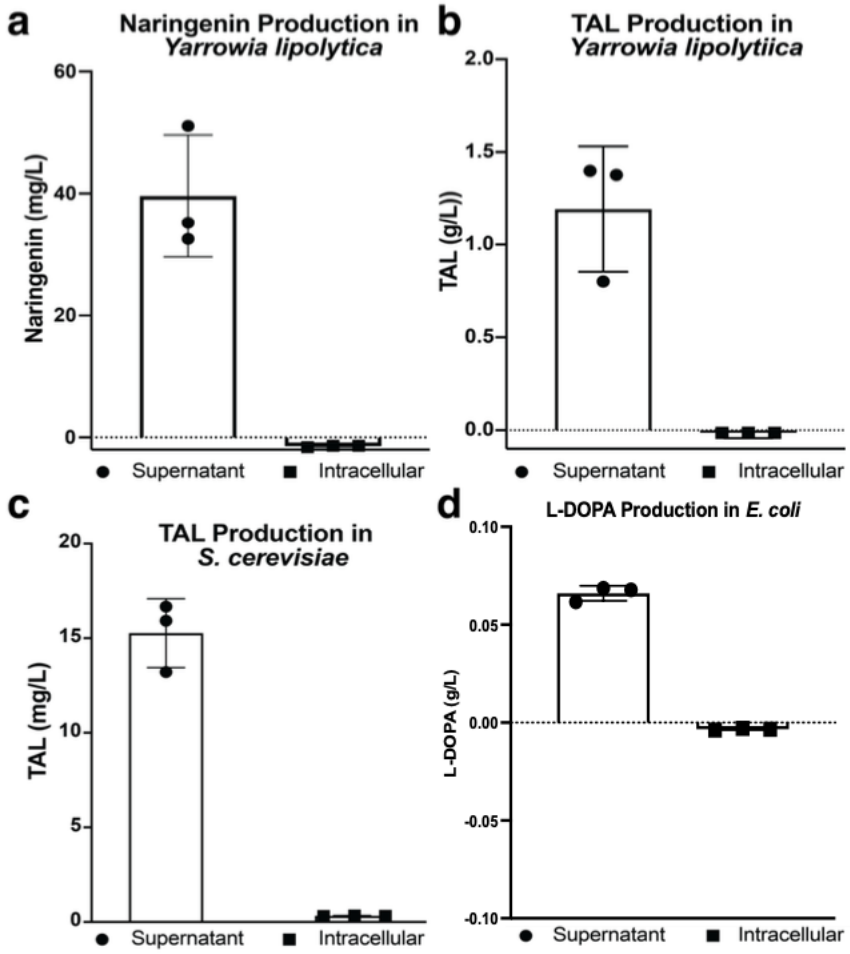

Figure 3-2: Intra-versus extra-cellular concentrations of small molecules of interest in *Y. lipolytica, S. cerevisiae,* and *E. coli.* a&b) Parent strain of Naringenin and TAL producing *Yarrowia lipolytica* secretes both products into the supernatant, making the use of an external cellular biosensor necessary for accurate identification of improved producers. Intracellular concentrations were below the linear range of the HPLC, thus show up as negative concentrations. b) *Saccharomyces cerevisiae* readily secretes most of the TAL it produces into the supernatant as well, thus resulting in the need for an extra cellular biosensor to identify changes that would be seen in improved producers. d) Wild type BL21DE3 *E. coli* readily secretes much of its produced L-DOPA. In order to screen for improved secretors as well as improved overall producers, external biosensors in droplets must be used to sort through as much of a gene perturbation library as possible.

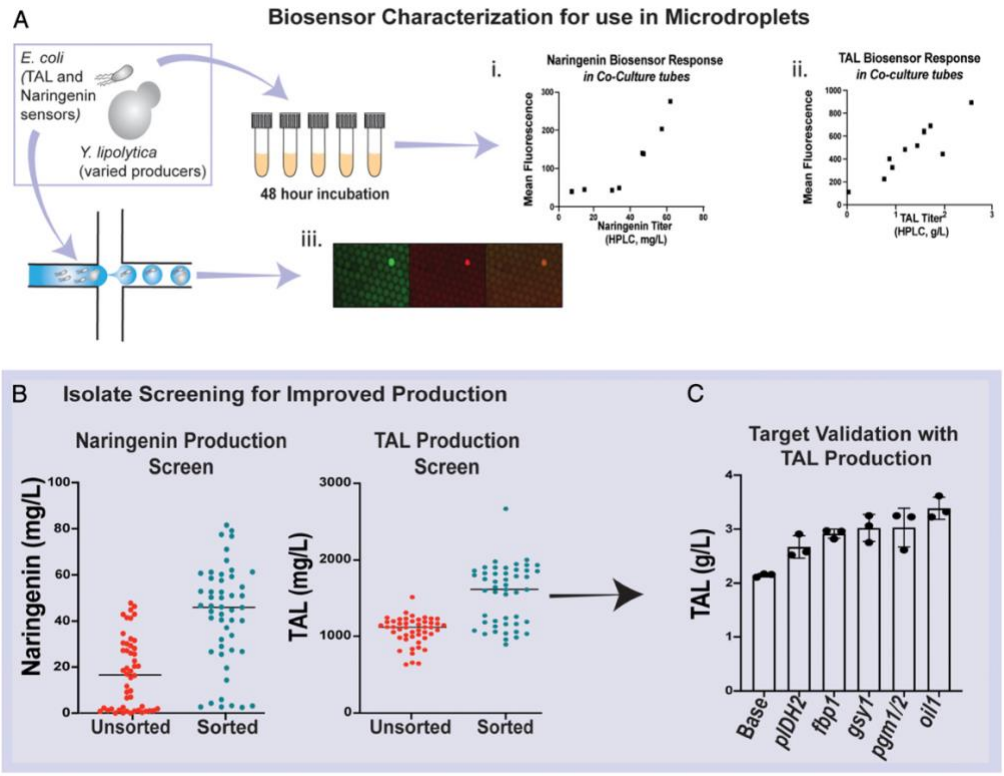

Figure 3-3: Improved *Y. lipolytica* production of TAL and naringenin. (A) Biosensor Characterization for use in microdroplets with *Y. lipolytica.* (*i)* Varied *Y. lipolytica* naringenin producers were cocultured with *E. coli* naringenin biosensors. Final titers were measured for naringenin via HPLC, and biosensor response was quantified via FACS. (*ii*) Varied *Y. lipolytica* TAL producers were cocultured with *E. coli* naringenin biosensors. Final titers were measured for TAL via HPLC and biosensor response was quantified via FACS. Initial titers were previously identified in other metabolic engineering efforts, aiding in isolate selection to demonstrate a range for each molecule of interest[91,95]. (*iii*) TAL Biosensor response in coculture droplets. The *Y. lipolytica* is tagged with mCherry and can be seen overlapping with the droplet showing the highest GFP response in the merged panel. (B) Fermentation of 56 individual isolates from a mock/unsorted control and the sorted sample show significant improvements in TAL and naringenin production (determined via two tailed *t* test with Welch's correction, P value < 0.01). Individual targets were identified from these pools that also demonstrate an overlap in the pathway for production of these small molecules in *Y. lipolytica.* (C) Targets were confirmed via TAL production in a fresh strain using biological triplicate. Error bars represent SD from biological triplicates.

The TAL biosensor is a single plasmid-based system, which was previously characterized, containing GFPmut2 downstream from an araC promoter along with a coexpressed mutant AraC-TAL transcription factor[87]. Coculturing the *E. coli* biosensor with various *Y. lipolytica* TAL producers resulted in a linear response to TAL production (Figure 3-3A, *ii*). The naringenin *E. coli*-based biosensor is a dual plasmid biosensor system whereby one plasmid constitutively expresses a codon-optimized ttgR transcription factor and the other plasmid contains GFPmut2[96] downstream of a modified *ttg*ABC promoter[92]. This biosensor system likewise worked well in cocultures with varying *Y. lipolytica* producers of naringenin (obtained via random integration), accurately assessing naringenin concentration for each (Figure 3-3A). Furthermore, a cross-comparison of these biosensors with specific analyte concentrations in droplets show a similar response to that seen in the plate reader-based assays (Figure 3-4 and 3-5A). These results demonstrate that a coculture-based approach could be feasible in droplets for *Y. lipolytica.*

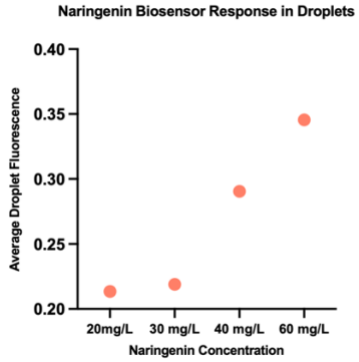

Figure 3-4: Naringenin biosensor response in microdroplets. Here, increasing concentrations of Naringenin (in DMSO) were dissolved in spent *Y. lipolytica* YPD media (20, 30, 40 and 60mg/L). Emulsions of these varying naringenin concentrations were made by co-encapsulating with *E. coli* naringenin biosensor at an estimated rate of 7 cells/droplet. After an 18-hour incubation, emulsions were re-injected into the Sphere pico-sort™ chip to collect average fluorescence of each emulsion.

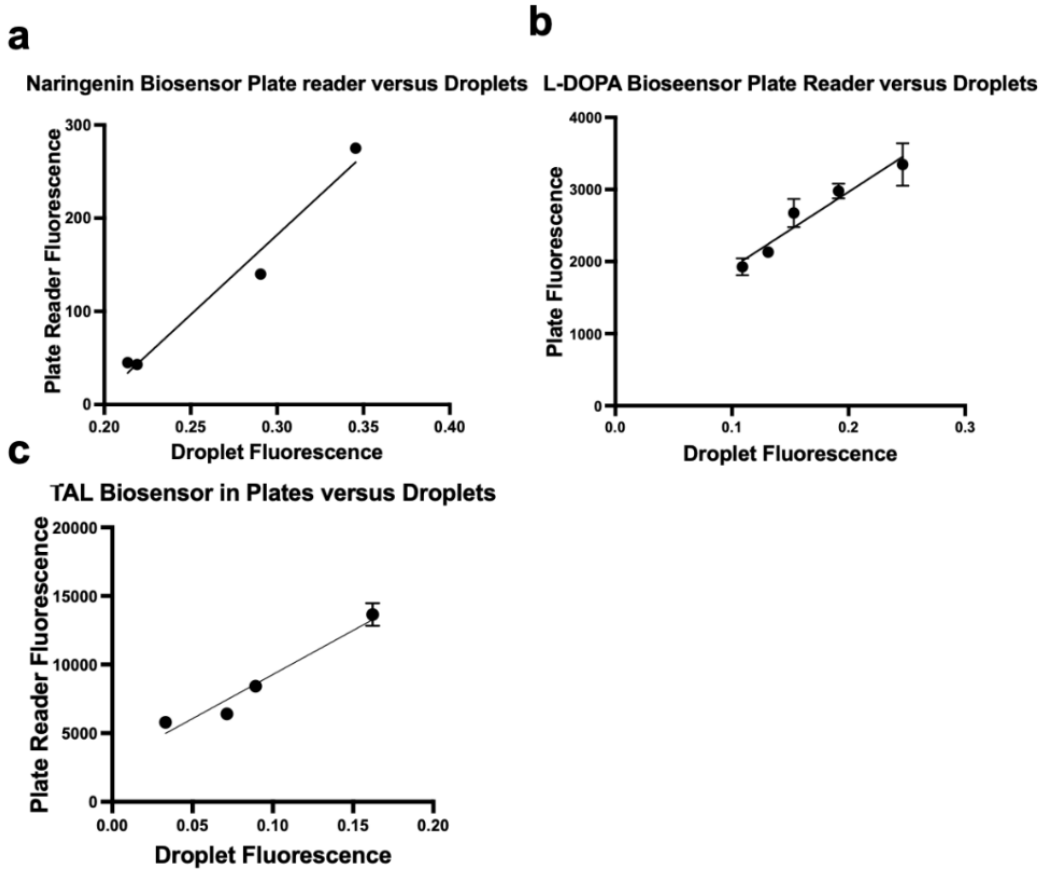

Figure 3-5: Droplet fluorescence vs. plate reader fluorescence of each cell-based biosensor. Here, varying concentrations of analyte were encapsulated wither with their biosensor or pico-injected with the biosensor immediately after encapsulation. In addition, the same concentration of analyte was spiked into spent media in a deep well plate followed by an addition of biosensor in, a) 5X LB-KAN+AMP (Naringenin), b) 5X CSM-URA or c) 5X LB-KAN (TAL). Droplet fluorescence of 100,000 droplets (for a total of 3 times, resulting in the average fluorescence of 300,000 droplets with the standard deviation between events shown), and plate fluorescence was collected by using a TECAN plate reader with the same 488nm excitation and 510nm emission settings as used with the droplet system, collection fluorescence measurements from biological triplicates for each concentration. Naringenin plate reader fluorescence has error bars too small to see from the data points. Linear regression was performed on each graph in order to determine the relationship between droplet and plate reader fluorescence (Naringenin r^2 = 0.9717, L-DOPA r^2 = 0.9411, TAL r^2 = 0.9500).

Attempts to coculture the TAL biosensor in a similar manner with *S. cerevisiae* resulted in almost no biosensor cell growth or response to varying TAL concentrations (Figure 3-6A). This restriction was remedied by the addition of a sodium phosphate buffer at pH 7.2 to the lysogeny broth with 50μg/mL kanamycin (LB-Kan) selective media (Figure 3-3A*i* and Figure 3-6B). With the system functioning, it was necessary to confirm the biosensor's response at low TAL concentrations to match with the production titers of *S. cerevisiae* compared with *Y. lipolytica.* Under this buffered condition, a linear low-range standard curve was observed for individual isolates (Figure 3-6B). These results demonstrate the need for a pico-injection-based approach that controls the time and restricts the duration for biosensor coculture.

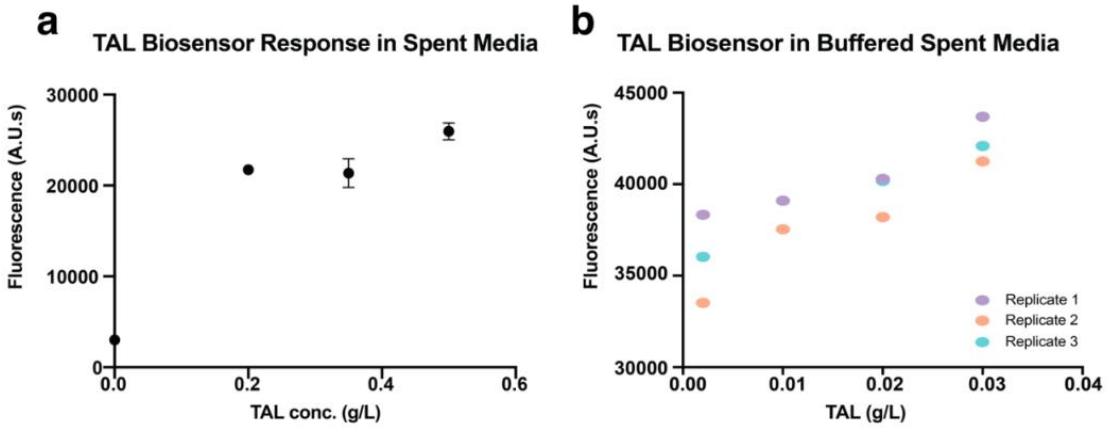

Figure 3-6: *TAL biosensor response in S. cerevisiae spent media cultures.* a) Initial testing of biosensor response to TAL standards in spent *S. cerevisiae* YSC-L media resulted in no distinguishable signal at varying TAL concentrations. b) Here, biological triplicates were taken to measure the fluorescence at lower ranges of TAL concentration, resulting in three distinct curves for each replicate-still showing resolution between TAL concentrations. Given that libraries are sorted based on an individual measurement, this variation was not a concern as we rely on linearity of response across a singular experiment (which is conserved in each replicate).

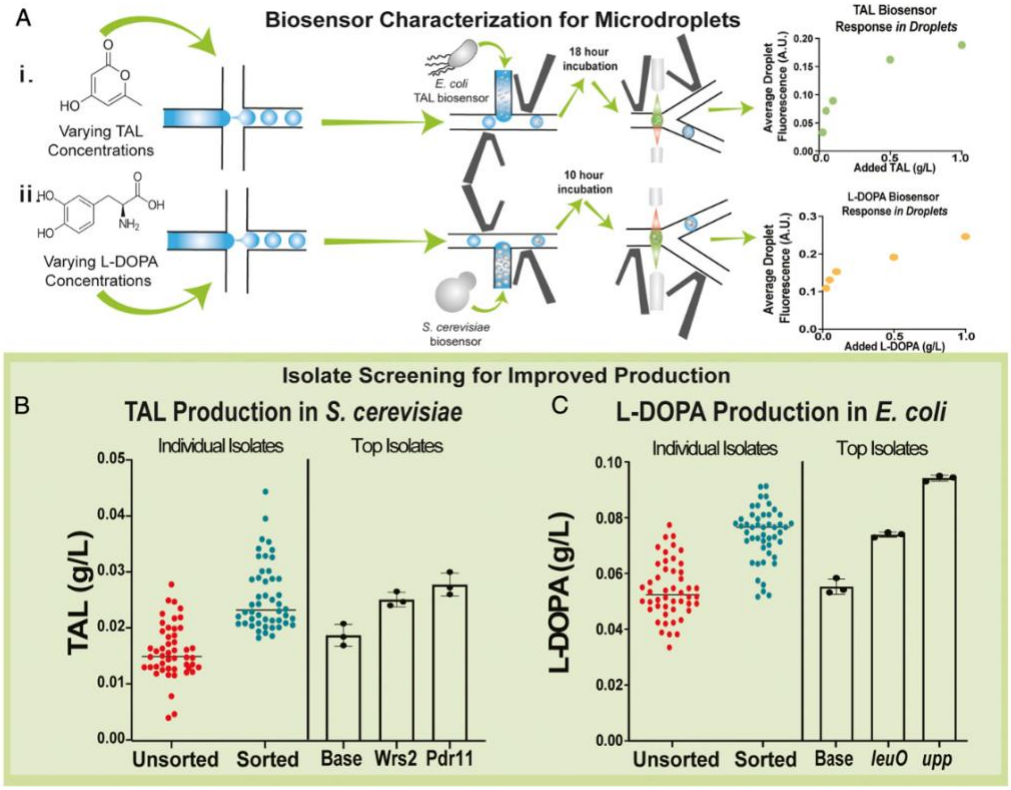

Figure 3-7: Improved production of TAL and L-DOPA by *S. cerevisiae* and *E. coli*, respectively. (A) Biosensor characterization for both TAL and L-DOPA sensing in droplets. (*i*) Here, *S. cerevisiae* was cultured in ysc-L for 72h to make spent media. Varying amounts of TAL standard were added to the spent media to generate emulsions of different TAL concentrations. *E. coli* TAL biosensors were pico-injected into these emulsions at an OD_{600} of 5.0 (final in-droplet OD_{600} of 1.0) in concentrated LB (5x LB-Kan selective media) to the spent media TAL standard emulsions along with a sodium phosphate buffer at pH 7.2. (*ii*) Here, BL21DE3 spent LB media was used with varying concentrations of L-DOPA. Biosensor response at concentrations flanking that were already made by the base strain was linear. These emulsions were run through the sorting chip to collect fluorescence measurements of 100,000 droplets for each concentration of analyte on three separate occasions (totaling 300,000 droplets). Weighted averages of each concentration are plotted along with their respective SDs (Calculated separately from the weighted average). Note: Triplicate data are shown, but it is difficult to distinguish as the fluorescence measurements were very similar. (*B*) Fermentations of individual isolates from an unsorted

control and sorted pool for TAL producing *S. cerevisiae* show population enrichment. The most frequently occurring sgRNAs identified by this screen were evaluated alongside controls to demonstrate enhanced production. Error bars represent SD. (C) Fermentations of individual isolates from an unsorted control and sorted pool for L-DOPA producing *E. coli* likewise shows population enrichment. Individual isolates identified via TAIL PCR were evaluated alongside the parental strain to confirm the effects of these perturbations. Error bars represent SD.

Finally, we sought to evaluate L-DOPA production from *E. coli* using an *S. cerevisiae* strain that can produce fluorescent betaxanthins through an L-DOPA intermediate that reacts to form betalamic acid via MjDOD. Such a coupling of cell types was more recently demonstrated using hydrogel-based cell encapsulation efforts[97], thus demonstrating the facile secretion and uptake of L-DOPA (Figure 3-A, *ii*). While coculture was possible in a hydrogel system[97], microdroplet environments as well as function dynamics and product oxidation in *E. coli* spent media (Figure 3-7A, *ii*) warrant a pico-injection type system.

3.3.2 Improving TAL and Naringenin Production in *Y. lipolytica*

As a first example, we chose to pair producer and biosensor cells that were growth compatible, thus enabling a simple coencapsulation demonstration (Figure 3-3). To do so, we selected individual strains of *Y. lipolytica* producing ~0.9 to 1.0g/L TAL and ~45 to 50 mg/L naringenin as starting points for diversity generation using a piggyBac random insertion transposon approach[98]. Initial tests indicated that these strains produced sufficient titer to induce previously reported *E. coli* biosensors with a fluorescence signal above background level while still leaving room to detect improved titers within the approximate linear range (Figure 3-3*A, i* and *ii*)[87,88].

Following transposon library generation, microdroplet emulsions were generated by utilizing an aqueous phase that contained both the producer cell (at a cell density

calculated for droplet loading to be roughly 0.1 cells/droplet based on the Poisson distributions[99]) and the biosensor cell (at a cell density calculated for droplet loading to be roughly 10 cells/droplet) together in rich yeast media, ultimately resulting in 1 in every 10 droplets containing a single secretor cell with roughly 10 biosensor cells. This condition was selected to allow for saturation of biosensor cells in each droplet while still avoiding co-encapsulation of more than one library member in most droplets.

Once created, we incubated these coencapsulated cells in droplets to allow for product accumulation and subsequent biosensor response before sorting via FADS to identify top producers (Figure 3-3A). To do so, FADS was performed by reinjecting the incubated emulsion into a microfluidic sorting chip (Sphere Fluidics) containing both a sorted and waste channel (Figure 3-1). For the first round of sorting, 1,800 individual droplets were collected from 4 x 10^5 droplets, followed by 200 collected in the final sort from a total of 1 x 10^5 droplets. Post sort, the emulsion was broken and plated to isolate individual strains. An analogous process (without sorting gates) was conducted with a control emulsion to establish a comparative baseline. Analysis of isolated clones demonstrated that library enrichment did indeed occur (Figure 3-3B and C).

Characterizing the top producing clones via inverse PCR identified six unique targets: *oil1, gsy1, pgm1/2,* p*IDH2, fbp1* and an intergenic region of chrD. Given that TAL and naringenin share precursor metabolites, we anticipated some overlap of identified gene targets and used both screens to identify putative targets to subsequently test. Among these six targets, *gsy1* and *pgm1/2* were found in both library enrichments whereas the *oil1* disruption was found in the TAL library enrichment and the remaining three targets were identified in the naringenin library. The two overlapping targets, as well as *fbp1,* had been previously identified as targets that improve TAL production in *S. cerevisiae* [100]. Likewise, *IDH2* has been associated with lipogenesis, another pathway that uses acetyl-CoA as a

precursor[101]. To confirm this collective set of targets and evaluate their potential to beneficially reroute metabolic flux, each target was reintroduced into a fresh TAL-producing strain, as baseline TAL production in this strain far exceeded naringenin baseline production. In doing so, we found that all of these targets functioned, except for the intergenic region of *chrD*, and could improve TAL production by up to 1.5-fold (Figure 3-3B). Confirmation of improved naringenin production was also performed for the *gsy1* target and achieved a 2.5-fold improvement, thus demonstrating transferability (Figure 3-8). Identification of these gene targets also demonstrates the power of coupling microdroplet sorting with biosensor cells.

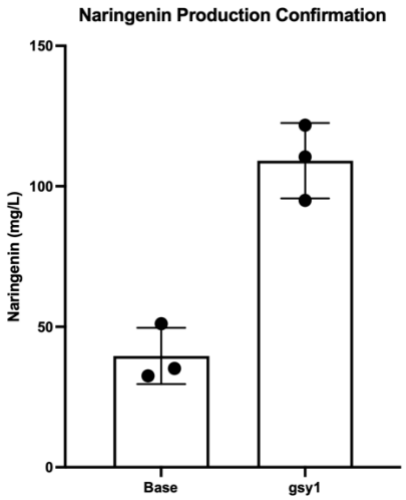

Figure 3-8: *Target confirmation for naringenin production.* Confirmation of the *gsy1* perturbation on naringenin production was performed by recreating the perturbation in a clean strain of *Y. lipolytica*. This isolate was incubated along with a control in 2X YPD for 48 hours, and the supernatant diluted 1:2 in 100% ethanol and run on the HPLC to obtain final titers. The *gsy1* target showed over a 2.5-fold improvement when compared to the base strain.

3.3.3 Optimizing a Pico-injection Workflow for Method Generalizability

As was demonstrated in the biosensor characterization, not all producer and biosensor cell systems will be amendable to coculture. In order further generalize the

microfluidic sorting approach for any producing-sensing cell combination (as well as to offer temporal spacing between production phase and detection to maximize biosensor operating range), we established a pico-injection workflow to introduce biosensor cells into microdroplets containing producing cells. To do so, we utilized a microfabricated chip specifically designed here to be compatible with the Sphere Fluidics system. Fundamentally, pico-injection utilizes a low voltage to merge an aqueous channel containing the biosensor in its preferred media with droplets containing individual library members as they pass through the site of pico-injection (Figures 3-1 and 3-7A). The key microfluidics of the Pico-Mix chip comprise a droplet feeding channel, and droplet mixing channel, together with a pair of electrodes positioned diagonally across the pico-injection nozzle (Figure 3-9). After a droplet enters the constrain channel, the droplet is elongated into a slug shape (Figure 3-9B, *i*), which ensures sufficient contact with the front interface of the pico-injection liquid when it passes the pico-injection nozzle (Figure 3-9B, *ii*). The electric field between the two electrodes facilitates a reduction in the integrity of the interface between the droplet and the pico-injection liquid. The continuous aqueous liquid is thus injected into the droplet via the opening between the droplet and the pico-injection liquid (Figure 3-9B, *iii*). A larger droplet forms due to the shear force from the continuous oil flow in the droplet constrain channel (Figure 3-9B, *iv*). By carefully tuning the flow rates of input droplets and pico-injection liquid, pico-injection can be accomplished at over 300Hz with the output droplet volume 40% larger. The coefficient of variation (CV) of droplet diameters is less than 2% before and after pico-injection. For example, the CVs of input and output droplets' diameters are 1.16 and 1.86%, respectively, as shown (Figure 3-9C). In this setup, pico-injection of the biosensor cell toward the end of producer cell fermentation allows for quick incubation, detection, and FADS using an otherwise growth-incompatible biosensor.

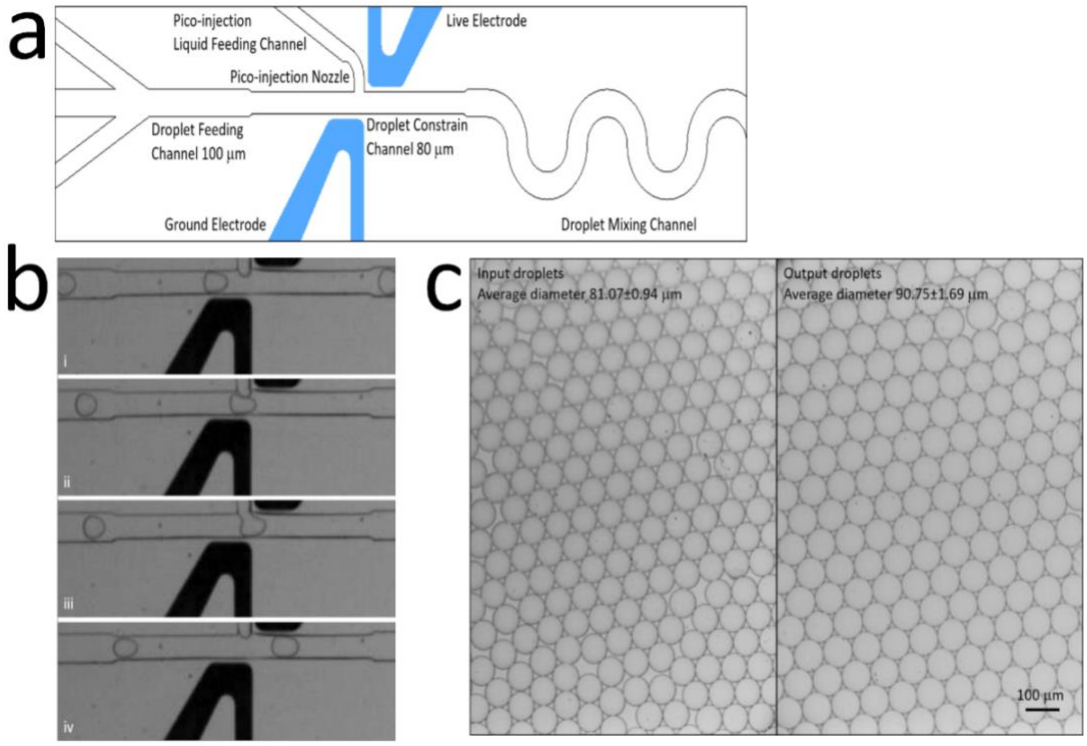

Figure 3-9: *Pico-injection chip design and performance.* a) The pico-injection chip is comprised of a droplet feeding channel that leads the droplets to the pico-injection nozzle flanked by two electrodes. The liquid feeding channel is where the biosensor cell mixture is fed in and merges with the in-coming droplets upon application of a low voltage. After pico-injection, the droplets are then collected via the droplet mixing channel and an outlet tube. b) In panel i, a droplet approaches the pico-injection site, in *ii*, it passes underneath the aqueous phase, and panel *iii*, pico-injection of individual droplets can be visualized. c) Change in droplet size post-pico-injection can be visualized and quantified, as seen here, with the original droplets having an average diameter of 81.07µm and the droplets post-pico-injection having an average diameter of 90.75µm.

To fully characterize the response of these biosensors when pico-injected into droplets, varying concentrations of each analyte were dissolved in spent media and encapsulated in separate emulsions of increasing concentrations (Figure 3-7A*i*, and *ii*). These droplets were then pico-injected with the respective biosensor. After incubation of the biosensor for 10 (L-DOPA Biosensor) or 18 (TAL Biosensor) h, the emulsions were

passed through the detector of the microfluidic sorting module to measure fluorescence levels of these droplets in a manner that would simulate sorting conditions (Figure 3-7A*i*,and *ii*). A similar analyte dose-response was exhibited by these biosensors in the droplets and plate reader experiments (Figure 3-5A-C), thus confirming that a pico-injection biosensor-based method is suitable for sorting libraries.

3.3.4 Improving TAL production in *S. cerevisiae*

Our first demonstration of a biosensor using pico-injection utilized *S. cerevisiae* TAL-producing cells along with an *E. coli* biosensor. As previously mentioned, these two cell types are coculture incompatible under most ordinary growth conditions without vastly altering screening conditions that may inadvertently select for a target that is irrelevant in ordinary conditions. Through optimization, it was determined that seeding the biosensor at high OD_{600} and in pH-buffered media was required to obtain a signal in spent *S. cerevisiae* media (Figure 3-6A & B). Thus, pico-injection of the biosensor cell via the Pico-Mix chip allowed for temporal and environmental control to establish a proper fluorescent signal with which to sort.

To utilize this TAL biosensor in pico-injection format, we established an *S. cerevisiae* library of varied TAL production through the use of an Mxi1 pooled CRISPR Guide RNA approach, containing just under 4,000 members[102]. This library offers graded expression levels for every basal metabolic enzyme and can thus be used to identify individual key rates, as expressions are altered for one enzyme at a time[102]. As a result, such an approach can enable modest changes to metabolism for each target identified.

After encapsulation of this library at a frequency of 0.1 cells/droplet, emulsions were incubated at 30°C for 3 days to allow for TAL production and then pico-injected with the *E. coli* biosensor followed by FADS utilizing the same final approach applied to the *Y.*

lipolytica library (Figure 3-7A). To improve enrichment rates, reduce noise, and account for user error during sorting, this library was sorted three times, the first resulting in 2,000 droplets collected from 8 x 10^5 sorted droplets. This library was then grown and re-encapsulated, followed by sorting 6 x 10^5 droplets, collecting 4,000 total droplets. These sorted droplets were then directly reinjected into the Pico-Sort (Sphere Fluidics) chip for further enrichment, collecting 200 droplets in final rounds. Strains isolated from this final sort displayed a statistically significant enrichment over members selected at random from the unsorted control pool (significance determined via two-tailed *t* test with Welch's correction, *P* value < 0.001) (Figure 3-7B).

All isolates with TAL production above the median of the sorted population were characterized via Sanger sequencing to identify the causative sgRNA. The stringent selection conditions used in this sort gave rise to many repeated sgRNA targets, including one targeting *pdr11* at the +76 base pair location and a second targeting *wrs1* at the -400 bp location, both indicative of medium-level knockdowns of expression[76,102]. These targets have not been previously implicated in TAL production in *S. cerevisiae.* Pdr11p is a multidrug transporter, thus implicating a potential role in the transport of TAL. Wrs1p is a tryptophan transfer ribonucleic acid (tRNA) synthetase, and its down-regulation likewise improves TAL production. The sgRNAs regulating *pdr11* and *wrs2* improved TAL production by 1.3 and 1.5 times respectively, when reconstructed into a clean strain and compared to a control vector with a scrambled guide (Figure 3-7B). To compare the magnitude and frequency of improvement obtained by this approach, we compared this approach to a random selection from the overall library. To do so, 384 members of the library were selected via 96-well deep plate screening, and these two droplet-identified targets outperformed the partial library screening by deep well plates, thus indicating that

the throughput of the droplet-based system was necessary in order to identify improved isolates with any reliability (Figure 3-10A & B)

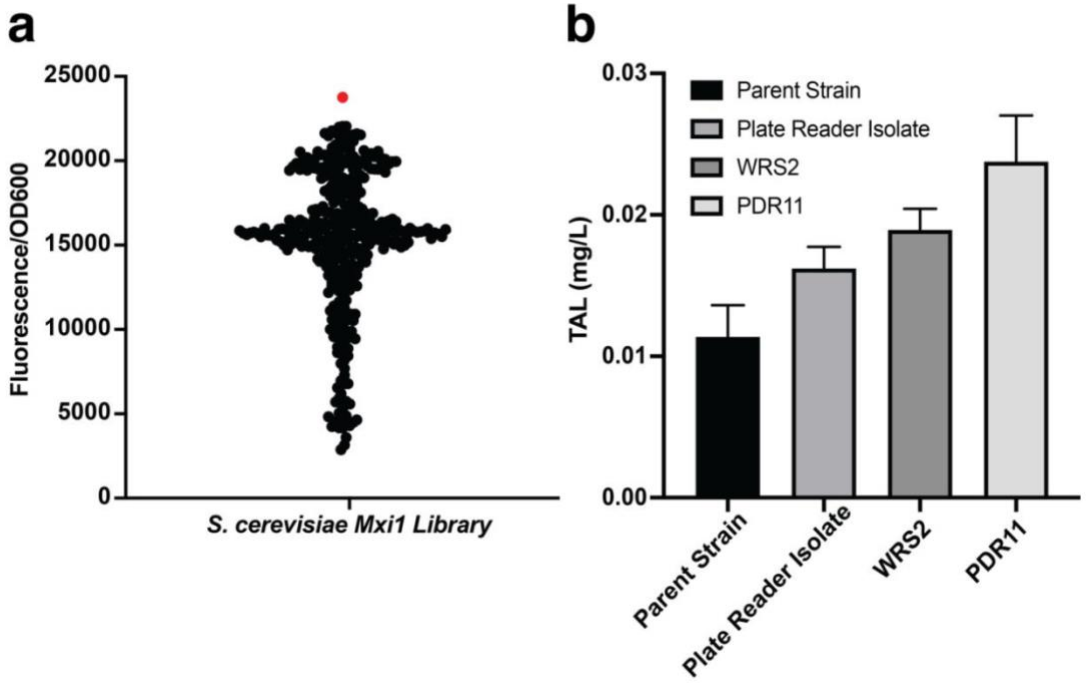

Figure 3-10: *Plate reader-based assay for identification of improved TAL producing isolates of S. cerevisiae in comparison to droplet-based methods.* a) Fluorescence of 384 isolates using *E. coli* TAL biosensor to identify improved production. Highlighted isolate was chosen for re-transformation and fermentation. b) Retransformation of highest performing isolate from deep well screen shows modest improvement in production compared to those identified via droplet screening. This demonstrates that a more comprehensive library coverage was necessary in order to identify best improvements to TAL production, indicating either increased screening via plate-based assays which is neither time nor cost effective, or more efficiently screening via microfluidics-based methods.

To demonstrate the scalable throughput capabilities of the droplet system, an EMS mutagenesis library of the *g2ps1*-expressing *S. cerevisiae* was created (Figure 3-11A, *i*). This library contained roughly 10^6 distinct members (calculated immediately after treatment via serial dilution plating). Encapsulation was carried out as described for the other library sizes, seeding at a rate of 0.1cells/droplet, and was evaluated at 1x coverage,

resulting in 10^7 droplets in total being evaluated on the microfluidic system (Figure 3011A, *ii*). After allowing the biosensor signal to mature for 18h, the emulsion was sorted by reinjecting all 10 million droplets (covering roughly a total of 1 million library isolates) and collecting the top 1% of the library based on fluorescence (Figure 3011A, *iv*, and B). This sorting process was completed in just over 6h, injecting the emulsion at a rate of 75μL/h. The unsorted control and sorted sample emulsions were broken and resuspended in 2X yeast extract peptone dextrose (YPD). After allowing the library to grow overnight, serial dilutions of sorted and unsorted samples were plated on nonselective media for single colony isolation. Random isolates were picked from the sorted and unsorted populations to obtain a total of 48 each for characterization on a singular 96 deep-well plate for 72h prior to high-performance liquid chromatography (HPLC) analysis (Figure 3-11C). Once again, these results demonstrated substantial enrichment of the selected pool over randomly selected members of the population (Figure 3-11C). These results showcase not only the generalizable and plug-and-play nature of this approach, but also demonstrates a throughput in the millions, easily exceeding the capacity of individual well-plate analyses.

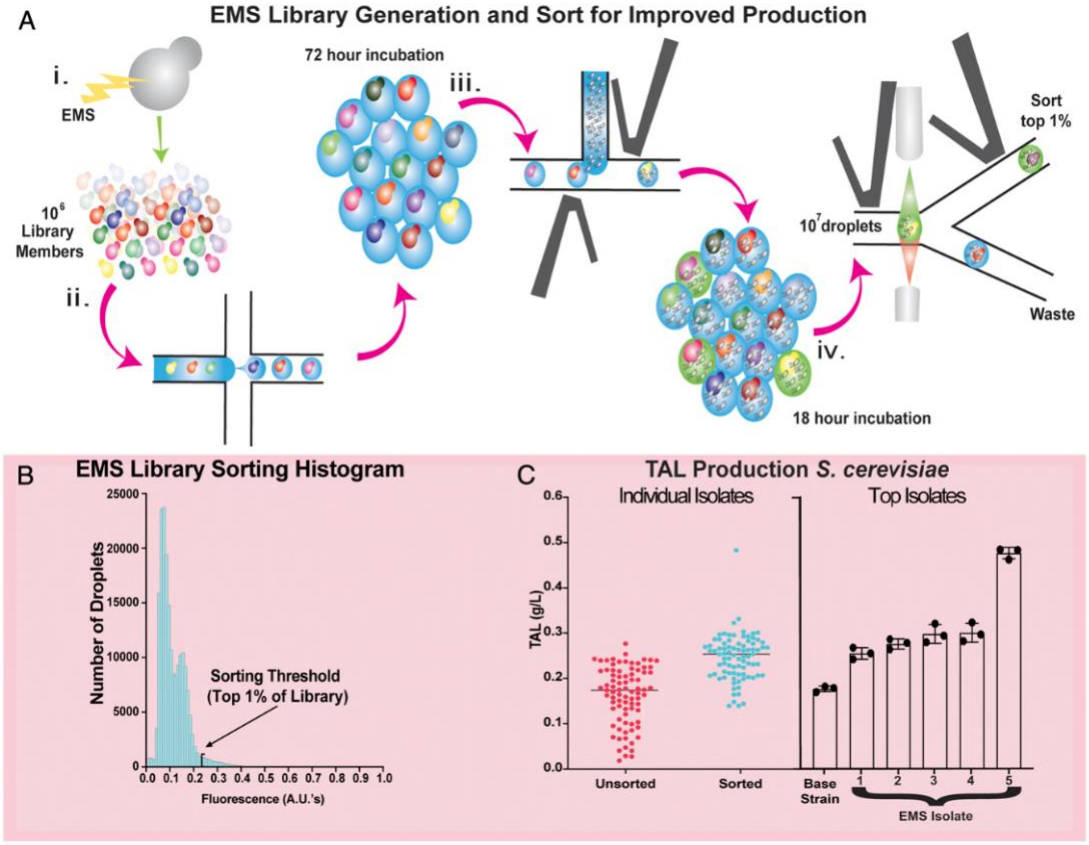

Figure 3-11: *Improved production of TAL in S. cerevisiae using an EMS library.* (A, *i*) EMS library synthesis resulted in a library with roughly 1 million members. (*ii*) Emulsion generation of the EMS library at a loading rate of an estimated 0.1 cells per droplet followed by incubation in droplets. (*iii*) Pico-injection of TAL *E. coli* biosensor followed by overnight incubation to allow for biosensor response. (*iv*) Sorting of the 10^7 droplet library based on fluorescence signal. (B) Histogram of FADS identifying sorting threshold at a value of 0.25. The two separate peaks indicate empty droplets (centered at 0.1) and droplets containing TAL-producing cells (centered at 0.15). (C, left) TAL titers via HPLC of 48 randomly selected isolates from the unsorted (not exposed to sorting gate) and sorted (Exposed to sorting threshold) populations. (Right) Performance of the top five EMS isolates identified from the sort compared to the parental strain titer. All sorts/fermentations performed in YPD, error bars represent SD from the mean.

3.3.5 Improving L-DOPA production in *E. coli*

As a final test demonstrating the generalizability of this pico-injection-based approach, we evaluated an opposite arrangement whereby a yeast cell was used as a biosensor to sort a library of *E. coli*-producing cells. In this case, we established a transposon library in the background of *E. coli* eBLTD1 that was engineered to overproduce L-DOPA (1-3,4-dihydroxyphenylalanine), a precursor to the neurotransmitter dopamine[103-105]. In this example, we sought to transduce the production of L-DOPA into a fluorescent signal via the synthesis of betaxanthins by *S. cerevisiae*. This yeast biosensor responded well in spent LB media to L-DOPA standards after a 10 h incubation (Figure 3-7A), and so we sought to perform a rapid production and pico-injection sorting scheme owing to the fast-doubling time of *E. coli* and to avoid oxidation of the L-DOPA product. To do so, fermentation time was limited to 18 h after encapsulation, followed by pico-injection of the biosensor and another 10 h incubation before sorting. In total, 118 high-fluorescence events were collected from sorting 8 x 10^5 droplets, and a final re-encapsulation and sorting step resulted in the collection of 40 high-fluorescence events. Once again, colonies isolated from this approach were compared with an unsorted pool to demonstrate a nearly 1.5-fold library-based enrichment (Figure 3-7C). Beneficial transposon insertion sites were identified via thermal asymmetric interlaced (TAIL) PCR[106], and among the loci isolated, we identified the targets of the *leuO* and uracil phosphoribosyl transferase (*upp*) as both resulting in at least 1.5-fold improved titer over the base strain (Figure 3-7C). The gene, *leuO*, encodes foor a transcription factor responsible for regulating leucine biosynthesis and has been identified previously for its role in stress response and pathogenesis expression[107], and thus its perturbation leading to

improved L-DOPA production implicates its role in other amino acid pathway regulation. Upp is responsible for catalyzing the conversion of uracil to UMP, and disruption of this gene being linked with higher L-DOPA production indicates a further interconnectedness between other amino acid production and the shikimate pathway in *E. coli.*

3.4 Discussion

Collectively, the demonstrations detailed in this chapter highlight a generalizable approach for using off-the-shelf biosensors in droplets via pico-injection or co-encapsulation to detect and sort for improved small molecule production/secretion. We showcase successful sorting across three different products in three different producer organisms (*Y. lipolytica, S. cerevisiae* and *E. coli*) through the use of both *E. coli* and *S. cerevisiae* as whole-cell biosensors. In all cases, we were able to sort through libraries that were commensurate with the size and scope of genome-wide and directed evolution experiments, ultimately reaching a droplet throughput of 10^7, thus successfully sorting a million-member library in a quarter of a day. In every case, microfluidic sorts resulted in gene targets that improved our phenotype of interest. We find that while co-encapsulation can be used in select cases, pico-injection allowed for more flexible control over biosensor timing, thus expanding the chemical palette accessible for small molecule library sorting in droplets. As a result, microfluidic sorting via externally introduced biosensors either by co-encapsulation or pico-injection represents a relief to the bottleneck in the "Test" step of the design-build-test cycle of metabolic engineering and synthetic biology.

3.5 Materials and Methods

PiggyBac Library Generation in *Y. lipolytica*

The piggy Bac plasmid (pMCS-[18dTEF-hyPBase]-[tr-5'ITR-URA3-3'ITR)-p1] was transformed into TAL and naringenin-producing PO1j-based strains. All colonies resulting from plasmid transformation were scraped into liquid SD-Leu media, and the resulting cultures were incubated for 4 d in a shaking incubator at 28°C/225rpm to allow for high level transposase expression and transposition. The resulting cultures were diluted into SD-Leu-Trp to select for transposon excision and incubated for 3-4d (to saturation) and then were diluted and outgrown again to saturation in nonselective 2X YPD two times to lose the replicative plasmid.

Mxi1 Library Generation for TAL Production in S. cerevisiae.

The BY4741 laboratory strain of S. cerevisiae was used as the base strain, transforming the Type III polyketide synthase, G2PS1 from *Gerbera hybrida*, into the His locus. This strain was then red-fluorescent protein (RFP)-tagged with m-cherry at the His locus. The final RFP-tagged, TAL-producing strain was then used to transform the Mxi1 plasmid-based library via a Gietz transformation, as previously described[102]. Libraries were stocked from an overnight culture, and individual stocks were used in their entirety to inoculate a culture for droplet encapsulation.

L-DOPA E. coli Library Generation Using pJA1

The E. coli eBLTD1 was used as a base strain in this work. Electrocompetent cells transformed with 1.5 μg of the pJA1 plasmid[108] were grown in 1.2 mL of super optimal broth with catabolite repression (SOC) at 37 °C with 225 rpm orbital shaking for 1 h. The recovered cells were subsequently transferred to a 125-mL flask containing 50 mL of LB

supplemented with 20 μg/mL kanamycin, 50 μg/mL spectinomycin, and 50 μM isopropyl ß-D-1-thiogalactopyranoside (IPTG) for 2-h induction of Tn10 trans- posase expression at 37 °C. The resulting libraries yielded a library size of ~10^5 variants. The culture was then centrifuged and reduced to a volume of 10 mL, and 2 mL of 20% glycerol stock aliquots were created. These aliquots were used for inoculation of the pJA1 transposon library cultures.

E. coli Naringenin Biosensor Characterization

Transformation of the PKSII cassette into *Y. lipolytica* resulted in multiple random integration clones with varying levels of naringenin production. Isolates were picked and incubated in 2× YPD + 2 mM 4-coumaric. The E. coli naringenin biosensor was added to the cultures during passage ratio of seven E. coli cells to one naringenin-producing *Y. lipolytica* cell. These clones were fermented in 2 mL culture tubes for 48 h at 28 °C in a rotary drum. Supernatant was collected from each and split and prepared for HPLC analysis or FACS for GFPmut2 fluorescence. Side scatter, forward scatter, and GFP fluorescence were measured using a BD LSRII Fortessa flow cytometer (488 nm laser, GFP filter), and the resulting 241 data points were analyzed using FlowJo software (FlowJo, LLC/BD).

Characterization of the *E. coli* naringenin biosensor in droplets was carried out using *Y. lipolytica* spent media with 0, 20, 40, and 60 mg/L of naringenin (in dimethylsulfoxide [DMSO]), with the *E. coli* naringenin biosensor in each sample diluted to 2.2 × 10^7 cells/mL. Emulsions were generated using a 60 × 60 μm Sphere Fluidics Pico-Gen chip with the aqueous phase set to 1,000 μL/h and 2.5% Pico-Surf in Pico-Wave oil

phase set to 1,800 μL/h to generate droplets with a diameter of 95 μm. Emulsions were then incubated overnight (18 h) and run through the Sphere Fluidics Pico-Sort chip with the photomultiplier tube sensitivity(PMT) setting set to 0.8 and the peak detection limit set to 0.06, consistent for every concentration of naringenin. Fluorescence data were collected on each population of droplets.

E. coli TAL Biosensor Characterization.

TAL biosensor characterization for use with Y. lipolytica was carried out as described for the naringenin biosensor. For characterization with S. cerevisiae, spent media was made by incubating BY4741 in defined media for 72 h and filter sterilizing. This media was used to create TAL standard curves for testing the E. coli biosensor response to TAL under these conditions. Two hundred microliters of biosensor was added to 800 μL of spent media in 5× LB-KAN (with and without 0.5 mM sodium phosphate buffer at pH 7.2) at an OD600 of 5.0 (final of 1.0) in order to mimic the 25% droplet volume increase that occurs via pico-injection. Fluorescence measurements were taken after 18 h of incubation using a TECAN Infinite 200 PRO plate reader with excitation/emission set to 488/510 nm, respectively, and autogain.

Characterization of the E. coli TAL biosensor in droplets was carried out using S. cerevisiae spent media with 0, 0.1125, 0.25, 0.5, and 1.0 g/L of TAL. Emulsions were generated using a 60 × 60 μm Sphere Fluidics Pico-Gen chip with the aqueous phase set to 1,000 μL/h and 2.5% Pico-Surf in Pico-Wave oil phase set to 1,800 μL/h to generate

droplets with a diameter of 95 μm. Im- mediately after generation, the emulsions were pico-injected with the E. coli TAL biosensor at an optical density at 600nm (OD_{600}) of 5.0 in 5× LB-KAN + 0.5 M sodium phosphate buffer. Emulsions were then incubated overnight (18 h) and run through the Sphere Fluidics Pico-Sort chip with the PMT setting set to 0.8 and the peak detection limit set to 0.06, consistent for every concentration of TAL. Fluorescence data were collected on each population of droplets.

S. cerevisiae L-DOPA Biosensor Characterization

S. cerevisiae L-DOPA bio- sensor characterization was performed similarly to the *E. coli* TAL characterization described. Spent LB media was made by incubation of BL21DE3 cells in LB for 24 h. This media was filter sterilized and used to create L-DOPA standard curves. Two hundred microliters of the *S. cerevisiae* biosensor in 5× yeast synthetic complete media (YSC)-Ura was inoculated into the spent media to mimic the pico-injection conditions. The signal was allowed to mature for 16 h and then read using a TECAN Infinite 200 PRO plate reader with excitation/emission set to 470/510 nm, respectively, and autogain.

Characterization of the *S. cerevisiae* L-DOPA biosensor in droplets was carried out using *E. coli* spent media with 0, 0.1125, 0.25, 0.5, and 1.0 g/L of L-DOPA. Emulsions were generated using a 60 × 60 μm Sphere Fluidics Pico- Gen chip with the aqueous phase set to 1,000 μL/h and 2.5% Pico-Surf in Pico-Wave oil phase set to 1,800 μL/h to generate droplets with a diameter of 95 μm. Immediately after generation, the emulsions were pico-

injected with the *S. cerevisiae* L-DOPA biosensor at an OD of 5.0 in 5× CSM-U + 20% OptiPrep. Emulsions were then incubated overnight (10 h) and run through the Sphere Fluidics Pico-Sort chip with the PMT setting set to 0.8 and the peak detection limit set to 0.06, consistent for every concentration of L-DOPA. Fluorescence data on each population of droplets were collected and weighted means were plotted in order to identify linear responses.

Co-encapsulation of *Y. lipolytica* and *E. coli* Biosensor

A Picodroplet Single Cell Encapsulation System (Sphere Fluidics) was used with a Pico-Gen 60 × 60 biochip (Sphere Fluidics) to generate homogeneous emulsions comprised of droplets with approximate initial diameters of 90 to 100 μm. A 2.5% (wt/vol) Pico-Surf surfactant in Pico-Wave diluent (Sphere Fluidics) was used as the fluorinated oil phase for emulsion formation. The final aqueous phase cell suspension consisted of 1.6× YPD after the addition of 20% vol/vol OptiPrep Density Gradient Medium (Sigma) to 2× YPD to minimize potential for settling of yeast cells during encapsulation. Kanamycin (TAL) or ampicillin plus kanamycin (naringenin) were also included in the growth medium to select for biosensor plasmids. After co-encapsulation at a ratio of approximately seven *E. coli* cells for each yeast cell, droplet emulsions were incubated with gentle shaking at 28 °C for 48 h to allow for product to build up enough to trip the biosensor but not so much as to saturate the signal. These times were determined via encapsulation of a range of *Yarrowia* producer cells with the corresponding biosensor, resulting in linear responses of HPLC-determined titer versus the droplet system fluorescence. All fermentation times

were chosen to allow for optimal product build up and biosensor maturation in as little time as possible to account for potential droplet shrinkage (previously identified) that could destabilize the emulsion[32]. The TAL biosensor plasmid was a gift from the Cirino laboratory, while the naringenin biosensor plasmid was from the Church laboratory.

FADS Sorting of *Y. lipolytica* Libraries

A Single Cell Assay and Isolation platform (Sphere Fluidics) equipped with a 488 nm laser 244 and 525/50 emission filter (GFP) was used for droplet sorting. After GFP-based sorting, Pico-Break reagent (Sphere Fluidics) was used to release cells from the droplets for recovery in 2× YPD test tube cultures containing 200 μg/mL spectinomycin (to inhibit growth of E. coli biosensor cells). Individual sorting events took less than half a day to sort, including the setup and takedown of the equipment. PMT setup of the system was set to a gain of 0.8 for both PMT 1 (GFP channel/bandpass 525/50) and PMT 2 (large bandpass 650/150). Peak detection minimum was set at 0.06 and maximum at 1. The minimum width was set to 0.1 and maximum set to 1 in order to avoid sorting/identifying large droplets that had merged together and therefore had a much higher signal. This exclusion allowed for a more accurate count of droplets containing single library members running through the sorting system. Sorting gates were determined based on the population distribution of GFP signal, and gates were manually drawn on the scatter plot to only collect the top percentage of droplets. The top 0.5% of 640,000 droplets (estimated 64,000 cell-containing) based on GFP intensity were collected in the first TAL droplet sorting cycle, and then the top 0.2% of 250,000 (estimated 25,000 cell-containing) droplets were

collected in the second TAL droplet sorting cycle. The top 0.6% of 600,000 (estimated 60,000 cell-containing) droplets was collected in the first naringenin droplet sorting cycle, and the top 0.2% of 200,000 (estimated 20,000 cell-containing) droplets was collected in the second naringenin droplet sorting cycle. After the second cycle of sorting and recovery in 2× YPD, the resulting cultures were diluted and plated on 2× YPD + spectinomycin agar to isolate single *Y. lipolytica* colonies. A control unsorted emulsion was made by drawing a gate around all fluorescent events in the sorting software. This resulted in every droplet with a fluorescent signal being collected, creating an unsorted control for comparative analysis.

Sequencing of Naringenin and TAL Top Producers

Individual piggyBac insertional mutant colonies from unsorted control and droplet-sorted populations were grown in suspension in 2× YPD for 1 to 2 d, and then genomic DNA was isolated using Zymolyase lytic enzyme (Zymo Research) and a Wizard Ge- nomic DNA Purification Kit (Promega). Approximately 1 μg of genomic DNA was digested with HaeIII (New England Biolabs) in a 15 μL reaction for 4 to 6 h, and the enzyme was heat inactivated at 80 °C for 20 min. Ten microliters of a 2× T4 Ligase master mix was then added to 10 μL of HaeIII-digested genomic DNA, and the ligation mixture was incubated overnight (16 to 18 h) at 16 °C before heat inactivation of ligase at 65 °C for 10 min. The resulting ligation mixture was used directly as a template in inverse PCR #1 (less stringent) with an annealing temperature of 55 °C and a 2-min extension time. Inverse 245 PCR #1 was then used directly as the template in inverse PCR #2 (more

stringent) with an annealing temperature of 63 °C and a 1-min extension time. After PCR purification, the products of inverse PCR #2 were directly subcloned into a linearized pUC19 vector using the In- Fusion HD cloning kit. Colony PCR products or isolated plasmid DNA from In-Fusion kit *E. coli* transformants were then subjected to Sanger Sequencing using M13F and/or M13R primers.

Generation of *S. cerevisiae* EMS Library

S. cerevisiae EMS library generation was carried out as previously described with some changes[109,110]. The same parental strain that was used for the Mxi1 library was used to generate the EMS library. It was grown in an overnight culture in 2× YPD to an OD_{600} of ~8 to 10. The cells were pelleted and washed twice with 0.1 M sodium phosphate buffer (pH 7.0) followed by resuspension to an OD_{600} of 30 in the same buffer. Liquid EMS (Sigma) was added to a 2.86% final concentration (vol/vol) and incubated with a control in which 50 μL of phosphate buffer was substituted for EMS for 1 h in a 30 °C/225 rotations per minute shaking incubator. After incubation, both the EMS-treated and control cells were pelleted and washed twice in 5% wt/vol sodium thiosulfate to inactivate the EMS and then resuspended in 0.1 M phosphate buffer (pH 7.0). Percent lethality due to EMS was estimated by plating both the library and control cells on nonselective YPD media. Size was similarly estimated from live cell counts on YPD adjusted for dilution factor, sample volume, and total volume of the final EMS pool.

Encapsulation of *S. cerevisiae* Libraries

Encapsulation of the Mxi1 and EMS S. cerevisiae libraries follow the encapsulation of *Yarrowia* with these changes: An overnight culture of the library was started from a glycerol stock in a 125-mL flask containing 50 mL CSM-L (complete supplement mixture without leucine), 20% Raffinose and YNB (Mxi1 library) or 2× YPD (EMS library). These cells were then diluted to 3×10^7 cells per milliliter in 1× YSC-leucine + 20% OptiPrep (vol/vol). The resulting cell suspension and fluorinated oil Pico-Wave (Sphere Fluidics) with 2.5% (wt/vol) Pico-Surf surfactant (Sphere Fluidics) were then separately loaded into 1-mL syringe (BD) and 10-mL steel syringe (SGE) injected at flow rates of 1,000 μL/h and 1,800 μL/h, respectively, into a 60 × 60 droplet maker (Pico-Gen) chip (Sphere Fluidics). After en- capsulation, emulsions were incubated at 30 °C and 225 rpm for 72 h to allow for maximum product accumulation without putting too much stress on the droplets.

Pico-injection of *E. coli* Biosensor *for S. cerevisiae* Library

Pico-injection of the *E. coli* biosensor was carried out the same way for both *S. cerevisiae* library types. After incubation of the emulsion, an overnight culture of the TAL biosensor was spun down and resuspended in 5× LB-KAN + 20 mM potassium phosphate buffer to a final OD_{600} of 5.0 and loaded into a 1 mL syringe (BD). The same Single Cell Assay and Isolation platform (Sphere Fluidics) used for sorting was used for pico-injection, this time equipped with a Pico-Mix (Sphere Fluidics) microfluidic chip. The emulsion was loaded into a 1 mL syringe (BD) and was reinjected into the chip at 100 μL/h, while the *E. coli* was injected between 7 and 50 μL/h. A set voltage level of 0.15 was continuously applied throughout the injection, resulting in merging of the aqueous with the droplets as

they passed through. Droplets were collected after injection and incubated overnight at 37°C and 225 rpm to allow biosensor maturation.

FADS Sorting of *S. cerevisiae* Libraries

After pico-injection, sorting of these libraries followed a similar protocol to the *Yarrowia* sort, with the same PMT parameter setups and the following changes: For the first sort of the Mxi1 library, the top 2% of 400,000 (estimated 40,000 cell-containing) droplets based on GFP intensity were collected, the top 0.5% of 400,000 droplets for the second sort, and the top 0.5% of 4,000 (estimated 400 cell-containing) for the final sort. This ensured 10× coverage of the plasmid-based library. After sorting, emulsions were broken with 100 μL Pico-Break (Sphere Fluidics) and cultured in YSC-L + Amp (to kill the biosensor) for 3 d before glycerol stocking at −80 °C and plating on YSC-L + Amp multiple dilutions to identify isolates. For sorting the EMS library, the top 0.1% of 10,000,000 total droplets (top 1% of the library) was isolated to sort through at least 1 million individual library members. After sorting, the emulsion was broken with 100 μL Pico-Break (Sphere Fluidics) and cultured in 2× YPD + Amp (to kill the biosensor) overnight before glycerol stocking at −80 °C and serial diluting on nonselective YPD to plate for isolates.

***S. cerevisiae* Library Sorting via Deep Well Plate**

After library transformation, samples were plated via serial dilution (10^{-5}, 10^{-6}, 10^{-7}) to obtain enough single colonies to cover one fourth of the library size (384). Colonies were then picked at random and inoculated into YSC-L–defined media to main- tain the

library plasmid. These cultures were then passaged at a starting OD_{600} of 0.1 and allowed to grow in a deep-well shaking incubator for 72 h. Finally, the *E. coli* TAL biosensor was added in buffered 5× LB-Kan to each well in a 1:5 ratio (biosensor:yeast culture). After an overnight incubation, plates were scanned for fluorescence to identify any improved isolates. Isolates showing significant improvement over the mean were then tested for improved TAL production via retransformation and subsequent tube fermentation and HPLC analysis of the supernatant.

TAL Production and Sequencing

Forty-eight colonies from a sorted plate and 48 from an unsorted plate were inoculated in YSC-L (Mxi1 library) or YPD (EMS library, *Yarrowia* libraries) and grown overnight in a 96 deep-well plate. These cultures were glycerol stocked and stored at −80 °C and then diluted to an OD_{600} of 0.1 and grown for an additional 72 h before collection for HPLC analysis. After identification of top producers, those cultures were inoculated from their corresponding glycerol stocks, and the plasmids were isolated via yeast mini prep II (Zymo). These plasmids were sent for Sanger Sequencing to identify the sgRNA of interest. Seed cultures were initiated by inoculating 2 mL of CSM-L (Mxi1 library), 2× YPD (EMS library, *Yarrowia* TAL library), or 2× YPD + 2 mM p-coumaric acid (*Yarrowia* naringenin library) in a 14-mL tube and incubating overnight in a rotary drum in a 28 °C incubator. The OD_{600} of seed cultures was measured using a spectrophotometer at 600 nm, and then 2 mL production cultures of 2× YPD were started in 14-mL tubes at an initial OD_{600} of 0.1 by direct inoculation from the corresponding seed culture. The resulting production tube cultures were incubated for 48 h in a rotary drum in a 28 °C incubator. At

harvest, production culture was diluted 5× with distilled water (TAL) or 2× with 100% ethanol (naringenin) and then the resulting sample was vortexed and centrifuged to pellet cells. The supernatant was then filtered through a 0.22 μm filter and loaded into sealed HPLC vials. A previously established Ulti- Mate 3000 (Thermo Fisher Scientific) HPLC protocol (Markham et al 2018) was used along with a pure TAL (Sigma) or pure naringenin (Sigma) standard curve to determine secreted polyketide titer. Production titers were calculated based on a standard curve linear fit and automatically adjusted for 5× dilution (TAL) or 2× dilution (naringenin) using Chromeleon software (Thermo Fisher Scientific).

E. coli Library Encapsulation

For emulsion droplet generation, the starter library cultures were inoculated from a glycerol stock into a 125-mL flask containing 50 mL of LB supplemented with 20 μg/mL kanamycin and 50 μg/mL spectinomycin and incubated at 37 °C for 10 h. Cells were then resuspended to 3×10^5 cells per mL in LB medium supplemented with appropriate anti-biotics, 10 mM vitamin C, 100 mM pipes, and 1% glucose for optimized L-DOPA production and droplet encapsulation[97]. The resulting cell suspension and fluorinated oil Pico-Wave (Sphere Fluidics) with 2.5% (vol/vol) Pico-Surf surfactant (Sphere Fluidics) were then separately loaded into 1-mL syringe (BD) and 10-mL steel syringe (CETONI) injected at flow rates of 1,000 μL/h and 1,520 μL/h, respectively, into a 60 × 60 droplet maker chip (Sphere Fluidics). Approximately 2×10^6 droplets (90 to 92 μm diameter) were collected in a 15-mL conical tube. The collection tube was then incubated at 37 °C with

225 rpm orbital shaking for 12 h to allow for L-DOPA production in droplets prior to pico-injection of a yeast biosensor.

E. coli Library Pico-injection of *S. cerevisiae* Biosensor

For pico-injection, the starter biosensor culture of sBY08 were grown in SD-URA medium at 30 °C for 1 d[97]. Yeast cells were then resuspended to 1.5×10^8 cells per mL with 5× SD-URA + 20% OptiPrep (vol/vol) Density Gradient Medium (Sigma) and loaded into a 1-mL syringe. Emulsion droplets after 12-h fermentation and fluorinated oil Pico-Wave were loaded into 1-mL BD syringe and Cetoni steel 10-mL syringe, respectively. Subsequently, yeast biosensors were pico- injected into droplets via a pico-injector chip at flow rates of 1,000 μL/h for oil phase, 50 μL/h for aqueous yeast sensor, and 350 μL/h for emulsion droplet with a voltage set at 0.25 V. Pico-injected droplets were finally collected in a 15-mL conical tube and incubated at 30 °C for 19 h to allow for L-DOPA accumulation to be sensed by yeast biosensors. The fluorescence of droplets was confirmed using a fluorescence microscope prior to running FADS.

FADS of *E. coli* Library

For droplet FACS, pico-injected emulsions and fluorinated oil Pico-Wave were loaded into a 1-mL BD syringe and Cetoni steel 10-mL syringe, respectively. Sorting was conducted at flow rates of 2,500 μL/h for oil phase and 25 μL/h for emulsion with a voltage set at 0.3 V using a Single Cell Assay and Isolation platform (Sphere Fluidics) equipped with a 488 nm laser and 525/50 emission filter (GFP). One hundred and eighteen high-

fluorescence events out of 7.9×10^5 (estimated 79,000 cell-containing) processed droplets (around top 0.015% of population) were collected in the first droplet sorting cycle, and then the top 0.1% of 2×10^5 droplets (estimated 20,000 cell-containing) were collected in the second sorting cycle. The unsorted and sorted samples were individually collected in 15-mL conical tubes, and afterward 100 μL of Pico-Break (Sphere Fluidics) was added to break emulsions. The resulting unsorted and sorted pools were separately grown in 2 mL of LB media supplemented with 20 μg/mL kanamycin and 50 μg/mL spectinomycin as well as 10 μg/mL nystatin (an antifungal antibiotic to inhibit the growth of yeast biosensor) and incubated at 37 °C for 7 h for cell recovery. Finally, suspension cultures were diluted and plated onto Petri dish agar plates containing LB media supplemented with 30 μg/mL kanamycin, 60 μg/mL spectinomycin, and 15 μg/mL nystatin to allow isolation of colonies.

L-DOPA Production and Sequencing

To screen and assess the isolates' capacities for L-DOPA production, 48 colonies were randomly picked from Petri dishes for each of the unsorted library and sorted pool and transferred to a Nunc 96 DeepWell Polystyrene Plate (Thermo Fisher Scientific) containing 800 μL of LB medium (supplemented with 20 μg/mL kanamycin and 50 μg/mL spectinomycin) per well. After incubation at 30 °C, 1,000 rpm for 18 h, 8 μL of each stationary-phase suspension culture was transferred to 800 μL of fermentation medium (LB + 20 μg/mL kanamycin + 50 μg/mL spectinomycin + 10 mM vitamin C [to prevent L-DOPA oxidation] + 1% glucose [wt/vol]) and incubated at 30 °C, 1,000 rpm for 15 h in a 96-deepwell plate. Fermentation samples were collected and supernatant fractions from pelleted cells were filtered and analyzed by HPLC for L-DOPA quantification.

Based on the titer screening results, cells identified as exhibiting increased L-DOPA content were selected and cultured with biological triplicates. The starter cultures of such isolates along with the control base strain eBLTD1 were grown in LB with appropriate antibiotics and incubated overnight at 37 °C. Seeding cultures were then transferred to test tubes containing 3 mL of LB medium supplemented with antibiotics, 10 mM vitamin C, and 1% glucose (wt/vol) with an initial OD600 of 0.005 and incubated at 30 °C for 15 h. Samples were collected and supernatants from pelleted cells were filtered and analyzed by HPLC. To confirm the transposon insertion sites of such high L-DOPA producers, each genomic DNA prepared using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI) was used as the initial template, primers used matched those previously described (Liu et al 1995). The gene fragments of the TAIL3 reaction were gel extraction purified and sequence verified using the sequencing primer TAIL-sEq (5′-CATCGCCTT- CTATCGCCTTCTT-3′) via Sanger sequencing[104].

HPLC Analysis of L-DOPA Titers

Samples were filtered with 0.2-μm nylon syringe filters (Wheaton Science) prior to running HPLC. HPLC confirmation of L-DOPA production was performed using a Dionex UltiMate 3000 (Thermo Fisher Scientific) equipped with an Agilent Eclipse Plus C18 column (3.0 × 150 mm, 3.5 μm) with detection wavelength set to 280 nm. Column oven was held at 30 °C with 1% acetic acid (vol/vol) in water or acetonitrile as the mobile phase over the course of the 20-min sequence under the following conditions: 5 to 15% organic (vol/vol) for 5 min, 15 to 100% organic (vol/vol) for 8 min, 100% organic (vol/vol) for 2

min, and 100 to 5% organic for 2 min followed by 5% organic for 3 min. The constant flow rate was set at 0.8 mL/min. A standard curve was prepared using 98.0% purity L-DOPA (l-3,4-dihydroxyphenylalanine) from Sigma.

Chapter 4: Temporal Sorting of Microdroplets can Identify Productivity Difference of Itaconic Acid from libraries of *Yarrowia lipolytica*[1]

4.1 Chapter Summary

As mentioned in the previous chapter, microdroplet screening of microorganisms can improve the rate of strain selection and characterization within the canonical Design-Build-Test paradigm. However, a full analysis of the microdroplet environment and how well these conditions translate to culturing conditions and techniques is lacking in the field. Quantification of three different biosensor/analyte combinations at 12-hour timepoints reveal the potential for extended dose-response ranges as compared to traditional *in vitro* conditions. Using these dynamics, in this chapter I detail an application and analysis of microfluidic droplet screening utilizing whole-cell biosensors, ultimately identifying an altered productivity profile of itaconic acid in a *Yarrowia lipolytica*-based piggyBac transposon library.

As mentioned in the previous chapter, we discovered that some cell-based biosensors are compatible with producers in droplets. The work in this chapter utilizes the ability to co-encapsulate *Yarrowia lipolytica* with a corresponding biosensor to shorten in-droplet fermentation and achieve enrichment for improved production at four distinct timepoints. Specifically, I show with this work that the timepoint for microdroplet selection can influence the outcome of the selection and thus shift the identified strain productivity

[1] Adapted from: Emily K Bowman, Phuong T.N. Hoang, Angela R. G. Sierra, Karoline M. V. Nogueira and Hal S. Alper, 2023. *Temporal sorting of microdroplets can identify productivity differences of itaconic acid from libraries of Yarrowia lipolytica*. ***Submitted.*** EKB built the piggyBac library, performed all time point sorting, sort confirmations, flask fermentation analysis, target validation and wrote the manuscript.

in flask scale, with the converse true as swell. Differences in response indicate microdroplet assays require tailored development to more accurately sort for phenotypes that are scalable to larger incubation volumes. Likewise, these results further highlight that screening conditions are critical parameters for success in high-throughput applications.

4.2 INTRODUCTION

High throughput microbial engineering relies heavily upon the canonical Design-Build-Test cycle[86,111]. While recent advances in DNA design and synthesis have helped usher in this capacity for high-throughput biology, they also create a substantial bottleneck in the test step, thus requiring more creative solutions for high throughput screening of microbial strains[23,56,112,113]. Efforts to switch away from bulk liquid handling and long chromatography/separation dependent analysis techniques help improve the throughput of the test step. Certainly, established methods such as fluorescence activated cell sorting, microtiter plate cultivation, and agar-plate based screens represent alternatives. In contrast, the fluidics miniaturization and parallel culturing capability of microfluidic droplet systems help bypass many of the limitations in high-throughput metabolic engineering that are inherent even in these alternative approaches[86,40,114,115]. To this end, microdroplet systems uniquely enable a physical encapsulation of cells and their extracellular phenotypes such as small molecule secretion.

A remaining challenge for microdroplet screening is the capacity to transduce a chemical concentration into a rapidly detectable signal. Recent applications accomplish this task through the use of cell based biosensors[29,40,86,92,114,115], chemical sensors[116], growth-based assays[29], and even mass spectroscopy for label-free sorting[117]. However, the

true utility and adoption of high throughput microdroplet screening and selection is predicated on its capability to deliver strains that perform as expected outside of the microdroplet. Limited studies have attempted to draw parallels between the microdroplet and liquid phase production, especially as it relates to the parameter of time and production.

Overall assumptions regarding microdroplet sorting systems have led to difficulties in scaling identified strains[32,113,118]. Specifically, the small-volume environments of microdroplets often identify condition-specific, beneficial gene perturbations that may not perform similarly in flask or bioreactor conditions[113]. Moreover, iterative and prolonged microdroplet-based high throughput screening (as with any screen with an inherent growth step) inadvertently creates a selective environment for increased biomass—a phenotype that often negatively correlates with metabolic flux for small molecule production[119,120]. Compounded on these challenges, previous work has observed inconsistencies in strain performance upon scale-up, specifically comparing small-scale cultivation to larger scale bioreactors[120]. These challenges aside, previous applications of fluorescent based microfluidic sorting have typically utilized the same incubation time used in larger cultures, including prior efforts detailed in the previous chapter[14,40,114]. However, full characterization of the microdroplet environment is still on-going, as analysis of microdroplets is limited by access to analytical devices for immobilizing microdroplets for prolonged observation, as well as sensitive sensing and measuring techniques such as Mass Spectrometry Assisted Droplet Sorting (MADS)[117,118].

Despite many challenges, the microdroplet environment, while clearly very different from other culturing methods including bioreactors, flasks, tubes, and even deep well plates[32,118],

is well-suited for selections and thus necessitating studies on cellular dynamics. More specifically, limited studies have investigated the importance of time-based selections when using these pico-liter sized microdroplets as a means of selecting for larger-scale productivity trends. To address these limitations, the work in this chapter seeks to answer the question of whether the time-dependent selections in microdroplets will correspond to the fermentation profiles of strains outside of microdroplet conditions. Put another way, this work seeks to identify whether microdroplet selections can be used to select for shifts in cell productivity and thus achieve more exponential or more stationary phase production.

As an initial characterization, we studied the biosensor dynamics of three separate whole cell biosensors with known production strains. Once we identified a product with the greatest dynamic range in microdroplets, we investigated the influence of time-dependent selections (in this case, 24, 48, 72 and 96 hour) in microdroplets on the identification and performance of *Yarrowia lipolytica* strains for itaconic acid production using a transposon mutagenesis library. The work in this chapter demonstrates that sorts conducted earlier indeed result in identifying strains that show improved early productivity outside of the microdroplet. Likewise, the corollary is also proven that later-timepoint sorts identify strains better suited for production later in the culturing process. Finally, through in microdroplet characterization of production, we can demonstrate the correspondence of the pico-liter culturing environments to larger-scale cultivation techniques. Ultimately, this work identified strains that were enriched for distinct production phenotypes including increased exponential productivity (a disruption in *Ypgm1)* and higher final titer/growth independent production (as disruption in *Ygsy1*).

4.3 RESULTS

4.3.1 Characterization of Biosensor Dynamics in Microdroplets

To elucidate the dynamic range of co-encapsulated biosensors, three established *Yarrowia lipolytica* strains producing TAL[91], Naringenin[95], and Itaconic Acid[121] were co-encapsulated with their corresponding *E. coli* biosensors[114,122] (Figure 4-1a). Fluorescence measurements were performed on the sorting instrument (SphereFluidics) to collect histograms of microdroplet count and fluorescence (Figure 4-2) throughout an 84-hour time course. Response curves at each condition indicate a wide dynamic range from each biosensor (Figure 4-2). Interestingly, the itaconic acid strain co-encapsulation with the itaconic acid biosensor shows a longer range with prolonged detection, with increases in fluorescence still occurring up to the 84-hour timepoint (Figure 4-1b).

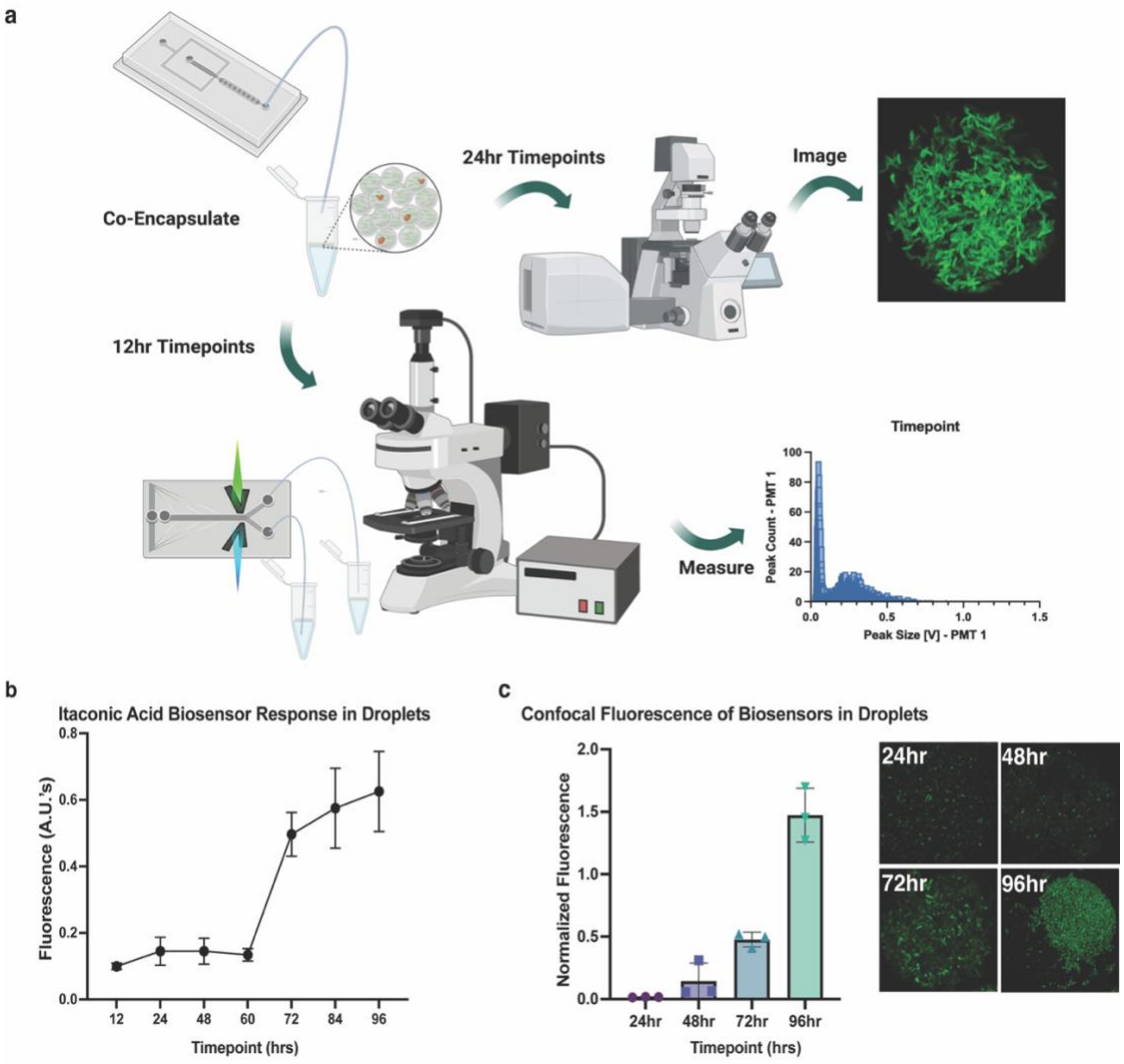

Figure 4-1: *Biosensor Dynamics in Microdroplets.* a) Workflow of emulsion generation and data collection for fluorescence at selected timepoints. Top performing isolates or wildtype were encapsulated with the same method as sorted isolates. Created with BioRender.com b) Biosensor response dynamics in droplets were measured over time. The itaconic acid producing parental strain of *Y. lipolytica* was encapsulated with the corresponding *E. coli* biosensor and fluorescence was measured on the microdroplet sorter in 12-hour increments. c) Parental strain fluorescence changes within microdroplets were captured with Nikon™ spinning disk confocal imaging. These measurements are normalized to the total area measured for fluorescence. To characterize the itaconic acid biosensor response in droplets more clearly, emulsions were imaged and fluorescence intensity at each timepoint was measured.

Based on these observations, we chose to also obtain confocal images of the microdroplets containing the *Y. lipolytica* itaconic acid production strain and the corresponding *E. coli* biosensor at 24-hour timepoints to provide an additional measurement biosensor fluorescent temporal response and homogeneity (Figure 4-1c). Throughout the timecourse, these images do suggest some biosensor growth within microdroplets. Therefore, it is possible that in these prolonged sensing conditions net sensor population growth can result in a dilution of analyte response. However, it is also clear from these images that both the total aggregate fluorescence and biosensor cell-specific fluorescence increases in a dose-responsive/time-dependent manner. Therefore, we chose to use itaconic acid production and sensing as a test case for time course-based sorting of *Yarrowia lipolytica* piggyBac transposon library.

4.3.2 Timepoint Sorting in Microdroplets to Enrich for Itaconic Acid Overproduction

A library of pathway-engineered *Y. lipolytica* containing piggyBac transposon insertions was used to screen for improved itaconic acid production. To achieve a detectable signal, an *E. coli* biosensor was co-encapsulated with these *Y. lipolytica* strains to achieve fluorescence-dependent readout of concentration. Prior work and efforts described above have demonstrated that *E. coli* and *Y. lipolytica* tolerate one another well in co-culture microdroplets[114], thus enabling both early and late sorting capacity with more accurate timepoints and less emulsion manipulation. Specifically, early timepoint sorting is achievable using co-encapsulation without concern for down-time when using approaches such as pico-injection. Likewise, extended emulsion incubations are enabled

without the consequence of disturbing microdroplet stability that can occur during pico-injections due to volume addition and voltage/temperature changes.

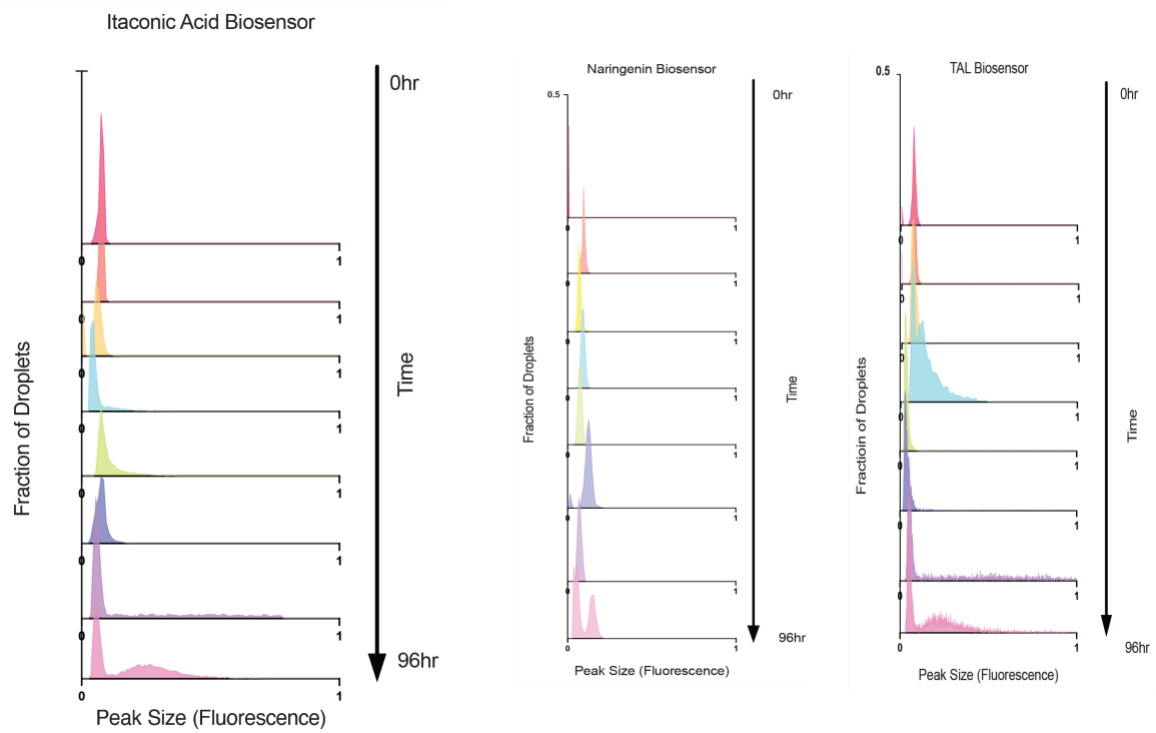

Figure 4-2: *Histograms of each biosensor at 12 hour timepoints.* Previously-characterized production strains of *Y. lipolytica* were co-encapsulated with corresponding *E. coli* biosensors for the production of itaconic acid, Triacetic acid lactone, and naringenin. Emulsions were subsequently incubated and run on the droplet sorting instrument (Sphere Fluidics) to collect fluorescence information from droplets at each timepoint. Droplet loading was 0.1 cell/droplet for *Y. lipolytica* and 10 cells/droplet for *E. coli* to reduce co-encapsulation of producers as modeled via the Poisson distribution. Therefore, only the top 10% of droplets represent those containing both the producer strain as well as the biosensor, resulting in a dose-based response.

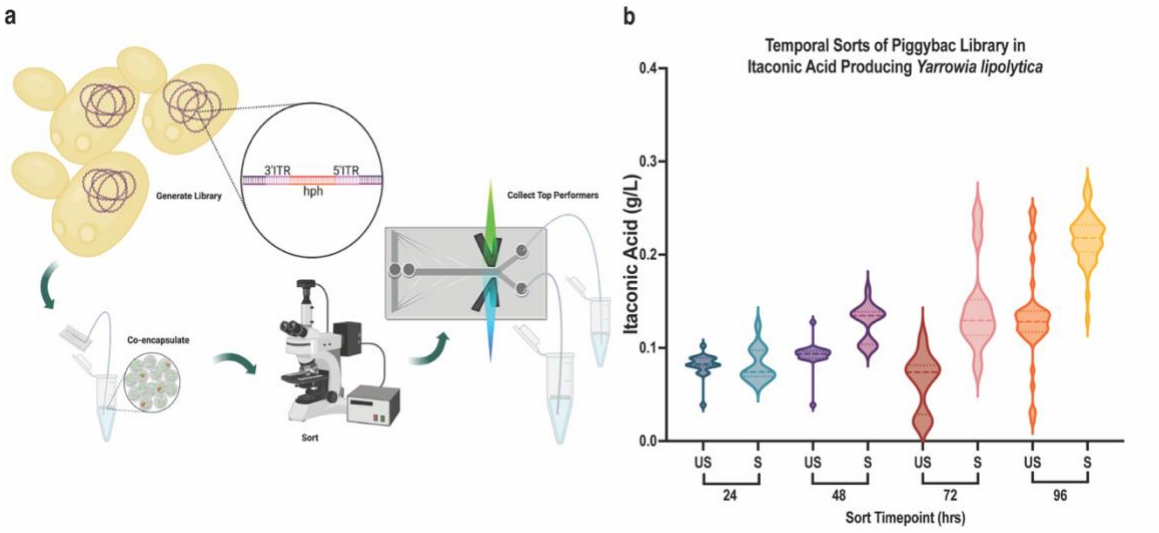

Figure 4-3: *Timecourse sorting of Y. lipolytica piggyBac library for itaconic acid production.* a) A piggybac transposon library was generated in an itaconic-acid producing strain of *Y. lipolytica* [121,123] and co-encapsulated with the itaconic acid *E. coli* biosensor at a ratio of 0.1:10 *E. coli: Y. lipolytica*). After timed incubation of either 24, 48, 72, or 96 hours, emulsions were then sorted using GFP signal. Created with BioRender.com. b) Post-sort, emulsions were broken and cultivated in a mixed population and plated for single colonies on YPD-agar plates. Random isolates from the sort and an unsorted control were picked from plates and grown in YPD on the tube-scale for their respective fermentation time that corresponded with the sorting timepoint. Supernatant was isolated and itaconic acid titer quantified via HPLC to determine if the sorts were successful.

Overall, microdroplet sorting was carried out as previously described[114] (Figure 4-3a) and isolated the top 1% of fluorescent microdroplets. In this experiment, the library sorting was conducted at four distinct timepoints (24, 48, 72 and 96 hours) to determine the impact of sorting time on productivity and final titer. These populations were characterized by first plating the enriched mixed culture after sorting and second selecting a representative set of 24 isolates from this subpopulation at random. The unsorted control was obtained via the same sorting method (including applied voltage), but with no gate applied followed by a similar plating and selection workflow. These isolated strains were

cultivated at the test-tube scale and supernatant samples were analyzed for itaconic acid quantification via liquid chromatography (Agilent Zorbax aq). In each of these cases, enriching the piggyBac library for improved itaconic acid production resulted in successful, statistically significant (Welch's correction *t* test) separation of populations at each timepoint compared with a control, unsorted pool isolated at that same timepoint (Figure 4-3b). Successful population enrichment from these sorts at all timepoints indicates that the dynamic range, sensitivity, and saturation properties of the biosensor were sufficient to achieve signal differentiation.

4.3.3 Flask-based Time Course Experiments Highlight Parallels with Sorting

After confirming successful sorts of each timepoint, all isolates demonstrating an itaconic acid production level above the mean from the respective sorted condition were selected for a flask-scale time course experiment. To do so, each isolate was grown in 250mL flasks containing 50mL YPD media with 2% glucose at 28°C with timepoints taken every 12 hours for a total of 96 hours. Collective production curves for each of the populations show clear trends with respect to production capability (Figure 4-4).

To further compare productivity trends, itaconic acid production from the flask time course is segmented and represented as a percent of total production at Early (48-36) Mid (72-48) and late (96-72) timepoints in the flask (Figure 4-5ai-iv) for strains that were isolated using early sort timepoints (24 and 48 hours in microdroplets), most of the productivity in the flask was likewise biased toward early timepoints (in this case, 36-48 hours) with strains peaking in production at around 60 hours (Figure 4-5ai-ii). For strains isolated at the late timepoints (72 and 96 hours), productivity was more gradual in the

culture with more itaconic acid produced late in the culture, especially at the 72-96 hour flask timepoints (Figure 4-5aiii-iv). It is worth noting that these late-sorting cells ultimately peak at a slightly higher final titer (Figure 4-4b). Most striking, however, is the clear difference in percent production profile between the strains isolated at 24- and 96-hour sorts (Figure 4-5ai, aiv).

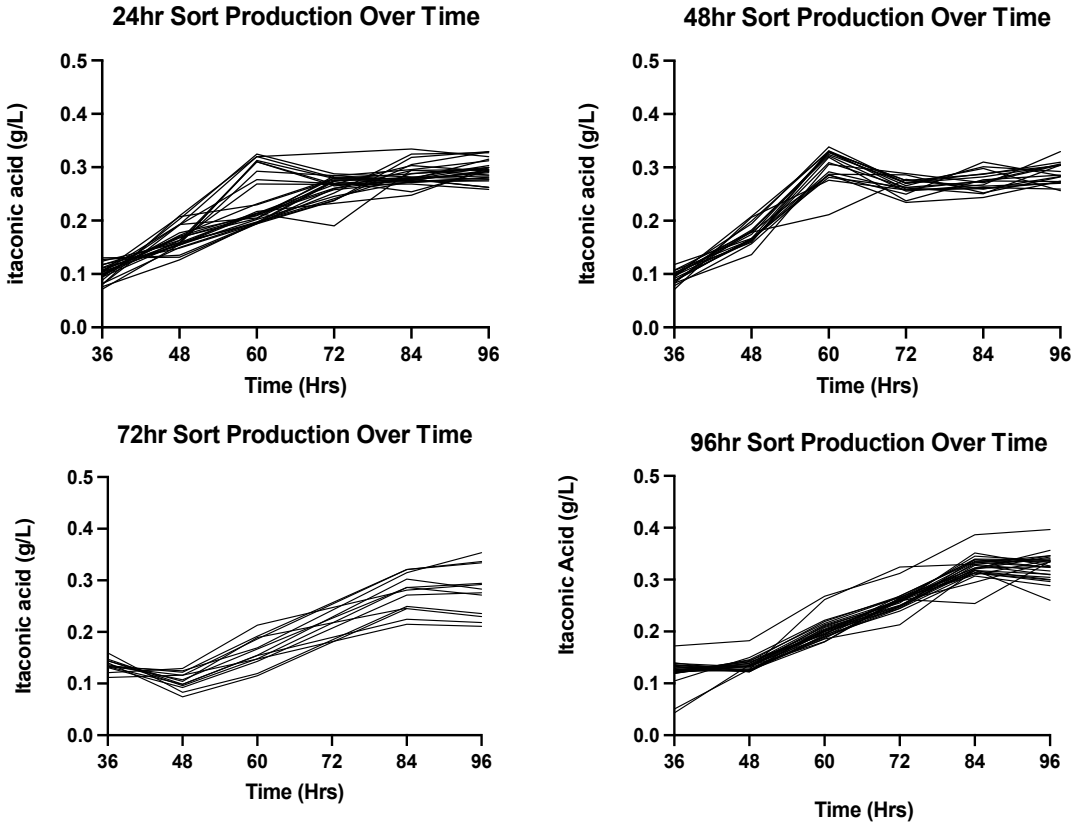

Figure 4-4: *Production Curves of Populations from Each Sort.* Top performing isolates from each sort were cultured for a total of 96 hours with itaconic acid production and OD sampled every 12 hours. a) Production curves of the 24- and 48-hour isolated population over time. Early sort timepoints (24 and 48) show a clear peak in production at the 60-hour timepoint followed by a plateau. b) Isolates sorted at late timepoints (72 and 96) show a slow increase in production that peaks at the 84-hour timepoint.

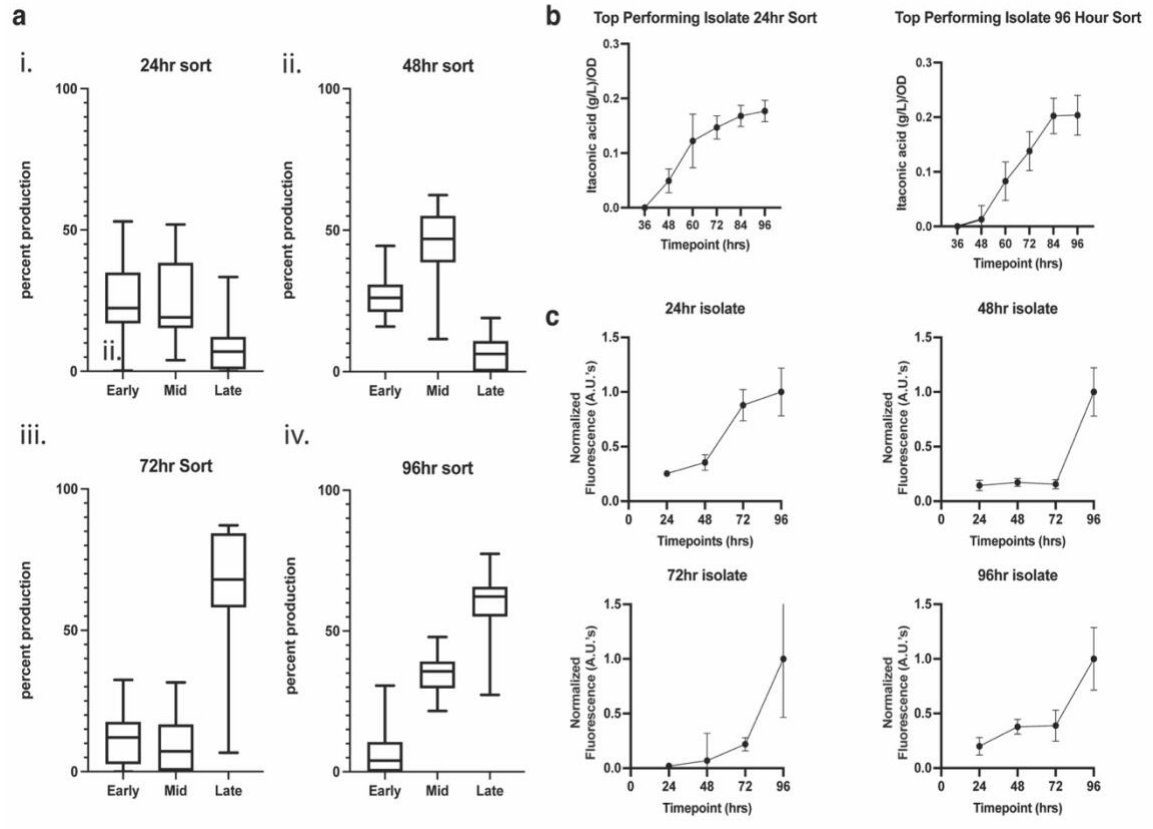

Figure 4-5: *Time course fermentations for production phenotype characterization.* Top performing isolates from each sort were cultured for a total of 96 hours with itaconic acid production and OD sampled every 12 hours. a) Percent production of itaconic acid at early (36-48hr) mid (72-48) and late (96-72) timepoints for 24 (i), 48 (ii), 72hr (iii) and 96 (iv) hour sorts. Isolates chosen for evaluation represent the top 24 identified from initial sort analysis (Figure 2b). Percent production at each segment of fermentation was calculated as the change in itaconic acid titer between the timepoints divided by the total final titer. b) Biological triplicate of defined media fermentations of top performing isolate from 24 and 96 hour timepoints show the same patterns of production as the rich media. In addition, significant production doesn't seem to begin until the 48-hour timepoint for both conditions. This indicates that rich media fermentations, while plagued by high background, still show a clear picture of what production looks like in a flask for these strains. c) The top performing isolate from each sorting condition was re-encapsulated with biosensor and fluorescence was collected via the droplet sorter. Fluorescence for each isolate was collected on separate days, therefore only comparisons can be made within each isolate, not between.

To determine if these phenotypes were linked to cell growth/biomass accumulation, itaconic acid titer was normalized to the OD for the 24 and 96-hour sort condition strains (Figure 4-5b). The production patterns obtained here are fairly like the non-normalized curves (Figure 4-4a and b) and once again highlight a higher productivity earlier in the fermentation for cells isolated in the 24-hour sort compared with the 96-hour sort. The normalized total production is also still higher for the strains isolated from the 96-hour sort. As a result, different productivity profiles and capacities are identified when distinct microdroplet selection conditions are imposed.

Top performing strains (Figure 4-3b) at each of the timepoints were preserved and isolated genomic DNA was used as a template in TAIL PCR[11,106,114,123,124] for transposon insertion site identification. It is interesting to note that the various selection conditions gave rise to both unique production phenotypes and distinct genetic targets. Specifically, an insertion in the *Ypgm1* gene responsible for phosphorylation in the glycolysis pathway was predominately isolated from the 24 and 48-hour best performing isolates. In contrast, the 72- and 96-hour best isolates both converged on the same insertion in the *Ygsy1* gene which controls glycogen storage. Interestingly, the 24-hour sort appears to exhibit a range of production phenotypes, the very late-stage production phenotype as well as the early-stage productivity (Figure 4-4a). It is thus not overly surprising that some lower-performing isolates from the 24-hour sort, but not the 48, show the same, *Ygsy1*, gene perturbation as the later sorts. Target convergence within each sort indicates the robustness of the screening system, and the 24-hour dual-phenotype isolation may also indicate the timepoint at which these phenotypes diverge to be isolated individually.

Interestingly, both of the piggyBac insertions identified here for itaconic acid production have also been identified for the production of naringenin and TAL in *Yarrowia lipolytica*[114]. These previous screens also utilized a similar biosensor-in-microdroplet approach to identify improved production of these small molecules. Re-identification of these gene targets increases the likelihood that their disruption could be beneficial for many small molecule production applications in this yeast.

4.3.4 Characterizing the In-Microdroplet Conditions for Top Performers to Understand Strain-Biosensor Dynamics

To better characterize the correspondence between flask and microdroplet scale, we re-encapsulated representative isolates from each of the major populations described above into *Y. lipolytica* mono-culture emulsions along with the *E. coli* biosensor. These mono-culture emulsions were incubated for a total of 96 hours once again and run on the FADS (Sphere Fluidics) at 24, 48, 72, and 96 hours (Figure 4-5c). Fluorescent measurements for isolates at each production stage were collected via the microdroplet sorting system, allowing for a direct comparison between fluorescence obtained and sorted in initial experiments to their best-performing timepoints. Patterns how similar responses to those seen in the wildtype experiment (Figure 4-1b), with the biosensor fluorescence increasing greatly at the final, 96-hour timepoint. However, this is likely due to the accumulation of product over time, and interestingly, the biggest differences between isolates can be seen in the fold-change of fluorescence between the earliest timepoint (24 hours), and the final timepoint (96 hours). In particular, the 24-hour isolate showed a 2-fold increase in fluorescence between the 24-hour timepoint and 96-hour, whereas the 96-

hour isolate shows an increase of almost 5-fold in fluorescence. This difference corresponds well with the production profiles that indicate an early plateau for the earliest-selected isolates (Figure 4-5a and 4-2).

4.4 Conclusion

Microdroplet sorting using cell-based biosensors has expedited the Design-Build-Test cycle for metabolic engineering and synthetic biology, especially for the case study of improved small molecule production[14,29,40,114]. By characterizing three distinct biosensor dynamics in microdroplets, we identified extended dose-response ranges to production. In particular, we recognized an opportunity to utilize an itaconic acid biosensor, with its increased response range in microdroplets, for production phenotype sorting. Utilizing this, we were able to showcase the flexibility of a microdroplet environment to uniquely allow for selections of strain productivity for itaconic acid producing *Yarrowia lipolytica.* This allowed us to connect in-microdroplet production to in-flask production and recapitulate those phenotypes from microdroplets to flasks.

In this chapter, we showcased distinct early versus late production phenotypes obtained from time course-based sorting and identified two insertions responsible for these different production changes. Applications of this time course-based system could be used to tailor high throughput screening for limited biosensor dynamic ranges, improving strain productivity, or identifying ideal sorting times for improving production using gene perturbation libraries.

4.5 Materials and Methods

Strain cultivation

All *Yarrowia lipolytica* production strains[86,114,121] were grown in either YPD containing 2% glucose (for microdroplet analysis and large flask time-course) or yeast synthetic defined (YSD) media (2% glucose, YNB (Difco 6.7g/L), CSM complete (Thomas Scientific)). Naringenin production strains also utilize p-coumaric acid at a final concentration of 2mM as previously described [95]. Flask cultures were grown at 28°C and shaken at 225rpm. Microdroplet encapsulated cultures were grown in a standing incubator at 28°C with oil-changes every 12-hour to maintain oxygenation of culture. *E. coli* itaconic acid biosensor[122] was pre-cultured before microdroplet encapsulation in LB-Kan (50ug/ml) at 37C for 18 hours. *E. coli* Naringenin and TAL biosensors were pre-cultured as previously described [114].

Cell Encapsulation in Microdroplets

Production strains used here were those developed previously for TAL[91], Naringenin [95] and Itaconic Acid production[121]. *Yarrowia lipolytica* production strains were encapsulated at a rate of 0.1 cells per microdroplet. Their corresponding biosensor cells were co-encapsulated at a rate of 10 cells per microdroplet. Encapsulation for the time courses was performed in two rounds, with timepoints taken from the same encapsulation at timepoint 0 as well as 12- and 24-hour intervals to achieve response curves.

Piggybac library generation in the itaconic acid producing strain of *Yarrowia lipolytica* took place as previously described[114]. *Yarrowia lipolytica* piggyBac library members were encapsulated in microdroplets at a rate of 0.1 cells/droplet. This was chosen based on *Poisson* distribution modelling to limit the number of co-encapsulation events, as previously described [99,114]. Cells were encapsulated in YPD +Kan + 20% Optiprep to

achieve neutral buoyancy and select for the ItcR biosensor plasmid [122]. *E. coli* biosensors were co-encapsulated with the *Yarrowia* library members at a rate of 10 cells/droplet. Ensuring saturation of biosensor cells within the microdroplet helped minimize fluorescent signal variation due to growth variability of the *E. coli*. Microdroplets were generated using a Sphere Fluidics Fluorescence Activated Droplet Sorter (FADS), in 2.5% Pico-Surf™ in 3M Novec 7500. Oil flowrate was set to 1000uL/hr, and aqueous was set to 800uL/hr to achieve a droplet diameter of roughly 95um. Microdroplets were incubated at 28°C without shaking and with 12-hour oil exchanges to ensure oxygenation of emulsions.

Emulsion Fluorescence Collection

To achieve biosensor response curves in microdroplets, emulsions of co-encapsulated production strains were run on the FADS (Sphere Fluidics) at a 0 timepoint and 12- or 24-hour intervals to collect histograms of the fluorescence response. No voltage was applied as we were only collecting fluorescent response information. The sorting apparatus was set up as previously described [114]. Histograms were exported from the sorting software (Figure 4-6).

PiggyBac Library Emulsion Sorting

Post-incubation, microdroplets were sorted at four different timepoints, 24, 48, 72 and 96 hours. These timepoints were chosen to represent early (24) mid (48,72) and late (96) fermentation stages. Microdroplets were re-injected into Sphere's Pico-sort chip at a rate between 15-30uL/hr, with spacing oil containing no surfactant (3M Novec 7500) at a rate between 2000-3000 uL/hr to ensure sufficient spacing. Sorting was carried out by applying a voltage to microdroplets as they passed through the sorting chamber as

previously described [114]. Gating was applied to the top 1% of the population based on fluorescence measurements. 400,000 microdroplets were sorted at each condition to achieve 10X library coverage (with a 0.1 cell/droplet loading frequency). This resulted in the collection of roughly 1000 total sorted isolates. This number varied based on which timepoint was sorted. Specifically, the 24hr timepoint was the most difficult to achieve resolution for and required two iterative rounds of sorting to achieve statistically significant enrichment for improved itaconic acid producers. The second round only collected a total of 800 microdroplets.

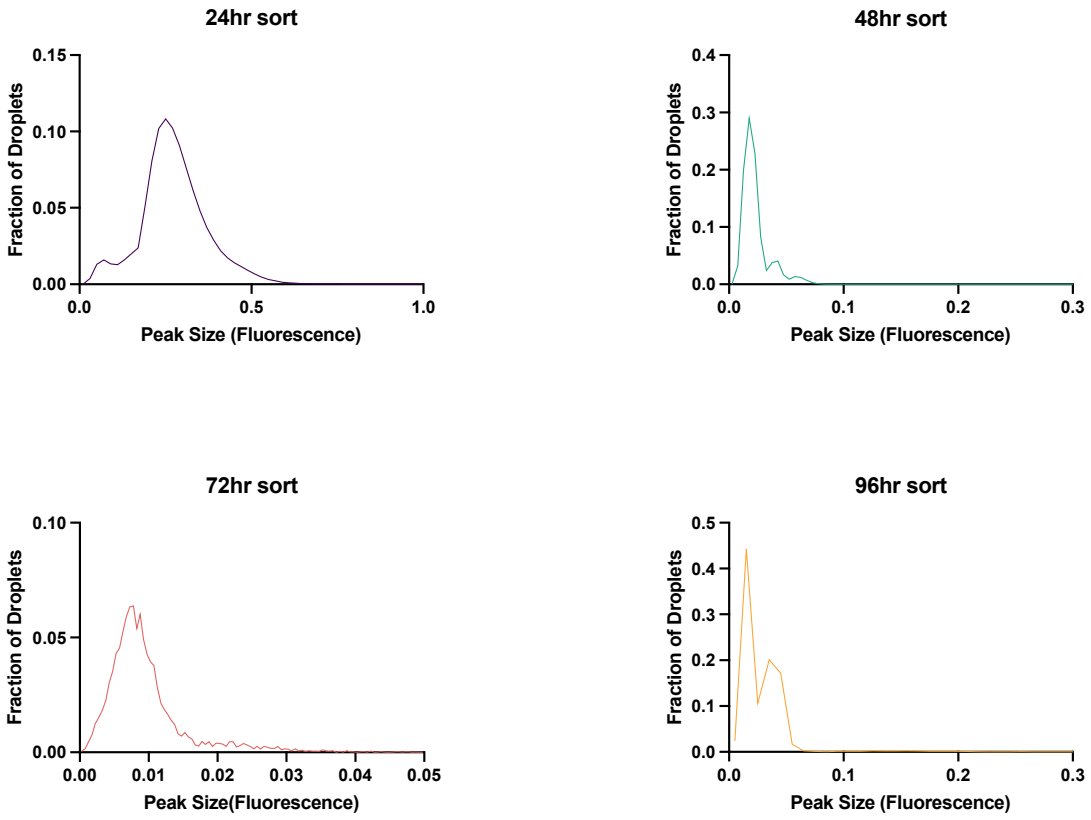

Figure 4-6: *Time course sorting for distinct production phenotype isolation.* Emulsions were generated and incubated at 28°C at 24-hour intervals. Sorts were carried out after their corresponding incubation times, and the top 10% of droplets were collected based on overall fluorescence, graphed here. Peak size/fluorescence are not comparable sample-to-sample as gain and laser strength were adjusted to account for signal strength. The 24-hour sample had to be iteratively sorted for a total of two sorts in order to achieve statistically significant differences between the unsorted and sorted population. This histogram represents that collected from the second sort, thus is not a full representation of an unsorted library at 24 hours.

HPLC Analysis

All supernatant samples were filtered with 0.2um nylon filters. Itaconic acid quantification was carried out on an Agilent HPLC, utilizing an Aminex Zorbax aq column with 0.1% TFA in Water as the aqueous phase and 0.1% TFA in Acetonitrile as the organic.

All standards for itaconic acid production in *Yarrowia lipolytica PO1J* were made in spent media that matched the composition and time of fermentation of the samples to reduce noise. However, for complex media (YPD), this increased total background in the spectra and rendered many of our timepoints before the 24-36 hour mark unreliable. Therefore, higher resolution productivity was gathered via fermentation in defined media (YSD) to reduce background.

Identification of Piggybac Insertion Sites

Thermal asymmetric interlaced (TAIL) PCR was performed as previously described [114]. Bands identified from round 3 of the protocol were gel extracted and sent to Eton Bio for sanger sequencing. Sequences received were blasted against the *Yarrowia lipolytica* genome and resulting gene disruptions were identified via KEGG[125]. Degenerate primers and nested primers used are in **Table 1** of the supplementary material.

Confocal Imaging of Droplets

For acquiring images, 10 μL of microdroplets were dropped on glass slides and covered with coverslips. The images were collected with a Yokogawa W-1 spinning disk confocal on a Nikon Eclipse-Ti2 inverted microscope (Nikon) equipped with a PLAN APO VC 60xA/1.20 water immersion objective. Biosensor emitted fluorescence was excited with a 488nm line from a 50mW Nikon's LUN-F Blues laser and collected with a quadruple bandpass dichroic mirror and a 525/30 emission filter. All images were taken with a Photometrics Prime95B EM-CCD camera controlled with NIS-Elements software. The brightness and contrast of the images captured were adjusted using NIS-Elements software. Image processing was performed using Fiji's ImageJ software.

Chapter 5: Single-Cell Phenotyping of Extracellular Electron Transfer via Cu(I)-catalyzed Alkyne-Azide Cycloaddition and Microfluidics[1]

5.1 Chapter Summary

Applications for the generalizable droplet system can span beyond screens exclusively for improved production. With the pico-injection based generalizable system discussed in Chapters 3 and 4, we have more control over addition of reagents beyond just cell-based biosensors. This injection schema, therefore, can be used to add in chemical sensors to screen for specific phenotypes. There is some precedent for this kind of microdroplet application[29,86,126,127] to perform bioprospecting for phenotypes of interest. In this chapter, I discuss the application of an oxygen-limited version of our droplet system in conjunction with a Cu(I)-catalyzed Alkyne-Azide Cycloaddition probe to identify extracellular election transfer[128]. Extracellular Electron Transfer (EET) is a cell function that is performed under anaerobic conditions; therefore, we tooled our microfluidic system to create and oxygen-limited environment. We first showcase controls that confirm our modifications lead to anaerobic gene expression, and then move forward with screening mixed cultures to enrich for EET-capable microorganisms.

These applications screen both a synthetic microbial consortia, composed of *S. cerevisiae, S. oneidensis* MR-1, and *E. coli* Nissle 1917, as well as a microbial consortium taken from lake water samples via the click-chemistry fluorescent probe to enrich the consortium for EET-performing organisms. Pre- and post-enriched samples were

[1] The work in this chapter was done in collaboration with Gina Partipilo in the Keitz lab at UT Austin. EKB performed microfluidic assays, data analysis, microscope imaging, and final ferrozine pourover assay.

sequenced to identify percent composition of microbial species, and the change in composition when exposed to the microfluidic screening method. In this chapter, I detail the establishment and utilization of an oxygen-limited, click-chemistry and microfluidics-based assay for identifying novel EET activity in microbial consortia from environmental lake water samples.

5.2 INTRODUCTION

Identification of EET capabilities in uncultured or uncharacterized microorganisms is important to increase our understanding of mechanisms and prevalence of EET[129]. Environmental microbial consortia analysis is typically limited to batch-based enrichments or full consortia phenotyping, losing the species-specific phenotyping necessary to elucidate novel or shared mechanisms of EET in unculturable or unidentified organisms[129]. Therefore, protocols for single-cell phenotyping of EET are needed to establish phenotype-genotype relationships. Some phenotyping strategies rely on either analysis of microbial consortia or isolation of specific microorganisms that are culturable within lab conditions[130-133]. Identification of unknown species capable of EET has been limited to these batch methods of enrichment. While these methods work well for identifying competitive respiration via EET, they do not allow for more nuanced analysis of individual genotyping. Recently, our collaborators developed a Cu(I)-catalyzed alkyne-azide cycloaddition assay capable of measuring EET via fluorescence[128].

In this Chapter, we utilize the CalFluor 488 probe[134] which undergoes a quenched-to-fluorescent transition upon triazole formation. Wild-type *S. oneidensis* MR-1 is a facultative anaerobe which preferentially respires on oxygen, when available, and under

anaerobic conditions will express the Mtr-pathway and begin EET. MR-1 utilizes a wide variety of both soluble and insoluble terminal electron acceptors including iron, copper, chromium and more[135-137]. As such, we redirected the electron flux to our copper redox active catalyst to generate Cu(I) *in situ* from Cu(II) and perform Cu(I)-catalyzed alkyne-azide cycloaddition. We saw that both kinetics and conversion could be incrementally limited through serial knockouts of the cytochromes in the Mtr-pathway. Thus, we set out to utilize this probe as a chemical assay for high throughput microfluidic screening of both synthetic and environmental consortia to isolate EET-capable organisms. Microfluidic screening of microbial consortia has been used for bioprospecting phenotypes of interest from environmental populations[127,138].

When compared to alternative growth-based assays for bioprospecting, the microdroplet screening has the potential to elucidate novel organisms not identified from the growth-based assay[138]. While growth-based enrichment for EET-capable organism identification may work, it likely principally leads to identification of organisms capable of out-competing other EET-capable microbes. In a microdroplet environment, however, single cells are isolated from their counterparts, removing growth-based competition seen in batch enrichments. Therefore, development of a microfluidic assay for EET identification is key in furthering discovery of these capabilities in novel or uncultured microorganisms.

In this chapter, I detail adapting existing microdroplet instrumentation for use with anaerobic cultures via an oxygen-limited protocol. While labs may have anaerobic-chamber-dedicated droplet systems, we wanted to expand the generalizability of our

system by creating an easy-to-assemble, closed-system to create an oxygen-limited environment *in situ*. We tested the oxygen leakage via purposefully oxygenated versus sealed protocols to show clear differences in the click-chemistry probe response. This allowed us to perform proof of concept experiments on mixed populations, before moving forward with an environmental sample of microorganisms. We also enrich these same microbial samples via an electrochemical cell in parallel, to ultimately compare population percentage changes between unenriched, electrochemical cell-enriched, and droplet-enriched samples. DNA sequencing results for these final analyses are still pending.

5.3 RESULTS

5.3.1 Optimization of CuAAC Fluorescent Probe Assay for Applications in Microdroplets

To port use of the CuAAC fluorescent probe into the droplet environment, we validated the reaction would not be inhibited by the reagents necessary for a stable emulsion. Previous probe applications ran experiments in defined M9 media[128], however our previous research as indicated rich media is helpful to preserve cell viability[114]. In addition, the probe is stored in DMSO, an organic solvent. Due to the nature of microdroplet-based emulsions, there was a concern that increased organic solvent within the droplets would lead to emulsion instability[128]. Finally, we tested whether the fluorinated oil (Sphere Fluidics) with fluorosurfactant (Sphere Fluidics) necessary to establish stable emulsions would also inhibit this reaction, rendering successful application of the probe in droplets unlikely.

These reactions were run anaerobically utilizing a plate reader assay with changes reflecting the environment of the microdroplets. We ran a control with increasing concentration of probe, as well as OD_{600} collections in the presence of increasing probe concentration, to determine if increasing this concentration would hinder cell growth (Figure 5-1a&b). While there was some minor reduction in growth, fluorescent response at the highest concentration of the CalFluour probe was significantly increased (Figure 5-1a). Finally, we ran a similar control in LB in the presence of fluorinated oil with or without the biosurfactant present (Sphere fluidics). Here, fluorescent response, while slightly lowered, was not entirely hindered by the extra reagents added to the reaction (Figure 5-1c).

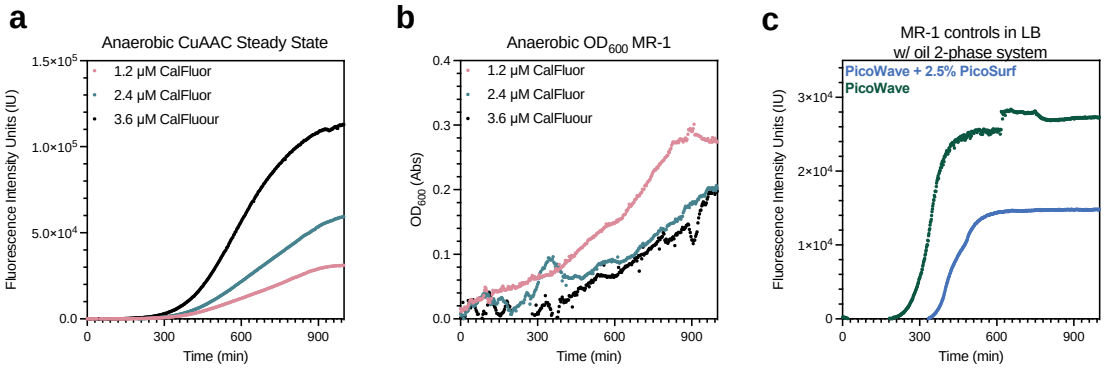

Figure 5-1: *Fluorescent and Growth data of CuAAC Probe Under Droplet-relevant Conditions in Plates.* a) Increasing concentrations of CalFluor probe were added to anaerobic cultures of MR-1, to determine the effect on total fluorescent signal via a plate-reader assay. b) The same samples from a were also measured at OD_{600} to determine cytotoxic effects of increasing probe concentration. c) Droplet-relevant materials (PicoWave™ with PicoSurf™ (Sphere Fluidics)) were added to MR-1 CuAAC reactions to determine if the fluorosurfactant or fluorinated oil would have inhibitory effects on the reaction.

5.3.2 Oxygen-Limited Protocol for CuAAC Probe Use and Fluorescent Signal Detection

Previous applications of the pico-injection system by us and other groups have shown its usefulness in both injection of cell-based biosensors[114], as well as chemical reagents for enzymatic activity[139-141]. We began by adapting our system for injection of the CalFluour reagents, by utilizing overnight anaerobic cultures, followed by atmosphere-exposed encapsulation at 1 cell per droplet. These emulsions were then incubated overnight in an anaerobic chamber and picoinjected the next day with the CuAAC reagents. A final overnight anaerobic incubation allowed for the maturation of signal. However, after this initial experiment, we saw no distinguishable fluorescent signal. While *S. oneidensis* is an

oxygen-sparging organism, we believe the metabolic burden of switching between aerobic and anaerobic metabolism led to poor cell growth within droplets. This hypothesis was supported by initial failed experiments where *sfGFP* containing MR-1 showed no cell viability when started aerobically and incubated anaerobically over 24 hours. Furthermore, when starting a mixed MR-1 and Δ*mtrC*Δ*omcA*Δ*mtrF* encapsulation aerobically, after a 24-hour incubation with CuAAC reagents, the knockout yielded less than 1% MR-1 (via ferrozine pour-over assay) of the recovered colonies. These data helped inform the transition into the oxygen limited encapsulation system.

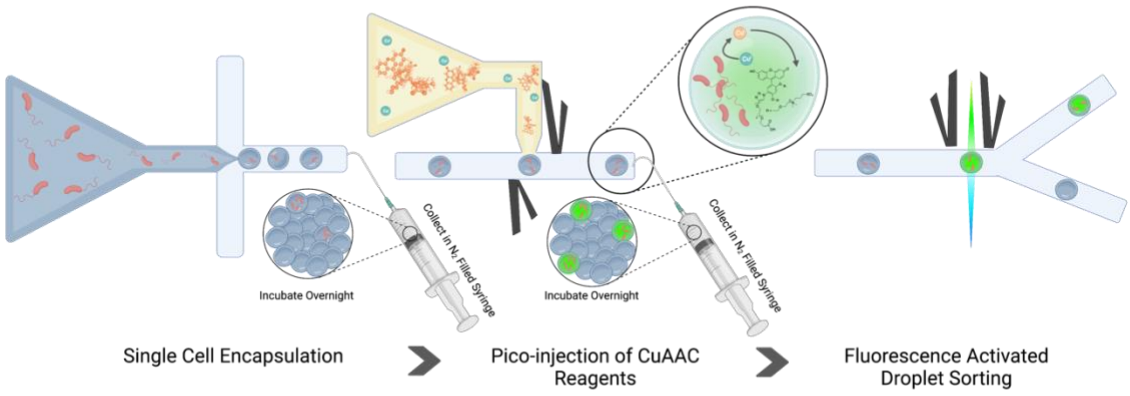

Figure 5-2: *Oxygen-limited protocol for Anaerobic cultivation of Microbes and EET activity analysis.* Microbial consortium or homogenous culture is encapsulated in water in oil emulsions. Overnight anaerobic cultures are diluted to account for a loading density of 1 cell per every 10 droplets. All samples are prepared in an anaerobic chamber with sealed tubing and syringes to reduce oxygen exposure. The chip is loaded by clipping tubing and applying positive pressure upon plumbing. Samples are collected in an anaerobic syringe filled with gas to avoid pressure build up. Samples are then incubated in the syringe anaerobically overnight, and a similar protocol is used to set up the pico-injection chip, with all reagents and samples loaded and prepped anaerobically. Samples are collected in an anaerobic

syringe with a gas mixture to avoid pressure build up and incubated overnight for signal maturation. Finally, samples are exposed to oxygen before sorting for improved fluorescence and collected for emulsion breaking an analysis via colony PCR or DNA isolation and sequencing.

To establish CalFluour probe response sufficiently in droplets, we chose to run control experiments with an oxygen-limited droplet set up in parallel with a purposefully aerobicized sample (Figure 5-2). The oxygen-limited droplet set up consisted of preparing samples within an anaerobic chamber (Coy Anaerobic) in pre-sealed syringes. Collection syringes were also sealed and filled with gas in the anaerobic chamber to reduce pressure buildup as the emulsion was collected. Pico-surf™ (Sphere Fluidics) in Pico-wave™ (Sphere Fluidics) was degassed under nitrogen and used to fill a fully sealed syringe within the anaerobic chamber as well. All tubing was clipped prior to plumbing the generation chip and loaded utilizing positive pressure to flush out any atmospheric gas that may leak into the lines as they're clipped and plumbed. This protocol allowed us to use a system that was already established for aerobic organism screening and provided flexibility beyond moving the whole instrument into an anaerobic chamber.

Once encapsulated, both MR-1 *S. oneidensis* controls were grown anaerobically overnight and pico-injected with a similar oxygen-limited set up the next day. However, the aerobicized control was pico-injected in the presence of non-degassed oil, exposing the sample to minor amounts of oxygen. Finally, reactions were incubated anaerobically overnight to allow for signal maturation. The next day, emulsions were exposed to atmospheric oxygen, imaged, and run on the droplet sorter utilizing previously developed methods[114] (Figure 5-3a). Fluorescent signal was collected for both samples to determine

if the fluorescent probe provided detectable fluorescent signal for one or both samples. Unsurprisingly, fluorescent signal was only detected in the fully oxygen-limited set up (Figure 5-3b). This allowed us to move forward in applying this assay for EET detection to both synthetic and environmental microbial consortia.

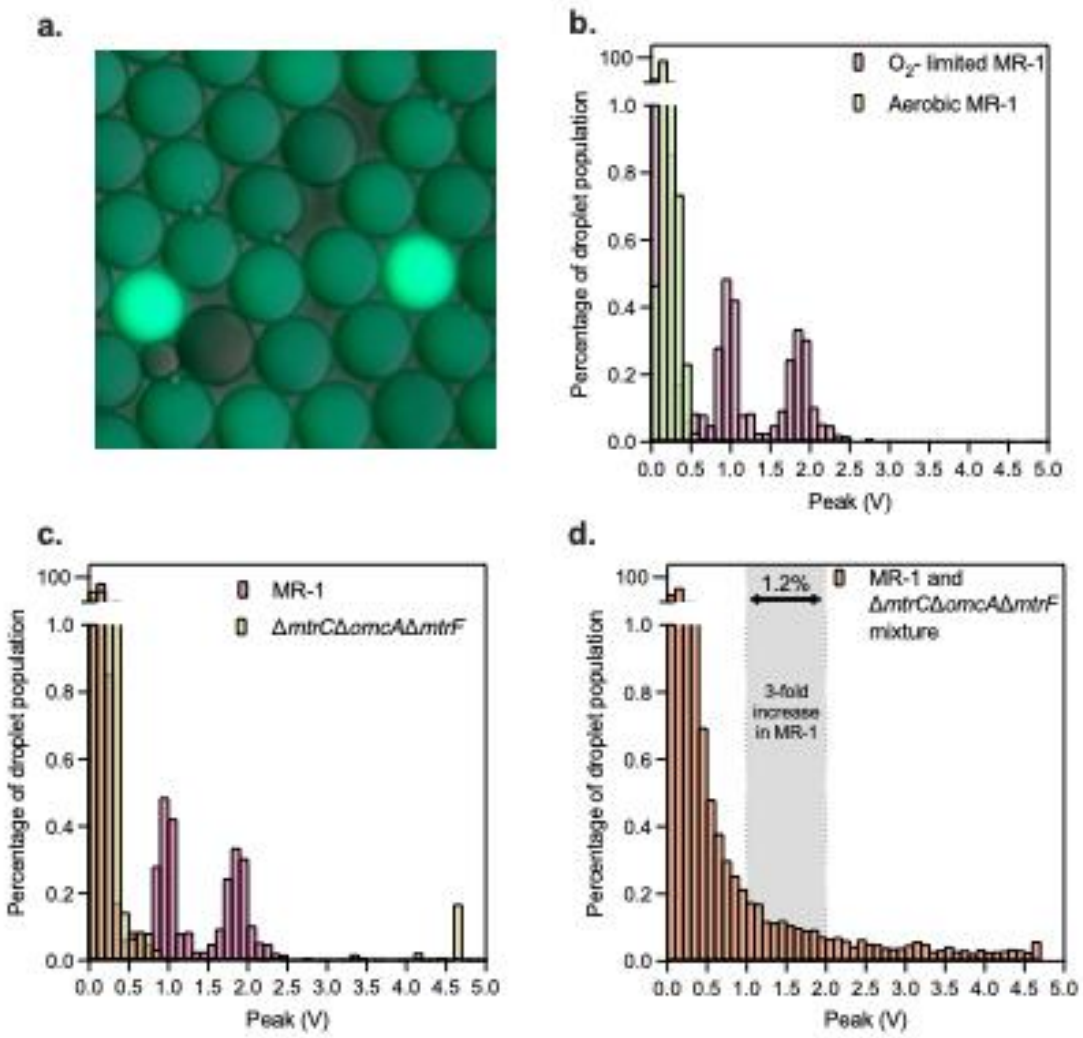

Figure 5-3: *Histograms of Fluorescence from Mixed and Separate Emulsions*. a) Image of fluorescent response in droplets after pico-injection of CuAAC materials. b) Oxygen-limited (pink) versus aerobic (green) prepared emulsion histograms. Fluorescent response shows a clear increase under anaerobic conditions versus aerobic. c) MR-1 (pink) versus JG596 knockout (yellow) samples run through the oxygen-limited protocol. While the JG596 sample

shows some moderate background fluorescence-as expected based on previous experimental development[128], the MR-1 emulsion shows distinguishable fluorescent signal at the 1-2V mark. c) Mixed emulsion of MR-1 and JG596 was run through the oxygen-limited protocol and sorted. The sorting gate is highlighted, and the emulsion was broken. Pre- and post-sort samples were plated for single colonies and colony PCR was run to identify enrichment of MR-1, resulting in a 3-fold enrichment in the sorted over the unsorted condition.

5.3.3 CuAAC for the detection of Extracellular Electron Transfer (EET) in microdroplets

Previous applications of the CuAAC fluorescent probe have been able to detect differences in EET output and capability in a 96-well plate format[128]. A triple knockout strain of *S. oneidensis* (JG596), missing *mtrC, omcA,* and *mtrF* was shown to have poor conversion, while still showing some background activity, likely due to *mtrA* not being knocked out and having some catalytic activity[128,142]. To establish this oxygen-limited protocol for sorting for EET activity, two individual emulsions containing either wildtype MR-1 or the JG596 knock out were run through the same protocol described above. Histograms of the droplet fluorescence were taken for each to determine the fluorescent signa from each phenotype (Figure 5-3c). Once the signal differences were established, a mixed culture of MR-1 and JG596 was encapsulated (Figure 5-3d).

The sorting protocol for this mixed emulsion followed that described above with a gate set to sort droplets identified in the previously defined range for MR-1 fluorescence in droplets, as well as an unsorted control that had no gate applied (Figure 5-4). Once collected, emulsions were broken and grown anaerobically overnight. These samples were plated and we performed both a ferrozine reduction assay (Figure 5-4a) and colony PCR on the *mtrC* locus from the 1-2V sample, to determine if there was enrichment of MR-1

wildtype over the JG596 knock out (Figure 5-4b). These PCRs showed a 3-fold enrichment of MR-1 from the unsorted control in the sorted sample, indicating two key points, the first is that where the voltage gate was applied is likely where the majority of the MR-1 signal lies, and the second is that this system is capable of enriching for EET phenotypes from a mixed culture (Figure 5-4b).

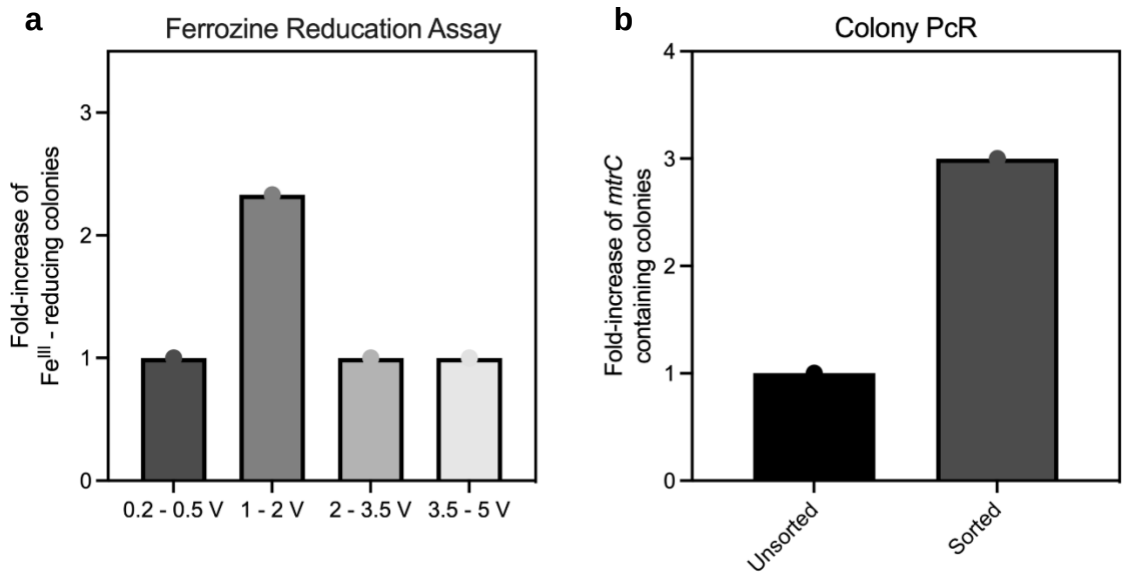

Figure 5-4: *Sort Confirmation of MR-1 JG596 mixed Emulsions.* a) Ferrozine reduction assays in a plate-based screen indicate MR-1 enrichment in the 1-2V sorted sample b) Colony PCR of the 1-2V sample shows a 3-fold increase in *mtrC* containing cultures over an unsorted control.

5.3.4 Microfluidic CuAAC enables the detection of EET-capable bacteria in multi-species co-cultures

Based on the success of the mixed knockout versus wildtype *S. oneidensis* screen, we wanted to establish the protocol for screening synthetic microbial consortia for EET capability. We chose to combine three different microorganisms all capable of anaerobic growth at 30°C, *S. cerevisiae*, *E. coli*, and *S. oneidensis.* These cultures were grown

anaerobically, separately, overnight and combined at roughly equal ratios immediately before encapsulation. This protocol followed that previously established for successful sorting of the MR-1/JG596 emulsion to enrich for EET. To determine enrichment, pre- and post-screening samples were DNA purified and sent for rDNA sequencing to determine enrichment of *S. oneidensis* in the sorted sample. This data is still pending.

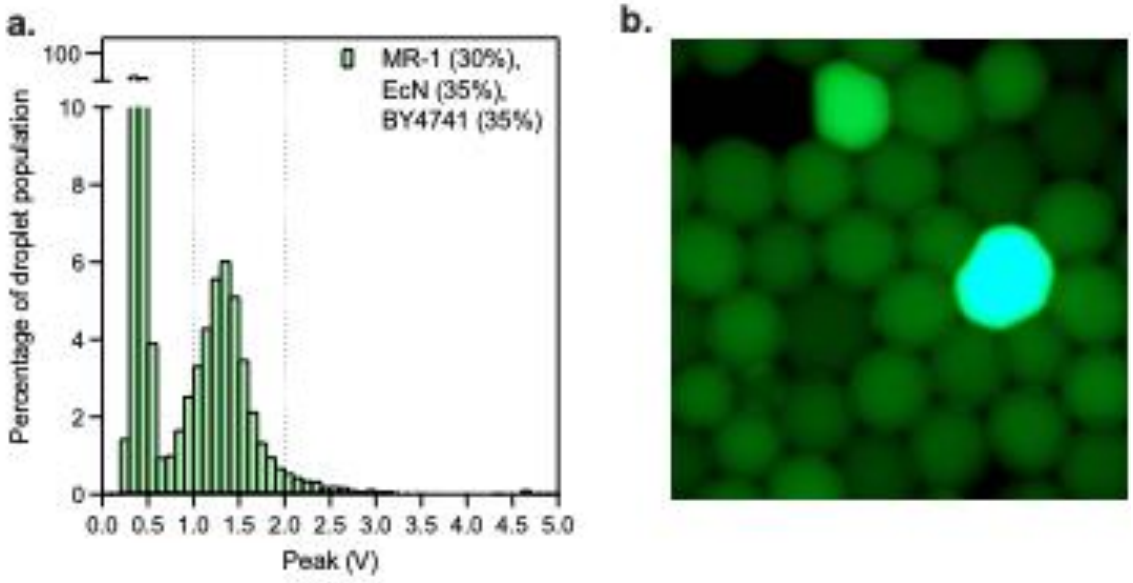

Figure 5-5: *Synthetic Microbial Consortium Screening via the Oxygen-limited Protocol.* a) fluorescent histogram of synthetic microbial consortia on the FADS instrument (Sphere Fluidics). The sorting gate is shown between the dotted lines and matches the sorting condition from the MR-1 JG596 mixed population. b) image of reacted droplets post-pico-injection but pre-sort. Signal differentiation is clear here, indicating the EET probe has seemingly no background in the presence of non-EET organisms.

To further showcase the generalizability and importance of this protocol, we chose to move forward with environmental microbial consortia to screen for EET. Lake water samples were collected, and cell counts determined via a live/dead cell stain (Figure 5-7). These counts were correlated with an OD spectrum to determine the best wavelength for

estimating cell count. Once determined, this microbial consortium sample was encapsulated at an estimated 1 cell per droplet, and treated with the same protocol described above. In parallel, electrochemical cell enrichments were run in a 200mL flask over 5 days to establish a field-relevant baseline for species previously shown to exhibit EET activity from this environmental sample.

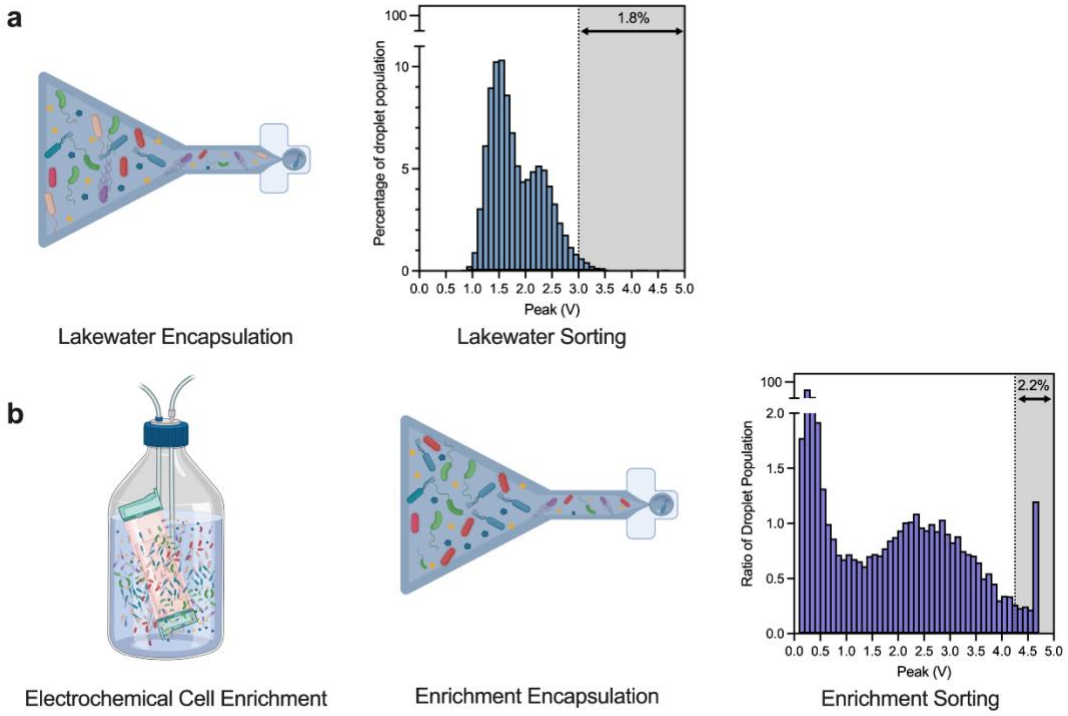

Figure 5-6: *Environmental and Electrochemical Cell Enrichment Utilizing the Oxygen-limited Droplet Protocol.* a) Environmental samples were directly encapsulated and screened via the oxygen-limited droplet protocol b) Environmental cells were initially enriched via an electrochemical cell, these samples were then used to seed an emulsion and screened via the same oxygen limited protocol. For the enriched sample, the gain had to be turned down to the intensity of the fluorescent signa, indicating pre-enrichment via electrochemical cell was successful for increasing EET-capable organisms.

Since EET activity within this microbial consortium was not previously characterized, multiple sorting gates were used to isolate the best EET performers from the sample (Figure 5-6). A histogram of this initial screen showed relatively low signal, indicating a low population of EET-capable organisms (Figure 5-6a). As a secondary screen, the electrochemical cell-enriched sample was used as a seed culture for a second screen, to determine if this culture known to be enriched for EET showed improved signal differentiation when compared to the unenriched sample (Figure 5-6b). This was the case, so much so, that we had to reduce the gain on our droplet sorter to obtain a distinguishable signal suitable for sorting the best EET performers based on fluorescence. This result indicates that the initial sample still likely contained many non-EET capable organisms. As a terminal step, we took this final droplet-enriched sample and ran a ferrozine pour-over assay on the plated cultures to look for iron reduction (Figure 5-7). This resulted in images indicating clear enrichment for EET-capable organisms. Finally, we wanted to confirm the population enrichment via sequencing analysis.

Samples from every step of enrichment, both droplet and electrochemical cell-based, were collected for rDNA analysis. We are still awaiting this data collection to determine levels of enrichment.

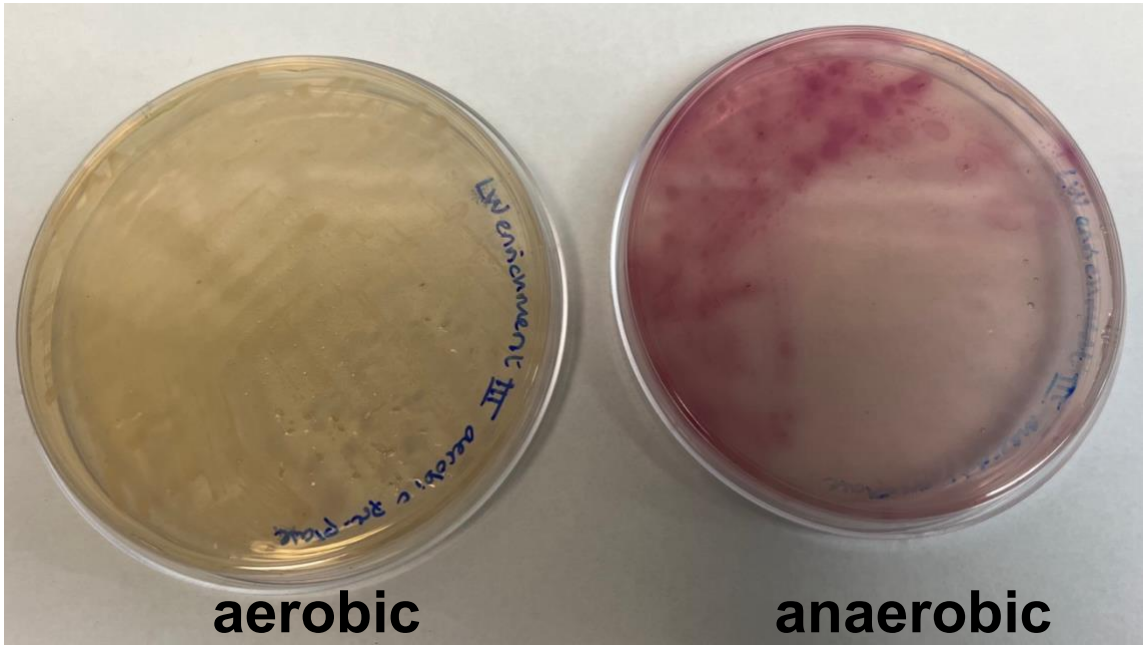

Figure 5-7: *Ferrozine pour-over assay of final enriched population.* As a final assessment of EET activity, the microbial population isolated after both electrochemical cell enrichment and our oxygen-limited protocol was plated. These plated colonies were used in a ferrozine pour-over assay run on plates pre-grown either in an anaerobic or aerobic environment. The anaerobic response and lack of aerobic response show that we were screening for exclusively anaerobic EET-capable microorganisms. Sequencing results from this population are still pending.

5.4 DISCUSSION

In this chapter, we established an oxygen-limited protocol for enrichment and identification of EET-capable microorganisms. This protocol allowed us to tool our existing instrumentation for cultivation and analysis of anaerobic microbes by utilizing a sealed-system to reduce atmospheric oxygen exposure. Proof of concept experiments with synthetic microbial consortia showed the impressive enrichment capabilities of this system, leading us to pursue high throughput phenotyping of an environmental microbial consortia.

DNA analysis of both synthetic and environmental microbial consortia in parallel with a traditional enrichment technique are still pending.

We foresee this protocol being useful for more bioprospecting applications, including more environmental samples from different locations, as well as potential microbiome analyses relevant to agriculture or human health. The easily-converted system means that groups already tooled with a microfluidic screening platform would easily be able to analyze anaerobic organisms and phenotype them in high throughput.

5.5 Materials and Methods

Reagents used in this chapter

CalFluor 488 (Click Chemistry Tools), alkyne-PEG4-acid (Click Chemistry Tools), copper(II) bromide ($CuBr_2$, Sigma-Aldrich, 99%), 2-(4-((bis((1-(tert-butyl)-1H-1,2,3-triazol-4-yl)methyl)amino)methyl)-1H-1,2,3-triazol-1-yl)acetic acid (BTTAA, Click Chemistry Tools >95%), sodium DL-lactate ($NaC_3H_5O_3$, TCI, 60% in water), sodium fumarate ($Na_2C_4H_2O_4$,VWR, 98%), EPES buffer solution ($C_8H_{18}N_2O_4S$, VWR, 1 M in water, pH = 7.3), potassium phosphate dibasic (K_2HPO_4, Sigma-Aldrich), potassium phosphate monobasic (KH_2PO_4, Sigma-Aldrich), sodium chloride (NaCl, VWR), dimethyl sulfoxide (cell culture grade, Sigma-Aldrich), ammonium sulfate ($(NH_4)_2SO_4$, Fisher Scientific), magnesium(II) sulfate heptahydrate ($MgSO_4 \cdot 7H_2O$, VWR), trace mineral supplement (ATCC), casamino acids (VWR), Lysogeny Broth (Thermofisher), OptiPrep (Sigma-Aldrich), Pico-Gen™ 60 x 60 single aqueous (Sphere Fluidics), Pico-Wave ™ (Sphere Fluidics), Pico-Break ™ 1 (Sphere Fluidics), Pico-Mix ™ (Sphere Fluidics), Pico-

Surf™ (Sphere Fluidics), Tubing (VWR), Needles (VWR). All media components were autoclaved or sterilized using 0.2 μm PES filters.

Click Chemistry Controls

Fluorescent assay controls were completed in 96-well plate format outlined previously (Partipilo 2022). Briefly, fluorescence emission was collected on a BMG LABTECH CLARIOstar plate reader with a 491 (±14) nm and an emission collection at 538 (±38) nm). Different to previously published methods, all reactions were performed in LB broth supplemented with lactate (20 mM) as a carbon source and fumarate (20 mM) as the primary electron acceptor. Stock solutions of 1 M sodium fumarate and 60 w/v% lactate solutions were stored at 4 °C until use. Aliquots of 3 mM of CalFluor 488 (46) were created in DMSO and stored frozen at −80 °C until use. Aliquots of 4 mM alkyne-PEG4-acid were created in sterile water and stored at −20 °C until use. An 8 mM copper bromide stock in sterile water was created and stored at 4 °C and mixed with an equal volume amount of 48 mM freshly made stock of BTTAA in sterile water. Control reactions included increasing the concentration of CalFluor 488, running with new buffer conditions, and using a dual phase system where both Pico-SURF ™ and Pico-WAVE ™ were added in 100 μL volumes to the 96-well plate and the LB-reaction mixture was run in a 100 μL volume floating on top of the oil layer. The reaction was then placed into the plate reader for analysis and allowed to react for between 10 and 24 h.

Oxygen-Limited Encapsulation of *S. oneidensis*

Overnight cultures were grown in a Coy Anaerobic Glovebox containing a humidified atmosphere at 3% hydrogen content and the balance nitrogen. The cultures were

either started by picking a single colony into argon-sparged LB broth supplemented with 20 mM sodium lactate (2.85 µL of 60% w/w sodium lactate per 1 mL culture). After overnight growth anaerobically at 30°C, cell cultures were diluted to an OD_{600} of 0.00006 into a solution of 40 mM lactate, 80 mM fumarate, 20 wt% OptiPrep in LB broth. Inside of the anaerobic chamber, 1 mL of the cell solution was loaded into the aqueous syringe (BD). The tubing was prepared by heat-sealing one end, and the remaining was loaded onto a needle (Thin-walled 0.7mm), inside of the anaerobic glove box, the needle was attached to an empty syringe and pressure was pulled three times and held for 30 seconds each time and then vented after each round to remove any O_2 present in the tubing. Additionally, 10 mL of argon-sparged 2.5% 008-FluoroSurfactant in Pico-Wave™ was loaded into the oil syringe (SGE) and capped with a sealed, sparged needle and tubing. A collection syringe was prepared by pulling 6 mL of the Coy Anaerobic Glovebox atmosphere and capped with a sealed, sparged needle and tubing.

A Pico Droplet Single Cell Encapsulation System (Sphere Fluidics) was used. All sealed syringes were removed from the anaerobic chamber. Each syringe was loaded under a 100 µL/h positive pressure before clipping the heat-sealed end and connecting to the 60 × 60 droplet maker (Pico-Gen) chip. To make the emulsion, the syringe pumps were increased to 1000 µL/h (aqueous) and 1200 µL/h (oil). The system was allowed to calibrate before cutting and plumbing the collection syringe (BD). After encapsulation, emulsions were incubated at 30 °C 24 h to allow for growth within the droplets.

Encapsulation of Mixed population

Overnight cultures were grown in a Coy Anaerobic Glovebox containing a humidified atmosphere at 3% hydrogen content and the balance nitrogen. The cultures were either started by picking a single colony into argon-sparged LB broth supplemented with 20 mM sodium lactate (2.85 µL of 60% w/w sodium lactate per 1 mL culture) or 20 mM dextrose (for *E. coli* Nissle 1917 and BY4741). After overnight growth anaerobically at 30°C, cell cultures were diluted to an OD_{600} of 1.8 x 10^5 (MR-1), 7.8 x 10^5 (EcN), and 0.008 (*S. cer*) into a solution of 40 mM lactate, 40 mM glucose, 80 mM fumarate, 20 wt% OptiPrep in LB broth. These reflect a loading density overall of 1 in 10 droplets, and the loading ration of 30:35:35 (MR-1:EcN:*S. cer*). The remaining encapsulation is the same as that of *S. oneidensis* outlined above.

Oxygen-Limited Pico-Injection of CuAAC components

A 5X Cu(I)-catalyzed Alkyne—Azide Cycloaddition solution (70 µM CalFluor 488, 2 mM Alkyne-Peg4-acid, 2 mM Cu:BTTAA (1:6)) was prepared from a 3.2 mM stock of CalFluor 488 in DMSO, 8 mM stock of $CuBr_2$ in water, 48 mM stock of BTTAA in water and a 4 mM stock of Alkyne-Peg_4-Acid in water in a Coy Anaerobic Glovebox containing a humidified atmosphere at 3% hydrogen content and the balance nitrogen. Into a 1 mL syringe, 200 µL of the solution was loaded into the aqueous syringe (BD) and capped with a sealed, sparged needle and tubing. Additionally, 10 mL of argon-sparged Pico-Wave™ was loaded into the oil syringe (SGE) and capped with a sealed, sparged needle and tubing. A collection syringe (12 mL BD) was prepared by pulling 6 mL of the Coy Anaerobic Glovebox atmosphere, adding in 500 µL of 008-Fluorosurfactant (5%) in Pico-Wave™, and capped with a sealed, sparged needle and tubing. The droplets were then

transferred to a 1 mL syringe (BD) and capped with a sealed, sparged needle and tubing. All sealed syringes were removed from the anaerobic chamber, loaded onto the syringe pumps set at 15 µL/h (aqueous), 100 µL/h (droplets) and 1200 µL/h (oil) before cutting off the sealed end and plumbing into a Pico-Mix™ chip. The system was allowed to calibrate before cutting and plumbing the collection syringe (PGE). After pico-injection, emulsions were incubated at 30 °C 24 h to allow for growth within the droplets.

Fluorescent Sorting of *S. oneidensis* heterogeneity

For droplet sorting, a Single Cell Assay and Isolation platform (Sphere Fluidics) with a 488 nm laser 244 and 525/50 emission filter (GFP) was used. PMT setup of the system was set to a gain of 0.8 (With the exception of the lakewater which was at a gain of 1.0) for both PMT 1 (GFP channel/bandpass 525/50) and PMT 2 (large bandpass 650/150). Peak detection minimum was set at 0.07 and maximum at 100. The minimum width was set to 0.18 and maximum set to 100. Sorting gates were determined based on population distribution and manually drawn above 1V(MR-1), 3.5V (Lakewater), 4.25V (LWEETE). The top 3000 droplets were then collected using an applied voltage of 0.3V. After sorting, the emulsion was broken with 100 µL Pico-Break (Sphere Fluidics) and cultured in LB broth supplemented with 20 mM lactate at 30 °C overnight. Each sample was split, freezing 500 µL as a cryogenic stock, and diluting 10 µL into the anaerobic chamber to start the creation of the next generation of sort.

Colony PCR of *S. oneidensis* Heterogeneity Sort

Colony PCR was conducted by picking single colonies from a streaked LB-agar plate into 20uL of 20mM NaOH, 0.1% Triton. Samples were boiled at 100°C for 10 minutes and briefly spun down. Once cooled, the supernatant was used as template in a PCR reaction with these primers: F-TACGGCGTTGAAGATGTTGTAG, R-ACAATGCCTTTACTGGTGGAAG, utilizing Q5 polymerase (NEB) per manufacturer's instructions. Samples were run on a 1% agarose gel and imaged via UV.

Mixed Population Sorting

Environmental microbial consortia sorting was carried out as described above. Samples for DNA analysis were collected at both pre- and post-enrichment steps. Sorting gates were applied at different voltages for lakewater and pre-enriched samples, 3.0V and above for lake water, representing 1.8% of the total population, and 4.25V and above for the pre-enriched sample representing 2.2% of the total population.

Lake water enrichment

Environmental and post-droplet enriched samples were enriched in an anaerobic batch system (Lu 2003). Briefly, 10 mL of sediment containing lake water was inoculated into a 250 mL batch electrochemical cell filled with filter-sterilized lake water. The cells were provided with lactate at 0.1mM as a carbon source, and iron-containing sediments were isolated in dialysis tubing. The cells were allowed to grow for 5 days and redox potential, pH, and aqueous Fe(II) concentrations were measured daily.

Sequencing and Analysis

DNA samples from both pre- and post-enriched samples via the oxygen-limited droplet protocol and electrochemical cell enrichments were purified via the AllPrep DNA/RNA kit from Qiagen per manufacturer's instructions. These samples included both the synthetic and environmental microbial consortia. Samples were sent to Mr. DNA for bacterial taxonomic analysis. Fungal analysis wasn't performed due to interest specifically in EET capable bacteria due to easier identification of potential gene clusters. Taxonomic analysis from Mr. DNA was used to identify fold changes between pre- and post-enriched samples and data was plotted using Prism™ software.

Chapter 6: Cytotoxicity Assay for Insect Cells in Microdroplets Utilizing IVTT

6.1 Chapter Summary

In the last chapter, applications for cell phenotyping in microdroplets were detailed. To expand the repertoire of droplet applications even further, this chapter will discuss work conducted to establish a cytotoxicity assay for Sf9 insect cells in microdroplets utilizing known active toxins. Agricultural research for novel insecticides requires screening methods to determine insecticide activity[143-145]. While there are some plate-based assays to determine toxin activities for insect cells, higher throughput methods are necessary to screen engineered protein libraries for potency. In this chapter, the lab-strain of insect cells, Sf9, is used as a model organism with known receptors that respond to toxins previously shown to kill these cells[1]. In this chapter, I detail establishing the use of a SYTOX cell death dye in droplets, along with utilizing *in vitro* transcription translation (IVTT) for toxin expression in microdroplets. Using sequential pico-injection, I show the capability of the droplet system to enrich for active toxin, confirmed via DNA isolation and a quantitative PCR protocol. This method allows for analysis of enrichment without having to maintain cell viability-a key part of determining improved cytotoxic compounds. In addition, use of IVTT for toxin expression results in easily-isolated DNA from enriched samples, indicating the possibility of utilizing this application in conjunction with next generation sequencing techniques for larger library analysis.

[1] Previously identified from Bayer Crop Sciences

6.2 INTRODUCTION

The need to develop assays for high throughput screening for bioactive compounds is on the rise[143,145,146]. Specifically, agricultural sciences are expanding screening efforts to identify novel pesticides and herbicides[143]. These screening efforts primarily rely on plate-based methods which can be time consuming and may require automation instrumentation[143]. To identify novel insecticides, developing screening tools for large library analyses is key. Inherent to the screening method for successful cytotoxic agents is the need to sort for cell death, rather than improved production[114], or growth[29,102]. At its face value, this work requires sorting for a negative signal rather than a positive signal, thus requiring techniques capable of separating true negatives from all false negatives and empty droplets in the system.

Cytoxocity assays for microdroplets are primarily being utilized to study mammalian cell behavior[147,148]. In depth single-cell characterization has been carried out with cell-death dyes that intercolate DNA when the cell membrane is compromised[147,148]. While many of these applications use microdroplets as nano-reaction chambers, observing the cell-cell interactions over time, the fluorescent dye response makes it an easy-to-adapt screen to utilize in fluorescence-activated droplet sorting (FADS). Thus, this approach can take the negative signal of a cell death and invert it to create a positive, sortable signal.

There is minimal precedent for cultivating insect cells in droplets [149]. There are few examples of cultivating Sf9 cells in pico-liter sized droplets, as many of the microfluidic approaches with these cells are not droplet based[150]. To this end, we decided to first establish what Sf9 cell viability looked like in microdroplets via encapsulation with

SYTOX dead cell stain. We also took timelapse pictures of growth in droplets, indicating the cells are not only still viable in droplets, but are also capable of proliferating, despite a slightly diluted media due to the addition of opti-prep to achieve neutral buoyancy. After establishing Sf9 viability baselines in droplets, we moved forward with cell death assays utilizing established toxin-cell interactions. Once the methods were established, we applied them to sequential-pico-injected emulsions to isolate active toxin versus either an mCherry or inactive toxin control.

6.3 RESULTS

6.3.1 Sf9 Viability in Microdroplets

While there have been some microfluidic protocols utilizing Sf9 cells as a model organism, growth and viability within pico-liter sized droplets has yet to be characterized[151,152]. Therefore, we chose to begin by determining cell viability within microdroplets via encapsulation and extended incubation to show cell growth. Sf9 cells were cultivated as described in the manufacturer's instructions. Cell viability was determined pre-encapsulation by utilizing a Nucleocounter^{TM} Y100. While this system is primarily intended for counting yeast cells, hemacytometer counts of Sf9 cells were taken and correlated with those identified via the Nucleocounter^{TM}. These measurements agreed and led to us primarily utilizing the Nucleocounter^{TM} for all cell count and viability analyses outside of microdroplets.

To determine cell-death ratios, SYTOX dye was co-encapsulated with Sf9 cells, both to act as an indicator of cell viability, and to determine if extended incubation with the dye had cytotoxic activity (Figure 6-1a). In both conditions, Sf9 cells were shown to continue to proliferate post-encapsulation, and viability was maintained at a similar

percentage to that determined pre-encapsulation (Figure 6-1b). This indicates that microdroplets are a reasonable environment for Sf9 cell analysis, and in particular, for analyzing activity of cytotoxic reagents, as viability is maintained at a base-line level.

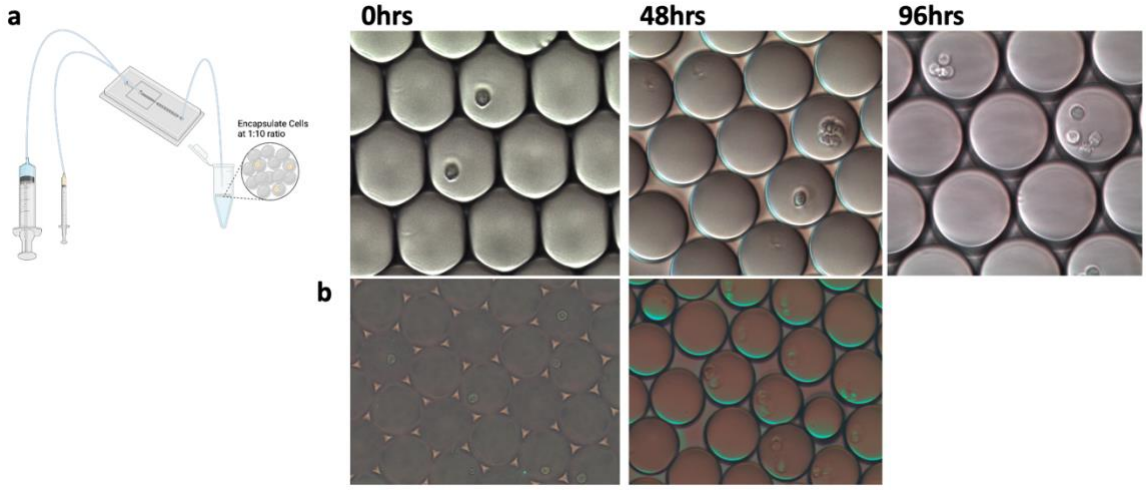

Figure 6-1: *Sf9 Insect Cell Encapsulation and Incubation.* a) Droplet schematic and time-lapse images of Sf9 cell growth in droplets. Growth can be seen continuing up to the 96-hour timepoint, indicating cell viability is maintained with extended droplet incubation. b) confirmation of cell viability with SYTOX dye during in-droplet incubation also confirms Sf9 cell viability in droplets with no fluorescent signal.

6.3.2 Establishing SYTOX Dye as a Cell-Death Indicator in Microdroplets

Utilizing SYTOX dye (Invitrogen) as an indicator of cell death, Sf9 cells were encapsulated in water-in-oil emulsions at a rate of 1 cell for every 10 droplets to avoid dual encapsulations. This encapsulation rate is based on calculations assuming Poisson distribution statistics. Immediately after encapsulation, increasing concentrations of soluble toxin 4, a toxin known to kill Sf9 cells, were pico-injected (Figure 6-2a). After injection, emulsions were incubated at 28°C for up to 18 hours to allow for signal resolution but reduce response to extended incubation with even limited levels of active toxin. This experiment resulted in a dose-response curve with fluorescence measured on our FADS

system, indicating the dye can be used as a biosensor for toxin efficacy (Figure 6-2b). This proof-of-concept experiment supported the ultimate application for encapsulating receptor-libraries and pico-injecting known toxins (Figure 6-2c).

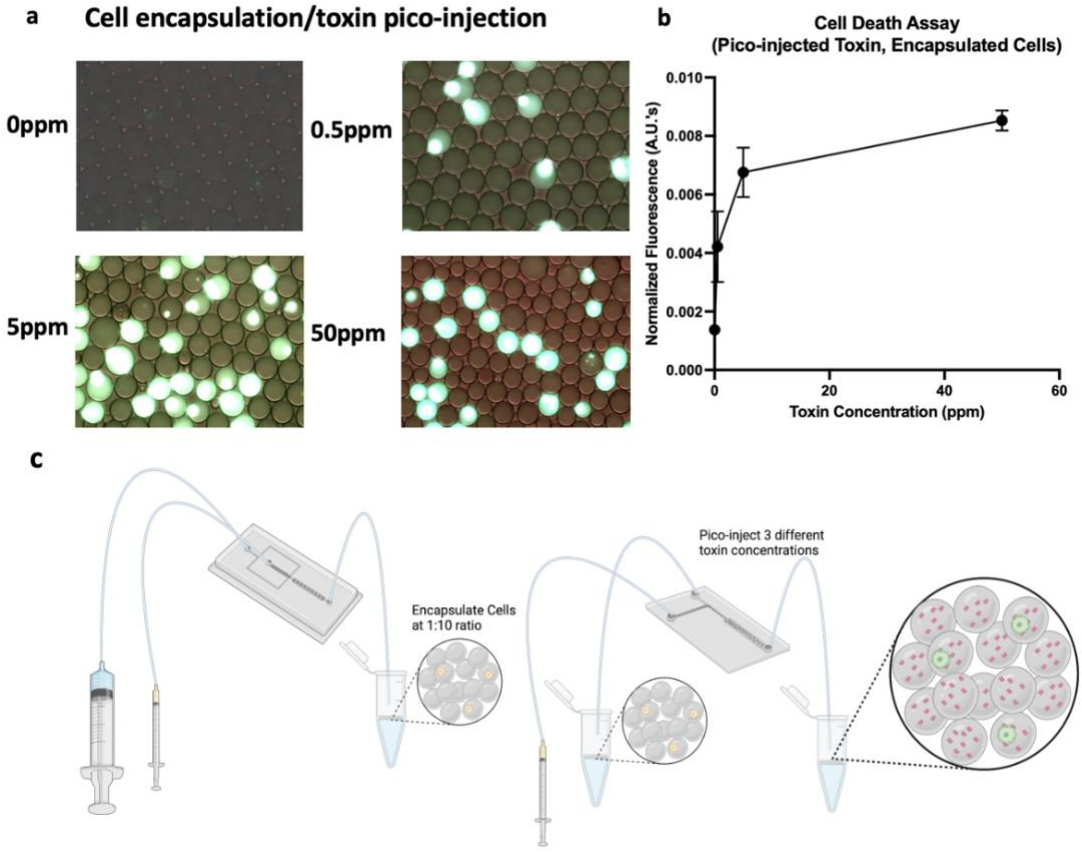

Figure 6-2: *SYTOX™ Dye as a Sf9 cell death indicator in microdroplets.* a) images of injected emulsion at increasing concentration of soluble toxin. Qualitative fluorescent analysis shows an increase in fluorescent signal as the toxin concentration increases. b) Quantitative fluorescence of emulsions injected with increasing concentration of soluble toxin. Emulsions were run on our FADS instrument and the top 10% of droplets were used to identify increased fluorescence corresponding with increased toxin. c) Schematic of droplet workflow for establishing a toxin dose-response in microdroplets.

6.3.3 IVTT for Pico-injected Toxin Response in Microdroplets

Due to the nature of this protocol, we needed to establish a method for identifying active toxin after successfully killing Sf9 cells. Soluble toxin injection and isolation would have necessitated western blot analyses to determine enrichment of active toxin. While this is not impossible, difficulties in protein degradation and isolation post-screening would lead to difficulty in enrichment confirmation. To this end, we chose to express toxins utilizing *in vitro* transcription-translation (NEB PURExpress), which would allow for DNA isolation post-screen to analyze enrichment based on presence of active toxin.

The use of IVTT in droplets has been reported in literature; however, we have not utilized it on our droplet system [153]. To confirm the IVTT reaction (NEB PURExpress™) would not be inhibited by our fluorinated oil or fluorosurfactant, we used a T7-mCherry linear DNA construct and measured fluorescence over time via a fluorescent microscope (Figure 6-3). Incubation was carried out at 37°C for three hours. This duration is one hour over the recommended incubation time from the PURExpress kit, however fluorescence measurements continued to rise significantly after the two-hour mark. It should be noted that this reaction did influence emulsion stability. At roughly the one-hour mark, the emulsion began to look more unstable, resulting in droplet merging and fragility when pipetting into the chamber slides for photographing. Despite this, we moved forward with this protocol to test toxin expression and efficacy in droplets.

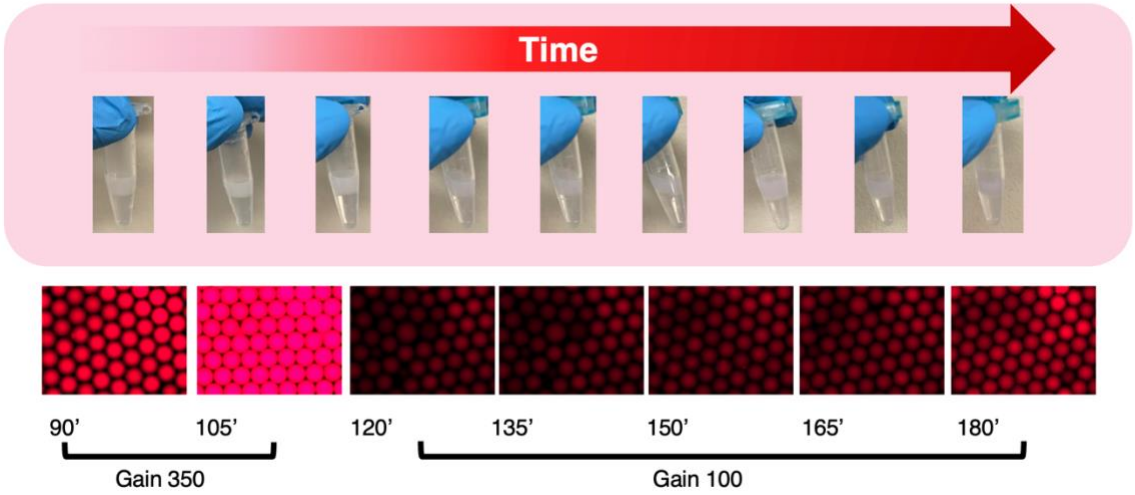

Figure 6-3: *mCherry IVTT reaction in microdroplets.* Timelapse images of a T7-mCherry construct utilizing NEBPURExpress™ IVTT reaction shows an increase in fluorescent signal up to the 180 minute mark when incubated at 37°C. Emulsion stability decreased over time leading us to switch fluorosurfactants for further IVTT reactions in microdroplets.

To test the IVTT-expressed toxin 4 efficacy against Sf9 cells, we approached this using the same avenues as described for the cell death assays with soluble toxin. First, individual cells were encapsulated, followed by injecting the pre-incubated toxin 4 IVTT reaction. This resulted in clear active toxin, as seen by the increase in fluorescence compared to a no-injection control (Figure 6-4b). However, this also destabilized the emulsion, resulting in an emulsion that we would not be able to sort reliably (Figure 6-4a). This same issue was seen in the IVTT encapsulation for toxin expression followed by pico-injection. However, here the emulsion was unstable before pico-injecting, thus making it more susceptible to severe droplet merging (Figure 6-4a). Previous IVTT applications in droplets have utilized an alternative fluorosurfactant from RAN biosciences. We chose to switch from the Sphere-fluidics Pico-surf™ biosurfactant, to that from RAN biosciences, which showed improved emulsion stability.

While the response to the IVTT made toxin was lower than the lowest concentration we tried (Figure 6-4a), it was enough to show a significant response in droplets. Thus, we moved forward in utilizing this scheme to sort for Sf9 cell death from sequential pico-injection of either mCherry or Toxin 4. Here, cells were encapsulated at a ratio of 1 cell for every 10 droplets, and the emulsion was split in half. Half of the emulsion was injected with an mCherry IVTT reaction versus a toxin 4 reaction. Droplets were sorted as previously described[114], and the histogram was collected to identify if there were distinct populations representing those expressing mCherry, and those that resulted in cell death/showed increased fluorescence upon SYTOX dye DNA intercolation (Figure 6-5a). This histogram showed three distinct populations, including empty droplets, droplets with cells but not responding to toxin, and droplets with cells that had responded to toxin/showed SYTOX dye response (Figure 6-5a). Due to the nature of this experiment, cell viability is inherently lost in the sorted population. Therefore, DNA isolation and PCR was performed on DNA from the sorted population versus an unsorted control. The sorted population showed a brighter band for the toxin, while the unsorted population showed both bands. The isolated population that was thought to be full droplets injected with the mCherry reaction showed exclusively the mCherry band on an agarose gel (Figure 6-5b).

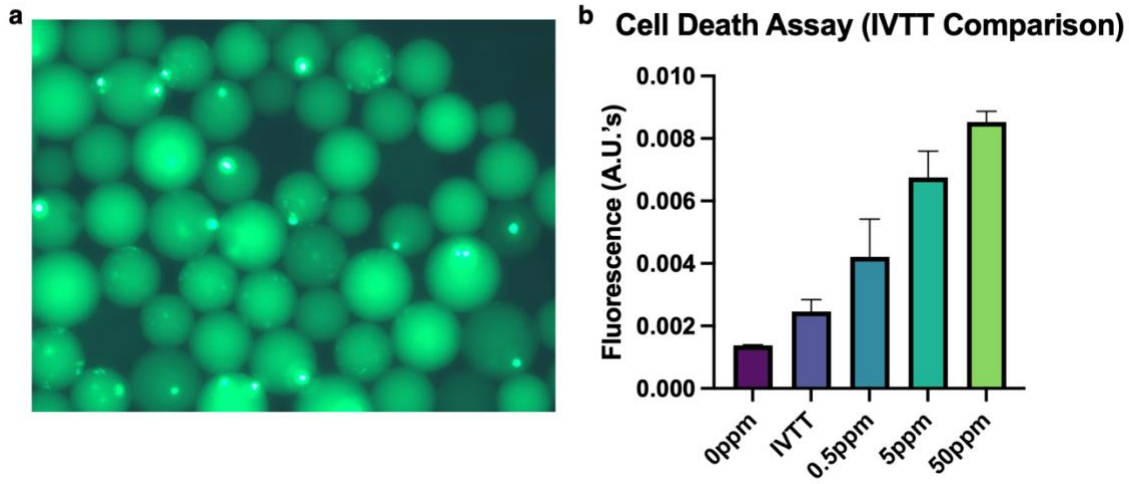

Figure 6-4: *IVTT-expressed Toxin injected compared to soluble toxin injection.* a) image of emulsion post-injection shows decreased emulsion stability, as seen with previous mCherry experiments (Figure 6-3). b) plotted fluorescence compared to soluble toxin injection shows a low fluorescent response that is still above the background seen from injection no soluble toxin, indicating it could be utilized to sort active toxin.

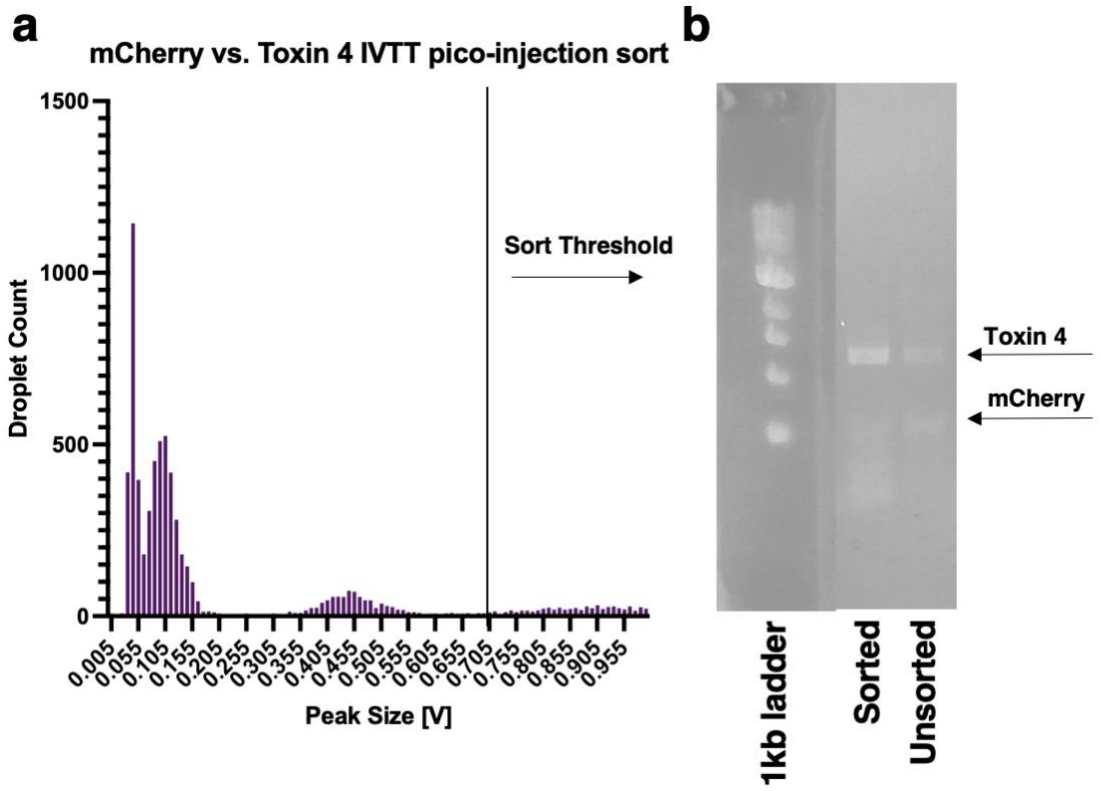

Figure 6-5: *Control Sort of pico-injected mCherry versus Toxin 4 IVTT reactions.* a) image of histogram collected during sort. This shows 3 distinct populations, the lowest indicating empty droplets, the middle showing droplets containing cells that did not show cytotoxic effects and the highest showing cells that had died and increased fluorescent response. The sorting threshold was applied to the highest fluorescent population to isolate only those droplets that contained active toxin and dead cells. b) gel of purified DNA from sorted and unsorted samples. Both show a band for each IVTT template (Toxin 4 and mCherry), but the band from the sorted sample is brighter for toxin 4. While not conclusive, this indicated our sorting system would work for identifying active versus inactive toxin in a future sort.

6.3.4 Mock sorting using RecA expressing Sf9 cells and IVTT-expressed toxin 4

After successfully sorting a sequentially-picoinjected sample utilizing IVTT reactions as the treatment, we wanted to determine if this system could be used to pull out active toxin from inactive toxin via using receptor A-expressing Sf9 cells and corresponding inactive toxin (toxin 5) or active toxin (toxin 1) IVTT reactions. The

emulsion was encapsulated at a ratio of 1 cell per droplet, and two separate emulsions were made, one of the RecA-expressing cells, and one of a blank control.

Due to low cell viability upon encapsulation, to reduce the number of false positive samples, emulsions were sorted immediately post-encapsulation to remove any fluorescence resulting from SYTOX dye response to encapsulated dead Sf9 cells. For sorting, the set up was similar to the mCherry experiment (Figure 6-5), involving splitting the emulsions into two, but rather than injecting an mCherry reaction as a control, inactive toxin 5 was injected. Post-injection, emulsions were mixed back together and allowed to incubate for toxin response up to 8 hours. The sample was then sorted for fluorescence and DNA was extracted from the sorted samples for both the RecA emulsion and the blank control.

To quantitatively determine sorting efficiency, a qPCR experiment was performed on a pre-sorted control from each emulsion, as well as sorted samples from the receptor-expressing and blank emulsions. This resulted in a clear enrichment of the active toxin 1 in the final, isolated, DNA sample (Figure 6-6a). Due to the inherent variability introduced in mixing the emulsions post-injection, the pre-sorted condition was used to normalize DNA concentrations to and indicate the level of enrichment for each condition.

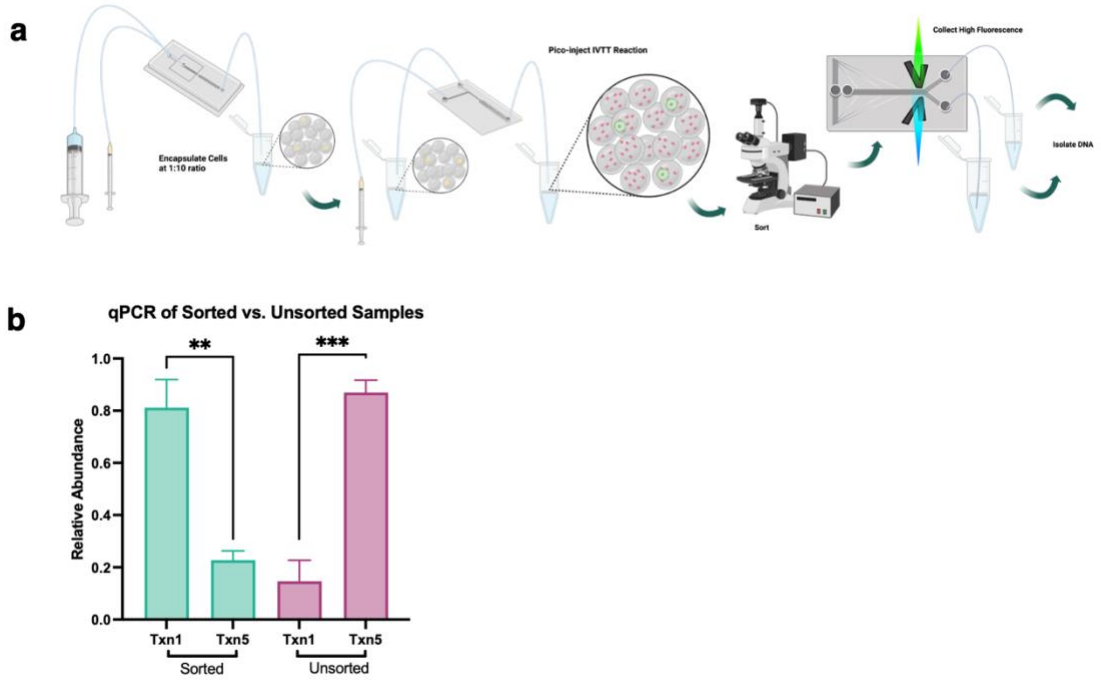

Figure 6-6: *Proof of Concept Sort to Screen for Active vs Inactive Toxin.* a) Workflow schematic of screen (Made with Biorender). b) Isolated DNA from screen was used as a template in a qPCR assay to determine relative abundance of DNA in each sort condition. This assay indicated a clear enrichment for active toxin (txn1) in the sorted sample versus the unsorted control, indicating isolated fluorescence from the screen did pull out those droplets containing active toxin and dead Sf9 cells.

6.4 Discussion

High throughput screening options to determine novel biocontrol agents are limited[143]. The ability to pair potential active toxins expressed via IVTT with their receptor pairs in microdroplets allows for a cytoxicity screen with an easy-to-isolate result that could be utilized for next generation sequencing approaches in the event of screening a library. Future applications of this screen could include looking for novel synthetic receptors in insect cells, as well as engineered toxin libraries to screen for efficacy. A simple, DNA sequencing read out makes it possible to screen for the desired phenotype-cell death- without losing the important information leading to that response.

Sf9 cell cultivation in microdroplets is limited, therefore our characterization of their growth and viability in microdroplets is paramount to move forward in other types of screens. While overall cell viability was a challenge in adding noise to a cell-death assay, pre-sorting emulsions to remove fluorescence not caused by active toxin allowed for a clean sort and enrichment of active toxin over inactive toxin in our final experiment.

6.5 Materials and Methods

Sf9 encapsulation

Sf9 cells were grown according to manufacturer's instructions (Thermofisher). Cell viability was determined via a nucleocounter Y100, which was confirmed via hemacytometer to correlate cell counts. Encapsulation was performed by first filtering cells via tube filtration (Tubes) and diluting for encapsulation of 0.1 cells per droplet in Sf900 media with 20% optiprep to maintain neutral buoyancy. Encapsulation took place as previously described[114]. Briefly, the cell-containing aqueous phase was run at 1000uL/hour along with 2.5% Pico-Surf in Pico-Wave (Sphere Fluidics). The emulsion was collected in a 15mL conical tube and incubated at 28°C for up to 96 hours. Images of cell growth were taken on a bright field microscope.

Cytotoxicity Assays using SYTOX dye

For cell-viability check emulsions, 1uL of undiluted SYTOX dye was added directly to the aqueous phase pre-encapsulation. These emulsions were utilized immediately for pre-sorting and pico-injection of active toxin. To determine dose-response with SYTOX dye, cells were encapsulated as described above, and incubated for up to 96hrs. Emulsions were then pico-injected with soluble toxin at the following

concentrations: 0ppm, 0.5ppm, 5ppm, 50ppm. Toxin was diluted with phosphate-buffered saline (pH 7.4) along with 1uL/10mL of SYTOX dye. Pico-injection occurred as previously described[114]. Briefly, the incubated emulsion was run at 200uL/hr, with spacing oil (Pico-wave) running at 1500uL/hr. The aqueous phase was run from 10-50uL/hr to maintain injection without creating a separate toxin-only emulsion. Post-injection, emulsions were incubated 18 hours after determining the ideal response-time. Emulsions were then imaged on a fluorescent microscope and run on our FADS system to collect fluorescence response based on injected toxin.

IVTT optimization in microdroplets

To allow for information-recovery from droplets, *in vitro* transcription-translation was used to express active toxin. Reactions were assembled per manufacturer's instructions (NEBPURExpress). To test emulsion stability of IVTT expression, a T7-mCherry linear DNA construct was used as template. The reaction was assembled on ice and encapsulated as described above, without dilution. The emulsion was incubated at 37°C for 3 hours to determine expression kinetics in the droplet environment. The emulsion was imaged on a fluorescent microscope and fluorescence was quantified via Fiji ImageJ. Emulsion stability was a problem for extended incubation, thus we chose to utilize Ranbiotech fluorosurfactant for all future IVTT reactions, as this surfactant showed better emulsion stability.

IVTT pico-injection

To showcase screening for active toxin, IVTT reactions were pre-incubated to express soluble active toxin, mCherry, or inactive toxin. Reactions were assembled and

incubated per manufacturer's instructions (NEBPURExpress), and used as the aqueous injection as described above for soluble toxin injection. Emulsions assembled as described above containing Sf9 cells were pico-injected with active toxin, mCherry, or inactive toxin. Fluorescent response was determined via running emulsions on our FADS system and collecting histograms of green fluorescence from SYTOX dye.

Sorting for Cytotoxicity via Fluorescent Response

For sorting experiments, emulsions of RecA-expressing Sf9 cells and a no-DNA control were encapsulated at 0.1 cells/droplet as described above, with SYTOX dye and incubated overnight. Emulsions were pre-sorted to remove fluorescent response before pico-injecting active or inactive IVTT toxin. The emulsions were then split in half, and either Toxin 1 or inactive toxin was then pico-injected. These samples were recombined and incubated overnight before sorting utilizing our FADS system. To determine where the change in fluorescent response occurs, the no-DNA control was run on our FADS system to collect a histogram of a control-population. This sample was followed by the RecA emulsion, and sorts were performed as previously described[114] with a gate applied at 0.5V to isolate only those droplets pico-injected with active toxin.

DNA isolation and qPCR analysis of sorted active versus inactive toxin

Sorted samples were collected in 1.7mL microcentrifuge tubes and the emulsion was broken as previously described[114] with the following changes: 100uL of pico-break (sphere fluidics) was added to the sorted emulsion and then vortexed at high speed for 30 seconds. The sample was then spun down briefly at max speed, and the aqueous phase was collected.

DNA purification was performed on the isolated aqueous phase via the Thermo Fisher DNA clean up kit per manufacturer's instructions. Isolated DNA was used as template in qPCR reactions. qPCR was performed utilizing primers provided by Bayer Crop Sciences, and the Takeda Biosciences 2X SYBR Green Master Mix per manufacturer's instructions. Assays were run on a Viia-7 qPCR instrument and data was analyzed via Prism software.

Chapter 7: Conclusions and Future Work

The field of metabolic engineering has progressed significantly thanks to modern DNA synthesis and sequencing. However, screening these information-rich libraries or populations poses a problem. High throughput screening for phenotypes that rely on extracellular measurements, whether small molecule production and secretion or extracellular electron transfer, or receptor expression, require pico-liter sized reaction chambers to miniaturize the environment and allow for high throughput screening of these pico-reaction chambers. Development of generalizable screens for these phenotypes of interest is paramount in furthering the depth and breadth of information gathered from each iteration of the canonical Design-Build-Test-Learn cycle. In order to learn more per cycle, design and screening of gene perturbation libraries for phenotypes of interest is key. We have addressed this by building and applying an sgRNA CRISPR-based library for gene regulation for improved growth on alternative carbon sources (**Chapter 1)**. Then, to achieve production-based information from this library and others, we developed a generalizable microfluidic screening method that pairs cell-based biosensors with production strains to sort for improved small molecule production (**Chapter 2**). Importantly, we also conducted investigation into the droplet environment by correlating biosensor response in droplets to temporal-based screens. This resulted in isolation of distinct production phenotypes based on when the *Yarrowia lipolytica* gene perturbation library was screened.

To expand our applications of this system, we then collaborated with the Keitz lab at UT, to utilize a click-chemistry probe for extracellular electron transfer into our droplet

system. This allowed us to screen not only synthetic microbial consortia, but also an environmental microbial consortium for EET phenotypes. Finally, we expanded the use of this system even further by applying it for cytotoxicity screens utilizing Sf9 insect cells. This work showcases the breadth of possibilities in the microfluidic screening field. In particular, identification of key aspects for portability of screens will inform future work porting off-chip assays into microdroplets for miniaturization and high throughput analysis.

CELL-BASED BIOSENSORS IN MICRODROPLETS FOR SMALL MOLECULE PRODUCTION SCREENING

We have showcased the use of cell-based biosensors for improved small molecule production for multiple production phenotypes. In particular, secreted small molecules are difficult to screen using traditional FACS-based methods. This is due to these screens being optimized almost exclusively for intracellular accumulation of product[28]. When transferred into droplets, however, this environment allows for small molecule accumulation, while still maintaining the phenotype-genotype connection when isolated via FADS (**Chapter 3**). Our generalizable platform showcased use of this system for three different production organisms, and three different cell-based biosensors. In addition, screening on this system of a million-member library led to identification of an isolate capable of producing significantly more product than the base strain. Further investigation of this droplet environment led to us to screen a piggyBac gene perturbation library in *Yarrowia lipolytica* for production phenotypes at different stages of growth **(Chapter 4).** This resulted in isolation of two distinct phenotypes (one growth-associated and one non-growth associated), caused by different piggyBac transposon insertions. These phenotypes were

maintained from droplet to flask scale, indicating that the miniaturization of these screens showcases the droplet condition is comparable to standard culturing methods for small molecule production.

Future applications of this work could investigate the droplet environment for small molecule production in even more depth. Applications that mimic batch-fed fermentations would likely allow for identification of gene perturbations that remain relevant when ported into bench-top bioreactors. In addition, in-depth analysis of cell growth in droplets could indicate which stages of fermentation are affected by the droplet environment. This work would improve the understanding of utilizing high throughput screening to not only test strains, but also learn more from them and increase the information obtained through each round of the Design-Build-Test-Learn cycle.

MICRODROPLET SCREENS FOR CONSORTIUM PHENOTYPING AND CYTOTOXICITY ASSAYS

At the foundation of metabolic engineering is the understanding of environmental microorganisms. A fundamental part of this understanding is the ability to phenotype novel environmental interactions with either engineered substrates, like insecticides, or with probes to identify phenotypes of interest that have the potential to be utilized or learned from in a laboratory environment. Adapting these screening methods into miniaturized assays utilizing microdroplets facilitates broadening this knowledge to improve designing and building synthetic consortia or substrates to mimic or inhibit these interactions.

Extracellular electron transfer in microorganisms has great potential for serving as energy sources[129]. However, this field has yet to fully characterized possible pathways for this phenotype in microbes not already acting as model organisms. Identification of microbes capable of this reaction from environmental samples can serve the field in two

distinct manners. Not only would identifying organisms better at EET than lab strains help in furthering synthetic biological applications of this in the future, but EET microbes can also serve as an indication of environmental health[129]. Thus, high throughput screening methods to identify these microbes in environmental samples help to lay a foundation for future work in characterization and applications in investigative bioprospecting.

We showcased the applicability of our microdroplet system by porting a click-chemistry based probe into droplets and developing an oxygen-limited platform for cultivating anaerobic microbes (**Chapter 5**). Through this system, we identified microbes that had not been previously characterized for EET but shared some homology to known EET pathway genes. In addition, based on data collected from our screening process, we could identify levels of EET ability in microbes by utilizing this probe, as well as a ferrozine pour over assay.

Future applications of this platform will likely include more bioprospecting of other environmental samples, as well as use for synthetic biology approaches. Not only is it capable of identifying EET-capable microbes, but it could also be used to engineer improved EET in these microorganisms. *S. oneidensis* has a well-characterized pathway for EET, and protein engineering of this pathway may lead to either improvement of EET ability, or a deeper understanding of the pathway itself, through direct changes in the structure and function of the proteins. This could have implications for use of biochemical cells to produce energy that have as of yet, come up short[129].

The bioactive chemical space has recently become fairly small thanks to resistance to antibiotics, pesticides, and insecticides[143]. Staying ahead of this resistance requires human ingenuity to design and build novel reagents to achieve the same kinds of cell death seen with previous applications[143]. High throughput screening of cytotoxicity is inherently difficult, as traditional methods are either throughput limited (in the case of plate-reader

based assays), or lose the phenotype link to the reagents responsible for cell death (Flow cytometry). Thus, cytotoxicity assays in microdroplets allow for pairing microorganisms of interest with tried-and-true cell death dyes, and libraries of potential cytotoxic reagents.

In this dissertation, I detailed our ability to develop a cytotoxicity assay for Sf9 insect cells in microdroplets. This resulted in a final screen that enriched for known active toxin over an inactive control (**Chapter 6**). Future applications of this work would expand upon what we have accomplished, by developing libraries of potential active toxin to identify novel toxins with insecticidal qualities. This would require further development of IVTT in droplets to create the library, which does have precedence in the literature[153]. In addition, Sf9 cells are much larger than the biosensor cells we typically utilize for pico-injection, thus design and development of a pico-injection chip with a larger channel would allow for introduction of toxin heterogeneity through encapsulation, and cell homology via pico-injection. However, with the foundation of a cytotoxicity screen laid, technical challenges could be overcome swiftly to identify novel insecticidal toxins for agricultural applications.

APPENDIX 1

EKB0065	GPI18 Fwd A/E sgRNA -522	GCTATTTCTAGCTCTAAAACGATGATGACGAG CTTACTCGTTTCGT
EKB0066	GPI18 Fwd A/E sgRNA -332	GCTATTTCTAGCTCTAAAACGCAGTCGACGAG CTTACTCGTTTCGT
EKB0067	GPI18 Fwd A/E sgRNA -199	GCTATTTCTAGCTCTAAAACATAACTGACGAG CTTACTCGTTTCGT
EKB0068	GPI18 Fwd A/E sgRNA 68	GCTATTTCTAGCTCTAAAACTACATCGACGAG CTTACTCGTTTCGT
EKB0069	FUN26 -542 A/E Fwd	GCTATTTCTAGCTCTAAAACCCTTAGGACGAG CTTACTCGTTTCGT
EKB0070	DIM1 -503 Fwd A/E	GCTATTTCTAGCTCTAAAACTGGTATGACGAG CTTACTCGTTTCGT
EKB0071	DIM1 -419 Fwd A/E	GCTATTTCTAGCTCTAAAACGAAGAAGACGA GCTTACTCGTTTCGT
EKB0072	DIM1 -172 Fwd A/E	GCTATTTCTAGCTCTAAAACATGGATGACGAG CTTACTCGTTTCGT
EKB0073	DIM1 58 FWD A/E	GCTATTTCTAGCTCTAAAACTCAGTAGACGAG CTTACTCGTTTCGT
EKB0074	GUS1 87 Fwd A/E	GCTATTTCTAGCTCTAAAACTAAGTTGACGAG CTTACTCGTTTCGT
EKB0075	ALG12 -361 FWD A/E	GCTATTTCTAGCTCTAAAACTGAAGAGACGAG CTTACTCGTTTCGT
EKB0076	COR1 -500 FWD A/E	GCTATTTCTAGCTCTAAAACTCTAACGACGAG CTTACTCGTTTCGT
EKB0077	GRS2 -525 FWD A/E	GCTATTTCTAGCTCTAAAACTCGAAGGACGAG CTTACTCGTTTCGT
EKB0078	GRS2 -393 FWD A/E	GCTATTTCTAGCTCTAAAACACATCAGACGAG CTTACTCGTTTCGT
EKB0079	GRS2 -170 FWD A/E	GCTATTTCTAGCTCTAAAACAGGTTAGACGAG CTTACTCGTTTCGT
EKB0080	GRS2 69 FWD A/E	GCTATTTCTAGCTCTAAAACGGTGGTGACGAG CTTACTCGTTTCGT
EKB0081	GPI18 Rev A/E sgRNA-522	aagtttaccccatcgatcATTCAGCAACGGACAGATGAT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0082	GPI18 Rev A/E sgRNA -332	aagtttaccccatcgatcATCATCGTCGTCATGGCAGTC CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0083	GPI18 Rev A/E sgRNA -199	aagtttaccccatcgatcAATAAATAGGAAAGTATAACT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0084	GPI18 Rev A/E sgRNA -199	aagtttaccccatcgatcAATTAGGCAGTTTGATACATC CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0085	FUN26 -542 A/E Rev	aagtttaccccatcgatcAAATAGCCCACGCACCCTTAG CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG

EKB0086	DIM1 -503 Rev A/E	aagtttaccccatcgatcAAGTGGCAACGTCAGTGGTAT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0087	DIM1 -419 Rev A/E	aagtttaccccatcgatcATGGTGTTTACTAGTGAAGAA CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0088	DIM1 -172 Rev A/E	aagtttaccccatcgatcAGGTAATACACAAGGATGGAT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0089	DIM1 58 Rev A/E	aagtttaccccatcgatcAAGAAACATTTGAGTTCAGTA CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0090	GUS1 87 Rev A/E	aagtttaccccatcgatcAAACTCCATAGCTATTAAGTT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0091	ALG 12 -361 REV A/E	aagtttaccccatcgatcATGCCTCGTAACAGGTGAAGA CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0092	COR1 -500 REV A/E	aagtttaccccatcgatcAAAGTAACGCAATTATCTAAC CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0093	GRS2 -525 REV A/E	aagtttaccccatcgatcATACTCTTTCAGCAATCGAAG CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0094	GRS2 -393 REV A/E	aagtttaccccatcgatcATGTACCAAAATCACACATCA CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG
EKB0095	GRS2 -170 REV A/E	aagtttaccccatcgatcATTTATATTCTGTCAAGGTTAC TGATGAGTCCGTGAGGACGAAACGAGTAAGC TCG
EKB0096	GRS2 69 REV A/E	aagtttaccccatcgatcACGTTTGAGATATATGGTGGT CTGATGAGTCCGTGAGGACGAAACGAGTAAG CTCG

REFERENCES

1 Ekas, H., Deaner, M. & Alper, H. S. Recent advancements in fungal-derived fuel and chemical production and commercialization. *Curr Opin Biotechnol* **57**, 1-9, doi:10.1016/j.copbio.2018.08.014 (2019).

2 Shepelin, D., Hansen, A. S. L., Lennen, R., Luo, H. & Herrgard, M. J. Selecting the Best: Evolutionary Engineering of Chemical Production in Microbes. *Genes (Basel)* **9**, doi:10.3390/genes9050249 (2018).

3 Sun, J. & Alper, H. S. Metabolic engineering of strains: from industrial-scale to lab-scale chemical production. *J Ind Microbiol Biotechnol* **42**, 423-436, doi:10.1007/s10295-014-1539-8 (2015).

4 Yu, T. *et al.* Reprogramming Yeast Metabolism from Alcoholic Fermentation to Lipogenesis. *Cell* **174**, 1549-1558 e1514, doi:10.1016/j.cell.2018.07.013 (2018).

5 Bittihn, P., Din, M. O., Tsimring, L. S. & Hasty, J. Rational engineering of synthetic microbial systems: from single cells to consortia. *Curr Opin Microbiol* **45**, 92-99, doi:10.1016/j.mib.2018.02.009 (2018).

6 Chen, F., Yuan, L., Ding, S., Tian, Y. & Hu, Q. N. Data-driven rational biosynthesis design: from molecules to cell factories. *Brief Bioinform*, doi:10.1093/bib/bbz065 (2019).

7 Choi, K. R. *et al.* Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends Biotechnol* **37**, 817-837, doi:10.1016/j.tibtech.2019.01.003 (2019).

8 Behler, J., Vijay, D., Hess, W. R. & Akhtar, M. K. CRISPR-Based Technologies for Metabolic Engineering in Cyanobacteria. *Trends Biotechnol* **36**, 996-1010, doi:10.1016/j.tibtech.2018.05.011 (2018).

9 Shi, T. Q., Huang, H., Kerkhoven, E. J. & Ji, X. J. Advancing metabolic engineering of Yarrowia lipolytica using the CRISPR/Cas system. *Appl Microbiol Biotechnol* **102**, 9541-9548, doi:10.1007/s00253-018-9366-x (2018).

10 Crook, N. *et al.* In vivo continuous evolution of genes and pathways in yeast. *Nat Commun* **7**, 13051, doi:10.1038/ncomms13051 (2016).

11 Liu, D. *et al.* Construction, Model-Based Analysis, and Characterization of a Promoter Library for Fine-Tuned Gene Expression in Bacillus subtilis. *ACS Synth Biol* **7**, 1785-1797, doi:10.1021/acssynbio.8b00115 (2018).

12 Tarasava, K., Liu, R., Garst, A. & Gill, R. T. Combinatorial pathway engineering using type I-E CRISPR interference. *Biotechnol Bioeng* **115**, 1878-1883, doi:10.1002/bit.26589 (2018).

13 Sciambi, A. & Abate, A. R. Accurate microfluidic sorting of droplets at 30 kHz. *Lab Chip* **15**, 47-51, doi:10.1039/c4lc01194e (2015).

14 Wang, B. L. *et al.* Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. *Nat Biotechnol* **32**, 473-478, doi:10.1038/nbt.2857 (2014).

15 Luo, T., Fan, L., Zhu, R. & Sun, D. Microfluidic Single-Cell Manipulation and Analysis: Methods and Applications. *Micromachines (Basel)* **10**, doi:10.3390/mi10020104 (2019).
16 Hindson, C. M. *et al.* Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat Methods* **10**, 1003-1005, doi:10.1038/nmeth.2633 (2013).
17 Roberts, C. H. *et al.* Development and evaluation of a next-generation digital PCR diagnostic assay for ocular Chlamydia trachomatis infections. *J Clin Microbiol* **51**, 2195-2203, doi:10.1128/JCM.00622-13 (2013).
18 Lan, F., Demaree, B., Ahmed, N. & Abate, A. R. Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. *Nat Biotechnol* **35**, 640-646, doi:10.1038/nbt.3880 (2017).
19 Habib, N. *et al.* Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods* **14**, 955-958, doi:10.1038/nmeth.4407 (2017).
20 Moon, H. S. *et al.* Inertial-ordering-assisted droplet microfluidics for high-throughput single-cell RNA-sequencing. *Lab Chip* **18**, 775-784, doi:10.1039/c7lc01284e (2018).
21 Kim, S., Masum, F. & Jeon, J. S. Recent Developments of Chip-based Phenotypic Antibiotic Susceptibility Testing. *BioChip Journal* **13**, 43-52, doi:10.1007/s13206-019-3109-7 (2019).
22 Liu, X., Zheng, W. & Jiang, X. Cell-Based Assays on Microfluidics for Drug Screening. *ACS Sens* **4**, 1465-1475, doi:10.1021/acssensors.9b00479 (2019).
23 Rothbauer, M., Zirath, H. & Ertl, P. Recent advances in microfluidic technologies for cell-to-cell interaction studies. *Lab Chip* **18**, 249-270, doi:10.1039/c7lc00815e (2018).
24 Savitskaya, J., Protzko, R. J., Li, F. Z., Arkin, A. P. & Dueber, J. E. Iterative screening methodology enables isolation of strains with improved properties for a FACS-based screen and increased L-DOPA production. *Sci Rep* **9**, 5815, doi:10.1038/s41598-019-41759-0 (2019).
25 Lee, J. J., Crook, N., Sun, J. & Alper, H. S. Improvement of lactic acid production in Saccharomyces cerevisiae by a deletion of ssb1. *J Ind Microbiol Biotechnol* **43**, 87-96, doi:10.1007/s10295-015-1713-7 (2016).
26 Markert, S. & Joeris, K. Establishment of a fully automated microtiter plate-based system for suspension cell culture and its application for enhanced process optimization. *Biotechnol Bioeng* **114**, 113-121, doi:10.1002/bit.26044 (2017).
27 Rosa, P., Tenreiro, S., Chu, V., Outeiro, T. F. & Conde, J. P. High-throughput study of alpha-synuclein expression in yeast using microfluidics for control of local cellular microenvironment. *Biomicrofluidics* **6**, 14109-141099, doi:10.1063/1.3683161 (2012).
28 Wagner, J. M. *et al.* A comparative analysis of single cell and droplet-based FACS for improving production phenotypes: Riboflavin overproduction in Yarrowia lipolytica. *Metab Eng* **47**, 346-356, doi:10.1016/j.ymben.2018.04.015 (2018).

29 Saleski, T. E. *et al.* Syntrophic co-culture amplification of production phenotype for high-throughput screening of microbial strain libraries. *Metab Eng* **54**, 232-243, doi:10.1016/j.ymben.2019.04.007 (2019).
30 Kerner, A., Park, J., Williams, A. & Lin, X. N. A programmable Escherichia coli consortium via tunable symbiosis. *PLoS One* **7**, e34032, doi:10.1371/journal.pone.0034032 (2012).
31 Zhang, X. R., J. Adaptive Evolution of Synthetic Cooperating Communities Improves Growth Performance. *PLoS One* **9** (2014).
32 Boitard, L. *et al.* Monitoring single-cell bioenergetics via the coarsening of emulsion droplets. *Proc Natl Acad Sci U S A* **109**, 7181-7186, doi:10.1073/pnas.1200894109 (2012).
33 Chung, M. T., Nunez, D., Cai, D. & Kurabayashi, K. Deterministic droplet-based co-encapsulation and pairing of microparticles via active sorting and downstream merging. *Lab Chip* **17**, 3664-3671, doi:10.1039/c7lc00745k (2017).
34 Abatemarco, J. *et al.* RNA-aptamers-in-droplets (RAPID) high-throughput screening for secretory phenotypes. *Nat Commun* **8**, 332, doi:10.1038/s41467-017-00425-7 (2017).
35 Fabienne Courtois, L. F. O., Graeme Whyte, Ashleigh B. Theberge, Wilhelm T. S. Huck, Florian Hollfelder, and Chris Abell. Controlling the Retention of Small Molecules in Emulsion Microdroplets for Use in Cell-Based Assays. *Anal Chem* **81**, 3008-3016, doi:10.1021/ac802658n (2009).
36 Chen, Y., Wijaya Gani, A. & Tang, S. K. Characterization of sensitivity and specificity in leaky droplet-based assays. *Lab Chip* **12**, 5093-5103, doi:10.1039/c2lc40624a (2012).
37 Alkayyali, T., Cameron, T., Haltli, B., Kerr, R. G. & Ahmadi, A. Microfluidic and cross-linking methods for encapsulation of living cells and bacteria - A review. *Anal Chim Acta* **1053**, 1-21, doi:10.1016/j.aca.2018.12.056 (2019).
38 Capretto, L., Mazzitelli, S., Luca, G. & Nastruzzi, C. Preparation and characterization of polysaccharidic microbeads by a microfluidic technique: application to the encapsulation of Sertoli cells. *Acta Biomater* **6**, 429-435, doi:10.1016/j.actbio.2009.08.023 (2010).
39 Håti, A. G. *et al.* Microarrays for the study of compartmentalized microorganisms in alginate microbeads and (W/O/W) double emulsions. *RSC Advances* **6**, 114830-114842, doi:10.1039/c6ra23945e (2016).
40 Zhu, X.-D. *et al.* High-throughput screening of high lactic acid-producing Bacillus coagulans by droplet microfluidic based flow cytometry with fluorescence activated cell sorting. *RSC Advances* **9**, 4507-4513, doi:10.1039/c8ra09684h (2019).
41 Tovar, M. *et al.* One Sensor for Multiple Colors: Fluorescence Analysis of Microdroplets in Microbiological Screenings by Frequency-Division Multiplexing. *Anal Chem* **91**, 3055-3061, doi:10.1021/acs.analchem.8b05451 (2019).

42 Gielen, F. *et al.* Ultrahigh-throughput-directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proc Natl Acad Sci U S A* **113**, E7383-E7389, doi:10.1073/pnas.1606927113 (2016).
43 Abate, A. R., Hung, T., Mary, P., Agresti, J. J. & Weitz, D. A. High-throughput injection with microfluidics using picoinjectors. *Proc Natl Acad Sci U S A* **107**, 19163-19166, doi:10.1073/pnas.1006888107 (2010).
44 Vallejo, D., Nikoomanzar, A., Paegel, B. M. & Chaput, J. C. Fluorescence-Activated Droplet Sorting for Single-Cell Directed Evolution. *ACS Synth Biol* **8**, 1430-1440, doi:10.1021/acssynbio.9b00103 (2019).
45 Qiao, Y. *et al.* Fluorescence-activated droplet sorting of lipolytic microorganisms using a compact optical system. *Lab Chip* **18**, 190-196, doi:10.1039/c7lc00993c (2017).
46 Sung, Y. J., Kim, J. Y. H., Choi, H. I., Kwak, H. S. & Sim, S. J. Magnetophoretic sorting of microdroplets with different microalgal cell densities for rapid isolation of fast growing strains. *Sci Rep* **7**, 10390, doi:10.1038/s41598-017-10764-6 (2017).
47 Wink, K. *et al.* An integrated chip-mass spectrometry and epifluorescence approach for online monitoring of bioactive metabolites from incubated Actinobacteria in picoliter droplets. *Anal Bioanal Chem* **410**, 7679-7687, doi:10.1007/s00216-018-1383-1 (2018).
48 Chen, J. Finding the Needle in the Haystack - the Use of Microfluidic Droplet Technology to Identify Vitamin-Secreting Lactic Acid Bacteria. *American Society for Microbiology mBio* **8**, doi:10.1128/mBio (2017).
49 Ma, F. *et al.* Substrate Engineering Enabling Fluorescence Droplet Entrapment for IVC-FACS-Based Ultrahigh-Throughput Screening. *Anal Chem* **88**, 8587-8595, doi:10.1021/acs.analchem.6b01712 (2016).
50 Fenneteau, J., Chauvin, D., Griffiths, A. D., Nizak, C. & Cossy, J. Synthesis of new hydrophilic rhodamine based enzymatic substrates compatible with droplet-based microfluidic assays. *Chem Commun (Camb)* **53**, 5437-5440, doi:10.1039/c7cc01506b (2017).
51 Huang, M. *et al.* Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc Natl Acad Sci U S A* **112**, E4689-4696, doi:10.1073/pnas.1506460112 (2015).
52 Huang, M., Joensson, H. N. & Nielsen, J. High-Throughput Microfluidics for the Screening of Yeast Libraries. *Methods Mol Biol* **1671**, 307-317, doi:10.1007/978-1-4939-7295-1_19 (2018).
53 Kim, H. S. *et al.* High-throughput droplet microfluidics screening platform for selecting fast-growing and high lipid-producing microalgae from a mutant library. *Plant Direct* **1**, e00011, doi:10.1002/pld3.11 (2017).
54 Dagkesamanskaya, A. *et al.* Use of photoswitchable fluorescent proteins for droplet-based microfluidic screening. *J Microbiol Methods* **147**, 59-65, doi:10.1016/j.mimet.2018.03.001 (2018).

55 Brechun, K. E. *et al.* Detection of Incorporation of p-Coumaric Acid into Photoactive Yellow Protein Variants in Vivo. *Biochemistry* **58**, 2682-2694, doi:10.1021/acs.biochem.9b00279 (2019).
56 Lin, J. L., Wagner, J. M. & Alper, H. S. Enabling tools for high-throughput detection of metabolites: Metabolic engineering and directed evolution applications. *Biotechnol Adv* **35**, 950-970, doi:10.1016/j.biotechadv.2017.07.005 (2017).
57 Siedler, S. *et al.* Development of a Bacterial Biosensor for Rapid Screening of Yeast p-Coumaric Acid Production. *ACS Synth Biol* **6**, 1860-1869, doi:10.1021/acssynbio.7b00009 (2017).
58 Qin, Y. *et al.* A Fluorescence-Activated Single-Droplet Dispenser for High Accuracy Single-Droplet and Single-Cell Sorting and Dispensing. *Anal Chem* **91**, 6815-6819, doi:10.1021/acs.analchem.9b01017 (2019).
59 Li, Z. Y. *et al.* Nanoliter-Scale Oil-Air-Droplet Chip-Based Single Cell Proteomic Analysis. *Anal Chem* **90**, 5430-5438, doi:10.1021/acs.analchem.8b00661 (2018).
60 Gao, D., Jin, F., Zhou, M. & Jiang, Y. Recent advances in single cell manipulation and biochemical analysis on microfluidics. *Analyst* **144**, 766-781, doi:10.1039/c8an01186a (2019).
61 Heinemann, J. *et al.* On-chip integration of droplet microfluidics and nanostructure-initiator mass spectrometry for enzyme screening. *Lab Chip* **17**, 323-331, doi:10.1039/c6lc01182a (2017).
62 McLaughlin, J. E. *et al.* A genome-wide screen in Saccharomyces cerevisiae reveals a critical role for the mitochondria in the toxicity of a trichothecene mycotoxin. *Proc Natl Acad Sci U S A* **106**, 21883-21888, doi:10.1073/pnas.0909777106 (2009).
63 Jones, G. M. *et al.* A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nature Methods* **5**, 239-241, doi:10.1038/nmeth.1181 (2008).
64 Yoshikawa, K. *et al.* Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae. *Yeast* **28**, 349-361, doi:10.1002/yea.1843 (2011).
65 Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc* **8**, 2180-2196, doi:10.1038/nprot.2013.132 (2013).
66 Ni, J., Zhang, G., Qin, L., Li, J. & Li, C. Simultaneously down-regulation of multiplex branch pathways using CRISPRi and fermentation optimization for enhancing beta-amyrin production in Saccharomyces cerevisiae. *Synth Syst Biotechnol* **4**, 79-85, doi:10.1016/j.synbio.2019.02.002 (2019).
67 Aouida, M. *et al.* Transcription activator-like effector nucleases mediated metabolic engineering for enhanced fatty acids production in Saccharomyces cerevisiae. *J Biosci Bioeng* **120**, 364-371, doi:10.1016/j.jbiosc.2015.02.017 (2015).
68 Liebermeister, W., Klipp, E., Schuster, S. & Heinrich, R. A theory of optimal differential gene expression. *Biosystems* **76**, 261-278, doi:10.1016/j.biosystems.2004.05.022 (2004).

69 Xu, N., Wei, L. & Liu, J. Recent advances in the applications of promoter engineering for the optimization of metabolite biosynthesis. *World J Microbiol Biotechnol* **35**, 33, doi:10.1007/s11274-019-2606-0 (2019).
70 Du, J., Yuan, Y., Si, T., Lian, J. & Zhao, H. Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Research* **40**, e142-e142, doi:10.1093/nar/gks549 (2012).
71 Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* **102**, 12678-12683, doi:10.1073/pnas.0504604102 (2005).
72 Lu, C. & Jeffries, T. Shuffling of Promoters for Multiple Genes To Optimize Xylose Fermentation in an Engineered Saccharomyces cerevisiae Strain. *Appl Environ Microbiol* **73**, 6072-6077, doi:doi:10.1128/AEM.00955-07 (2007).
73 Jost, M. *et al.* Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nature biotechnology* **38**, 355-364 (2020).
74 Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* **12**, 326-328, doi:10.1038/nmeth.3312 (2015).
75 Österlund, T., Nookaew, I., Bordel, S. & Nielsen, J. Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. *BMC systems biology* **7**, 1-10 (2013).
76 Deaner, M., Mejia, J. & Alper, H. S. Enabling Graded and Large-Scale Multiplex of Desired Genes Using a Dual-Mode dCas9 Activator in Saccharomyces cerevisiae. *ACS Synth Biol* **6**, 1931-1943, doi:10.1021/acssynbio.7b00163 (2017).
77 Kim, S. R., Ha, S. J., Wei, N., Oh, E. J. & Jin, Y. S. Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol. *Trends Biotechnol* **30**, 274-282, doi:10.1016/j.tibtech.2012.01.005 (2012).
78 Klein, M., Swinnen, S., Thevelein, J. M. & Nevoigt, E. Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities. *Environ Microbiol* **19**, 878-893, doi:10.1111/1462-2920.13617 (2017).
79 Giaever, G. *et al.* Functional profiling of the Saccharomyces cerevisiae genome. *nature* **418**, 387-391 (2002).
80 Bro, C., Knudsen, S., Regenberg, B., Olsson, L. & Nielsen, J. Improvement of galactose uptake in Saccharomyces cerevisiae through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* **71**, 6465-6472, doi:10.1128/AEM.71.11.6465-6472.2005 (2005).
81 Mumberg, D., Müller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119-122, doi:10.1016/0378-1119(95)00037-7 (1995).
82 Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-451, doi:10.1016/j.cell.2013.06.044 (2013).
83 Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* **164**, 1185-1197 (2016).

84 Yuan, S.-F. & Alper, H. S. Metabolic engineering of microbial cell factories for production of nutraceuticals. *Microbial cell factories* **18**, 1-11 (2019).
85 DeLoache, W. C. *et al.* An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. *Nat Chem Biol* **11**, 465-471, doi:10.1038/nchembio.1816 (2015).
86 Bowman, E. K. & Alper, H. S. Microdroplet-Assisted Screening of Biomolecule Production for Metabolic Engineering Applications. *Trends Biotechnol* **38**, 701-714, doi:10.1016/j.tibtech.2019.11.002 (2020).
87 Tang, S. Y. *et al.* Screening for enhanced triacetic acid lactone production by recombinant Escherichia coli expressing a designed triacetic acid lactone reporter. *J Am Chem Soc* **135**, 10099-10103, doi:10.1021/ja402654z (2013).
88 Rogers, J. K. *et al.* Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic acids research* **43**, 7648-7660 (2015).
89 Wu, Y., Du, G., Chen, J. & Liu, L. Genetically encoded biosensors and their applications in the development of microbial cell factories. *Engineering of Microbial Biosynthetic Pathways*, 53-73 (2020).
90 Tu, R., Li, L., Yuan, H., He, R. & Wang, Q. Biosensor-enabled droplet microfluidic system for the rapid screening of 3-dehydroshikimic acid produced in Escherichia coli. *J Ind Microbiol Biotechnol* **47**, 1155-1160, doi:10.1007/s10295-020-02316-1 (2020).
91 Markham, K. A. *et al.* Rewiring Yarrowia lipolytica toward triacetic acid lactone for materials generation. *Proc Natl Acad Sci U S A* **115**, 2096-2101, doi:10.1073/pnas.1721203115 (2018).
92 Xiong, D. *et al.* Improving key enzyme activity in phenylpropanoid pathway with a designed biosensor. *Metab Eng* **40**, 115-123, doi:10.1016/j.ymben.2017.01.006 (2017).
93 Lv, Y., Edwards, H., Zhou, J. & Xu, P. Combining 26s rDNA and the Cre-loxP System for Iterative Gene Integration and Efficient Marker Curation in Yarrowia lipolytica. *ACS Synth Biol* **8**, 568-576, doi:10.1021/acssynbio.8b00535 (2019).
94 Curran, K. A., Leavitt, J. M., Karim, A. S. & Alper, H. S. Metabolic engineering of muconic acid production in Saccharomyces cerevisiae. *Metab Eng* **15**, 55-66, doi:10.1016/j.ymben.2012.10.003 (2013).
95 Palmer, C. M., Miller, K. K., Nguyen, A. & Alper, H. S. Engineering 4-coumaroyl-CoA derived polyketide production in Yarrowia lipolytica through a β-oxidation mediated strategy. *Metab Eng* **57**, 174-181, doi:10.1016/j.ymben.2019.11.006 (2020).
96 Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33-38, doi:https://doi.org/10.1016/0378-1119(95)00685-0 (1996).
97 Johnston, T. G. *et al.* Compartmentalized microbes and co-cultures in hydrogels for on-demand bioproduction and preservation. *Nature Communications* **11**, 563, doi:10.1038/s41467-020-14371-4 (2020).

98 Wagner, J. M., Williams, E. V. & Alper, H. S. Developing a piggyBac Transposon System and Compatible Selection Markers for Insertional Mutagenesis and Genome Engineering in Yarrowia lipolytica. *Biotechnol J* **13**, e1800022, doi:10.1002/biot.201800022 (2018).
99 Mazutis, L. *et al.* Single-cell analysis and sorting using droplet-based microfluidics. *Nat Protoc* **8**, 870-891, doi:10.1038/nprot.2013.046 (2013).
100 Cardenas, J. & Da Silva, N. A. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. *Metabolic engineering* **25**, 194-203 (2014).
101 Liu, H., Marsafari, M., Deng, L. & Xu, P. Understanding lipogenesis by dynamically profiling transcriptional activity of lipogenic promoters in Yarrowia lipolytica. *Applied microbiology and biotechnology* **103**, 3167-3179 (2019).
102 Bowman, E. K. *et al.* Bidirectional titration of yeast gene expression using a pooled CRISPR guide RNA approach. *Proc Natl Acad Sci U S A* **117**, 18424-18430, doi:10.1073/pnas.2007413117 (2020).
103 Nadjar, A., Gerfen, C. R. & Bezard, E. Priming for l-dopa-induced dyskinesia in Parkinson's disease: a feature inherent to the treatment or the disease? *Progress in neurobiology* **87**, 1-9 (2009).
104 Alper, H., Miyaoku, K. & Stephanopoulos, G. Construction of lycopene-overproducing E. coli strains by combining systematic and combinatorial gene knockout targets. *Nature Biotechnology* **23**, 612-616, doi:10.1038/nbt1083 (2005).
105 Fordjour, E., Adipah, F. K., Zhou, S., Du, G. & Zhou, J. Metabolic engineering of Escherichia coli BL21 (DE3) for de novo production of L-DOPA from D-glucose. *Microbial Cell Factories* **18**, 1-10 (2019).
106 Liu, Y.-G. & Whittier, R. F. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* **25**, 674-681, doi:https://doi.org/10.1016/0888-7543(95)80010-J (1995).
107 Shimada, T., Bridier, A., Briandet, R. & Ishihama, A. Novel roles of LeuO in transcription regulation of E. coli genome: antagonistic interplay with the universal silencer H-NS. *Molecular microbiology* **82**, 378-397 (2011).
108 Badarinarayana, V. *et al.* Selection analyses of insertional mutants using subgenic-resolution arrays. *Nature Biotechnology* **19**, 1060-1065, doi:10.1038/nbt1101-1060 (2001).
109 Liu, L., Pan, A., Spofford, C., Zhou, N. & Alper, H. S. An evolutionary metabolic engineering approach for enhancing lipogenesis in Yarrowia lipolytica. *Metab Eng* **29**, 36-45, doi:10.1016/j.ymben.2015.02.003 (2015).
110 Winston, F. EMS and UV mutagenesis in yeast. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **Chapter 13**, Unit 13.13B, doi:10.1002/0471142727.mb1303bs82 (2008).
111 Liu, Y. & Nielsen, J. Recent trends in metabolic engineering of microbial chemical factories. *Curr Opin Biotechnol* **60**, 188-197, doi:10.1016/j.copbio.2019.05.010 (2019).

112 Guo, M. T., Rotem, A., Heyman, J. A. & Weitz, D. A. Droplet microfluidics for high-throughput biological assays. *Lab Chip* **12**, 2146-2155, doi:10.1039/c2lc21147e (2012).
113 Rienzo, M. *et al.* High-throughput screening for high-efficiency small-molecule biosynthesis. *Metab Eng* **63**, 102-125, doi:10.1016/j.ymben.2020.09.004 (2021).
114 Bowman, E. K. *et al.* Sorting for secreted molecule production using a biosensor-in-microdroplet approach. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2106818118 (2021).
115 Li, D., Liu, L., Qin, Z., Yu, S. & Zhou, J. Combined evolutionary and metabolic engineering improve 2-keto-L-gulonic acid production in Gluconobacter oxydans WSH-004. *Bioresour Technol* **354**, 127107, doi:10.1016/j.biortech.2022.127107 (2022).
116 Liu, W. W., Zhu, Y. & Fang, Q. Femtomole-Scale High-Throughput Screening of Protein Ligands with Droplet-Based Thermal Shift Assay. *Anal Chem* **89**, 6678-6685, doi:10.1021/acs.analchem.7b00899 (2017).
117 Holland-Moritz, D. A. *et al.* Mass Activated Droplet Sorting (MADS) Enables High-Throughput Screening of Enzymatic Reactions at Nanoliter Scale. *Angew Chem Int Ed Engl* **59**, 4470-4477, doi:10.1002/anie.201913203 (2020).
118 Bjork, S. M., Sjostrom, S. L., Andersson-Svahn, H. & Joensson, H. N. Metabolite profiling of microfluidic cell culture conditions for droplet based screening. *Biomicrofluidics* **9**, 044128, doi:10.1063/1.4929520 (2015).
119 van Tatenhove-Pel, R. J. *et al.* Serial propagation in water-in-oil emulsions selects for Saccharomyces cerevisiae strains with a reduced cell size or an increased biomass yield on glucose. *Metab Eng* **64**, 1-14, doi:10.1016/j.ymben.2020.12.005 (2021).
120 Bachmann, H., Pronk, J. T., Kleerebezem, M. & Teusink, B. Evolutionary engineering to enhance starter culture performance in food fermentations. *Curr Opin Biotechnol* **32**, 1-7, doi:10.1016/j.copbio.2014.09.003 (2015).
121 Blazeck, J. *et al.* Metabolic engineering of Yarrowia lipolytica for itaconic acid production. *Metab Eng* **32**, 66-73, doi:10.1016/j.ymben.2015.09.005 (2015).
122 Hanko, E. K. R., Minton, N. P. & Malys, N. A Transcription Factor-Based Biosensor for Detection of Itaconic Acid. *ACS Synth Biol* **7**, 1436-1446, doi:10.1021/acssynbio.8b00057 (2018).
123 Kalyani, D. *et al.* A Highly Efficient Recombinant Laccase from the Yeast Yarrowia lipolytica and Its Application in the Hydrolysis of Biomass. *PLOS ONE* **10**, e0120156, doi:10.1371/journal.pone.0120156 (2015).
124 Fonager, J. *et al.* Development of the piggyBac transposable system for Plasmodium berghei and its application for random mutagenesis in malaria parasites. *BMC Genomics* **12**, 155, doi:10.1186/1471-2164-12-155 (2011).
125 Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457-462, doi:10.1093/nar/gkv1070 (2016).

126 Burmeister, A. & Grünberger, A. Microfluidic cultivation and analysis tools for interaction studies of microbial co-cultures. *Current Opinion in Biotechnology* **62**, 106-115, doi:https://doi.org/10.1016/j.copbio.2019.09.001 (2020).
127 Najah, M. *et al.* Droplet-based microfluidics platform for ultra-high-throughput bioprospecting of cellulolytic microorganisms. *Chem Biol* **21**, 1722-1732, doi:10.1016/j.chembiol.2014.10.020 (2014).
128 Partipilo, G., Graham, A. J., Belardi, B. & Keitz, B. K. Extracellular Electron Transfer Enables Cellular Control of Cu(I)-Catalyzed Alkyne-Azide Cycloaddition. *ACS Cent Sci* **8**, 246-257, doi:10.1021/acscentsci.1c01208 (2022).
129 Lovley, D. R. & Holmes, D. E. Electromicrobiology: the ecophysiology of phylogenetically diverse electroactive microorganisms. *Nat Rev Microbiol* **20**, 5-19, doi:10.1038/s41579-021-00597-6 (2022).
130 Hartshorne, R. S. *et al.* Characterization of an electron conduit between bacteria and the extracellular environment. *Proc Natl Acad Sci U S A* **106**, 22169-22174, doi:10.1073/pnas.0900086106 (2009).
131 Koch, C. & Harnisch, F. Is there a Specific Ecological Niche for Electroactive Microorganisms? *ChemElectroChem* **3**, 1282-1295, doi:10.1002/celc.201600079 (2016).
132 Kiely, P. D., Regan, J. M. & Logan, B. E. The electric picnic: synergistic requirements for exoelectrogenic microbial communities. *Curr Opin Biotechnol* **22**, 378-385, doi:10.1016/j.copbio.2011.03.003 (2011).
133 Logan, B. E. Exoelectrogenic bacteria that power microbial fuel cells. *Nature Reviews Microbiology* **7**, 375-381, doi:10.1038/nrmicro2113 (2009).
134 Shieh, P. *et al.* CalFluors: A Universal Motif for Fluorogenic Azide Probes across the Visible Spectrum. *Journal of the American Chemical Society* **137**, 7145-7151, doi:10.1021/jacs.5b02383 (2015).
135 Ross, D. E. *et al.* Characterization of Protein-Protein Interactions Involved in Iron Reduction by Shewanella oneidensis MR-1. *Appl Environ Microbiol* **73**, 5797-5808, doi:doi:10.1128/AEM.00146-07 (2007).
136 Beblawy, S. *et al.* Extracellular reduction of solid electron acceptors by Shewanella oneidensis. *Mol Microbiol* **109**, 571-583, doi:10.1111/mmi.14067 (2018).
137 Dedrick, S. *et al.* Impact of Temporal pH Fluctuations on the Coexistence of Nasal Bacteria in an in silico Community. *Frontiers in Microbiology* **12**, doi:10.3389/fmicb.2021.613109 (2021).
138 Qiao, Y. *et al.* Fluorescence-activated droplet sorting of PET degrading microorganisms. *J Hazard Mater* **424**, 127417, doi:10.1016/j.jhazmat.2021.127417 (2022).
139 Su, X. *et al.* Directed Evolution of Laccase for Improved Thermal Stability Facilitated by Droplet-Based Microfluidic Screening System. *Journal of Agricultural and Food Chemistry* **70**, 13700-13708, doi:10.1021/acs.jafc.2c05048 (2022).
140 Beneyton, T., Coldren, F., Baret, J.-C., Griffiths, A. D. & Taly, V. CotA laccase: high-throughput manipulation and analysis of recombinant enzyme libraries

expressed in E. coli using droplet-based microfluidics. *Analyst* **139**, 3314-3323, doi:10.1039/C4AN00228H (2014).

141 Beneyton, T. *et al.* Droplet-based microfluidic high-throughput screening of heterologous enzymes secreted by the yeast Yarrowia lipolytica. *Microbial Cell Factories* **16**, 18, doi:10.1186/s12934-017-0629-5 (2017).

142 Fan, G., Graham, A. J., Kolli, J., Lynd, N. A. & Keitz, B. K. Aerobic radical polymerization mediated by microbial metabolism. *Nature Chemistry* **12**, 638-646, doi:10.1038/s41557-020-0460-1 (2020).

143 Raymaekers, K., Ponet, L., Holtappels, D., Berckmans, B. & Cammue, B. P. A. Screening for novel biocontrol agents applicable in plant disease management – A review. *Biological Control* **144**, 104240, doi:https://doi.org/10.1016/j.biocontrol.2020.104240 (2020).

144 Vontas, J., Katsavou, E. & Mavridis, K. Cytochrome P450-based metabolic insecticide resistance in Anopheles and Aedes mosquito vectors: Muddying the waters. *Pesticide Biochemistry and Physiology* **170**, 104666, doi:https://doi.org/10.1016/j.pestbp.2020.104666 (2020).

145 Li, R. *et al.* Development of S-Fluxametamide for Bioactivity Improvement and Risk Reduction: Systemic Evaluation of the Novel Insecticide Fluxametamide at the Enantiomeric Level. *Environmental Science & Technology* **53**, 13657-13665, doi:10.1021/acs.est.9b03697 (2019).

146 Gressel, J. Perspective: present pesticide discovery paradigms promote the evolution of resistance – learn from nature and prioritize multi-target site inhibitor design. *Pest Management Science* **76**, 421-425, doi:https://doi.org/10.1002/ps.5649 (2020).

147 Antona, S., Platzman, I. & Spatz, J. P. Droplet-Based Cytotoxicity Assay: Implementation of Time-Efficient Screening of Antitumor Activity of Natural Killer Cells. *ACS Omega* **5**, 24674-24683, doi:10.1021/acsomega.0c03264 (2020).

148 Sarkar, S. *et al.* Dynamic Analysis of Human Natural Killer Cell Response at Single-Cell Resolution in B-Cell Non-Hodgkin Lymphoma. *Front Immunol* **8**, 1736, doi:10.3389/fimmu.2017.01736 (2017).

149 Joensson, H. N. & Andersson Svahn, H. Droplet Microfluidics—A Tool for Single-Cell Analysis. *Angewandte Chemie International Edition* **51**, 12176-12192, doi:https://doi.org/10.1002/anie.201200460 (2012).

150 Kamiya, K. *et al.* Well-Controlled Cell-Trapping Systems for Investigating Heterogeneous Cell–Cell Interactions. *Advanced Healthcare Materials* **7**, 1701208, doi:https://doi.org/10.1002/adhm.201701208 (2018).

151 Choque-Guevara, R. *et al.* A recombinant SARS-CoV-2 RBD antigen expressed in insect cells elicits immunogenicity and confirms safety in animal models. *bioRxiv*, 2021.2011.2026.470043, doi:10.1101/2021.11.26.470043 (2021).

152 Yamaguchi, S., Ueno, A., Akiyama, Y. & Morishima, K. Cell patterning through inkjet printing of one cell per droplet*. *Biofabrication* **4**, 045005, doi:10.1088/1758-5082/4/4/045005 (2012).

153 Holstein, J. M., Gylstorff, C. & Hollfelder, F. Cell-free Directed Evolution of a Protease in Microdroplets at Ultrahigh Throughput. *ACS Synth Biol* **10**, 252-257, doi:10.1021/acssynbio.0c00538 (2021).