Harnessing the Power of Synthetic Microbial Consortia: New Approaches for Microbial Engineering in Biochemical Production

Tatyana E. Saleski

A dissertation submitted in partial fulllment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Engineering)
in the University of Michigan
2019

Doctoral Committee:

Professor Xiaoxia Nina Lin, Chair
Professor E. Neil G. Marsh
Professor Henry Y. Wang
Professor Fei Wen

Tatyana E. Saleski

tsaleski@umich.edu

ORCID iD: 0000-0001-8456-7781

Acknowledgments

First and foremost, I would like to thank my advisor, Prof. Nina Lin, for her mentorship and guidance throughout my time in her lab. I admire her enthusiasm to attack problems from new angles, her talent for communicating ideas in clear and compelling ways, as well as her thoughtful leadership style and ability to foster an inclusive and collaborative research environment. I have learned a great deal from her feedback and suggestions and I appreciate the many opportunities she has provided me to explore different areas of the projects on which I have worked and to develop new skills and perspectives.

I would also like to express my gratitude to the rest of my dissertation committee, Profs. Fei Wen, Henry Wang, and Neil Marsh, for providing valuable feedback on my research progress over the past several years. Their insights have significantly influenced the overall directions of my work.

I would like to thank the past and present members of Nina's lab for making it a great place to work, especially Dr. Jeremy Minty, Prof. Fengming Lin, Dr. Alissa Kerner, Dr. Mike Nelson, Dr. Steven Wang, Dr. Scott Scholz, Corine Jackman, David Carruthers, James Tan, Adam Krieger, Li Yuan, Dr. Chun Wan, Prof. Chang Kyu Byun, Luqman Mahir, Mathieu Rossion, and Amin Mohajeri. My work builds on that of Jeremy and Alissa and I am grateful to them for starting these interesting projects. I also appreciate Fengming and Mike helping me get started when I joined the lab, and Scott and Corine for their support and many helpful discussions throughout our time in Nina's lab. I am particularly grateful to Corine for collaborating on the microdroplet co-cultivation assay and teaching me to perform these experiments. I would also like to thank Li and Adam for making valuable

contributions to the integration project in its early stages, as well as Steven and James for helpful discussions regarding the microdroplet cultivation and sequencing data analysis.

I have also had the pleasure of working with several talented undergraduates who assisted with the work presented here. I would particularly like to thank Azzaya Khasbaatar, who worked on the cross-feeding screening project, Xiangyu Zhao, who worked on the isobutanol pathway integration, and Sarah Hammer, who worked on the adaptive evolution of *T. reesei/E. coli*. Each of them performed a great deal of high-quality work on these projects and I appreciate their enthusiasm and dedication.

I am also especially grateful to my collaborator, Meng Ting Chung, in Prof. Katsuo Kurabayashi's lab, who developed the microdroplet sorting assay. He performed many experiments, including design and fabrication of the microfluidic devices, and also taught me to operate the sorting system. I appreciate his patience and persistence since this assay format required more troubleshooting than we anticipated.

I would like to thank Profs. Katsuo Kurabayashi, Mark Burns, Jinsang Kim, Lola Eniola, Fei Wen, Erdogan Gulari, Allen Liu, and Dr. Chris Barr, each of whom generously allowed me to use equipment in their labs for various aspects of the work presented here, as well as Brian Johnson for help with the clean room and for 3D printing Petri dish holders for the plate scanning experiments.

I am also thankful to my friends in other labs, particularly Dr. Hui Gao, Dr. Lavinia Li, Christine Yee, Eshita Khera, Dr. Sarah Mena, Dr. Megan Dunn, Dr. Saadet Albayrak, Eric Holt, Dr. Wen-Chi Lin, Martin de Beer, Alyse Krausz, and Alison Banka. I have really enjoyed and benefited from their comradery and many interesting discussions. I would also like to thank Holly Chung and Divya Sanghi for being wonderful and incredibly supportive roommates and friends.

I would like to thank the Chemical Engineering staff members, especially Susan Hamlin, Kelly Raickovich, Shelley Fellers, Barbara Perry, Michael Africa, Pamela Bogdanski, Connie Bacus, and Rhonda Jent, for all their help on various aspects.

Outside of the University of Michigan, several labs kindly shared strains and plasmids that were essential for the work presented here. Prof. James Liao (UCLA) provided isobutanol production strains and plasmids, Prof. Keith Tyo (Northwestern U) shared the chemically inducible chromosomal evolution plasmid, and Dr. Donald Court (NIH) sent recombineering plasmids.

I would also like to extend my gratitude to my research mentors when I was an undergraduate: Dr. Sabine Brinkmann-Chen, Dr. Xiang Liu, and Prof. Frances Arnold, who introduced me to research. I admire their creativity and passion for their work and my experience working with them inspired me to wish to pursue a career in research. I am particularly grateful to Sabine, who trained me to work in the lab and is a very patient and thorough teacher.

Finally, I am sincerely grateful to my parents. I have always admired their integrity, generosity, and work ethics. I would like to thank them for all their love and support, for making my education a top priority, and for always encouraging me.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

Adh	Alcohol dehydrogenase
AFEX	Ammonia fiber expansion
amp	Ampicillin
BGL	β-glucosidase
CBP	Consolidated bioprocessing
CGSC	*E. coli* genetic stock center
CIChE	Chemically inducible chromosomal evolution
cm	Chloramphenicol
COMPACTER	Customized optimization of metabolic pathways by combinatorial transcriptional engineering
CRISPR	Clustered regularly interspaced short palindromic repeats
IPTG	Isopropyl β-D-1-thiogalactopyranoside
FADS	Fluorescence-activated droplet sorting
FBA	Flux balance analysis
gDNA	Genomic DNA
HFE	Hydrofluoroether
HMG-CoA	3-hydroxy-3-methyl-glutaryl-coenzyme A
HPLC	High-performance liquid chromatography
IME	Inverse metabolic engineering
kan	Kanamycin
Kdc	Alpha-ketoisovalerate decarboxylase
2-KIV	2-Ketoisovalerate
LB	Lysogeny broth (Lennox formulation)
M9IPG	M9 isobutanol production medium with glucose
MAGE	Multiplex automated genome engineering
MCC	Microcrystalline cellulose
NTG	N-Methyl-N'-nitro-N-nitrosoguanidine
PDMS	Polydimethylsiloxane
pNPG	4-Nitrophenyl β-D-glucopyranoside
qPCR	Quantitative polymerase chain reaction
SHF	Separate hydrolysis and fermentation
SnoCAP	Syntrophic co-culture amplification of production phenotype
SSCF	Simultaneous saccharification and co-fermentation
SSF	Simultaneous saccharification and fermentation
TCA cycle	Tricarboxylic acid cycle
tet	Tetracycline
TIGR	Tunable intergenic region
TMM	*Trichoderma* minimal medium (0.1 M maleate-NaOH, pH 6)
X-gal	5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Abstract

Microbial biosynthesis is a sustainable and high-specificity means of producing various bioproducts, including pharmaceuticals, specialty and commodity chemicals, and biofuels. Due to the complexity of microorganisms, it is frequently difficult to rationally engineer them, which necessitates iterative rounds of design, construction, and testing to generate high-producing strains. Furthermore, it can be difficult to optimize multiple functions in the same microorganism. Microbial consortia are abundant in natural environments and can offer unique properties that are not attainable by monocultures. Design principles have begun to be developed for synthetic consortia and further maturation of this field will lead to many exciting new opportunities in microbial bioprocessing. In this dissertation, we describe two cases of utilizing microbial consortia, one as a tool for screening microbial libraries, and the other as a division-of-labor approach for accomplishing the complex task of lignocellulosic biofuel production.

First, we demonstrate that a cross-feeding metabolic circuit can convert production phenotypes into growth phenotypes, which are highly screenable. This technology, which we term Syntrophic Co-culture Amplification of Production phenotype (SnoCAP), has two valuable properties that are not present in monocultures: (1) it has a highly tunable dynamic range, and (2) it amplifies small differences between strains. We implemented three different compartmentalization schemes of increasing throughput capability: microplates (10^2-10^4 strains evaluated/experiment), agar plates (10^4-10^5 strains evaluated/experiment), and microdroplets (10^5-10^7 strains evaluated/experiment). We demonstrated SnoCAP's ability to differentiate between *Escherichia coli* strains of differing production levels for 2-ketoisovalerate (2-KIV), a precursor of the drop-in biofuel isobutanol, and L-tryptophan, a

precursor for several pharmaceutically active compounds. We then used SnoCAP to screen a chemically mutagenized library and identify an efficient isobutanol production strain that reaches a 5-fold higher titer than the parent strain. We expect SnoCAP can be applied to the screening of a wide variety of target molecules for which high-throughput screening assays do not currently exist.

Second, we examine a previously developed co-culture of the cellulolytic fungus *Trichoderma reesei* and isobutanol-producing *E. coli* for consolidated bioprocessing of lignocellulosic biomass to biofuel. This approach provides division-of-labor, distributing the metabolic burden and allowing optimization of hydrolysis and fermentation separately. We work toward improving this co-culture by engineering the *E. coli* strains for improved performance under co-culture conditions. Due to observed issues with plasmid loss, we developed strains with the isobutanol pathway integrated into the genome. We used the chemically inducible chromosomal evolution (CIChE) method to achieve high copy number of the genes responsible for the conversion of 2-KIV to isobutanol. We then explored the use of position-dependent expression variation, in conjunction with SnoCAP screening, to optimize expression of another gene crucial for the synthesis of 2-KIV. Additionally, we developed a framework for adaptive evolution of the *T. reesei/E. coli* co-culture. We expect that this method may be used on a strain with the isobutanol pathway integrated into the genome to select for variants that are well-suited to production under co-culture conditions.

In summary, this work contributes to the development of synthetic microbial consortia for biochemical production. We have demonstrated that the properties of cross-feeding metabolic circuits can be exploited as a useful high-throughput screening tool. We have also explored a synthetic fungal-bacterial consortium that divides the labor of lignocellulosic biomass conversion between two specialist strains and developed new approaches to optimize the fermentation specialist for the unusual conditions it encounters in the co-culture.

Chapter 1: Background and motivation

1.1 Strategies for high-throughput metabolite detection

1.1.1 A brief overview of industrial microbiology and the need for phenotype screening techniques

Microbial biosynthesis offers an attractive approach to producing many chemicals. Compared to chemical synthesis, bioproduction generally does not require harsh chemicals or process conditions, making it safer and more environmentally friendly. Microbial processes also have a superior capacity for stereospecificity and regiospecificity compared to synthetic chemistry. Fermentation of food products has been practiced since prehistoric times. Scientific understanding of these processes began to emerge in the 1850s with Louis Pasteur's studies of yeast fermentation. Subsequent study of yeast led to the discovery of enzymes and multi-step pathways. New chemical demands caused by World War I led to efforts in industrial strain engineering. In these early days of industrial microbiology, strain development generally proceeded by starting with strains that could produce small amounts of a target molecule and subjecting them to iterative rounds random mutagenesis and screening or selection for higher-producing strains. This approach led to notable successes, including the production of penicillin by *Penicillium chrysogenum*. Rounds of mutagenesis and screening increased production from titers of less than 1 g/L to more than 20 g/L [1]. While effective, a drawback to such brute force screening approaches is their labor-intensive nature. An alternative is to use selections that enrich for the target population. In the 1950s and 60s, successes in constructing amino acid overproducers began to be reported by selecting

for *Corynebacterium* strains that were resistant to toxic amino acid analogs [2].

With the advent of cloning techniques, recombinant DNA technology and improved understanding of biochemistry, new strategies of rational design became possible, leading to the emergence of the field of metabolic engineering in the 1990s. Today, we use metabolic engineering to produce an ever-expanding collection of useful, value-added chemicals, including pharmaceuticals, specialty and commodity chemicals, and biofuels. The three major components of metabolic engineering are pathway design, construction, and optimization. Optimization of titer, rate, and yield is essential since the ultimate goal is an industrially viable strain [3]. If the product is not continuously removed during production then the strain's tolerance to the product may also require improvement since the cells need to be able to survive and continue to function at high concentrations of the produced compound. Even molecules that are essential at low concentrations (e.g., amino acids) can become inhibitory at high concentrations.

Combinatorial methods have proved effective for achieving the optimization step. Strategies for combinatorial genetic diversity generation are reviewed in [4]. Some examples include genome shuffling, gene knockout libraries, overexpression libraries, and global transcription machinery engineering. For fine-tuning gene expression, techniques such as multiplex automated genome engineering (MAGE) [5], customized optimization of metabolic pathways by combinatorial transcriptional engineering (COMPACTER) [6], and tunable intergenic regions (TIGRs) [7] enable construction of combinations of different expression levels of a pathway's genes. Libraries of great diversity can be created using these combinatorial methods and it is often important to be able to screen as many members as possible.

An illustrative example of successful metabolic engineering is the development of microbial production of the antimalarial drug precursor artemisinic acid by the Keasling group at UC Berkeley ([8] provides a summary of the work). While introduction of the pathway initially led to low titers and cell viability issues, intensive optimization efforts produced an *Saccharomyces cerevisiae* strain capable of producing 25 g/L artemisinic acid. Production

was further optimized by Amyris and industrial-scale production of this product is now underway by Sanofi. During one stage of the work, the yeast mevalonate pathway was expressed in *Escherichia coli*. Low titers were assumed to be due to a pathway imbalance leading to the buildup of an inhibitory intermediate [9]. The TIGR strategy, which generates combinations of post-transcriptional control elements, was developed and utilized to modulate the expression levels of three pathway genes combinatorially. It was found that a variant with reduced expression of two of the enzymes achieved seven-fold higher mevalonate production [7]. Further analysis of the strains revealed that the intermediate 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG-CoA) had been the inhibitory compound. Subsequently the gene for HMG-CoA reductase was overexpressed, leading to increased titers. Thus the combinatorial method led to insights that could not have been predicted purely by the rational design methods available.

In addition to pharmaceuticals, biofuels are another area of interest for metabolic engineering efforts. In 2008, the Liao group at UCLA reported production of various higher alcohols by *E. coli* through their amino acid biosynthesis pathways by expression of an alpha-ketoisovalerate decarboxylase (Kdc) and an alcohol dehydrogenase (Adh) and overexpressing amino acid biosynthesis genes (Fig. 1.1A, [10]). Isobutanol was produced at high yield and titer, making it a promising biofuel candidate. To optimize isobutanol production, two parallel approaches were taken. One was a rationally guided approach to delete competing pathways (Fig. 1.1B, [10]). The other was an evolutionary approach, involving chemical mutagenesis followed by growth on a toxic analog of leucine/valine to select for high flux through this pathway, followed by overexpression of the pathway genes, and screening for isobutanol production [11]. Both strategies proved highly effective. Further work has been done to optimize the process, including *in situ* isobutanol removal that enabled production of >50 g/L isobutanol in 72 h [12] and resolution of a cofactor imbalance that led to anaerobic isobutanol production at 100% theoretical yield [13]. Industrial-scale microbial production of isobutanol has begun by Gevo and is planned by Butamax.

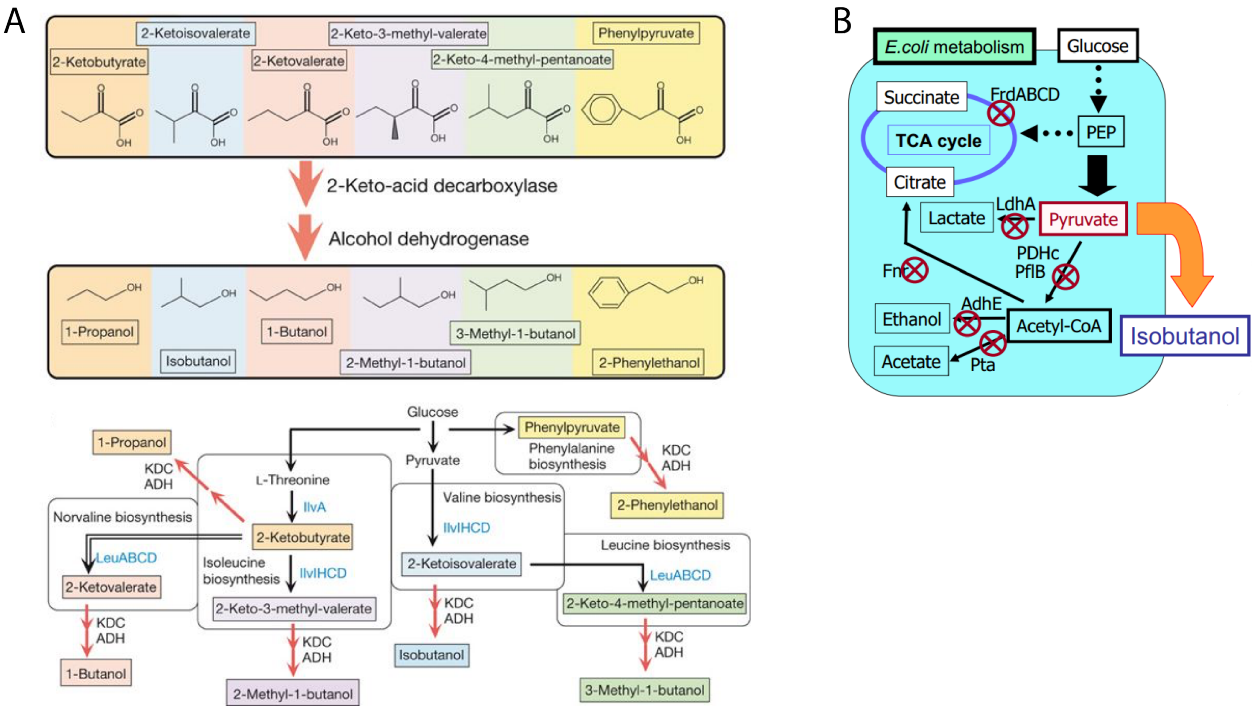

Figure 1.1: Non-fermentative alcohol production in *E. coli*. (A) Various higher alcohols that can be produced in *E. coli*. Reproduced from [10]. (B) Strategy for increasing isobutanol production in strain JCL260 by deletion of competing pathways [10]. Reproduced from: http://www.arpa-e.energy.gov/

Once a strain of high production-level is attained, either by random mutagenesis or by more rational methods, it is often of interest to identify the specific mutations that have caused the desired phenotype. Separating the causative mutations from silent mutations can improve understanding and aid future strain engineering efforts. Inverse metabolic engineering (IME), introduced by Bailey *et al.* in 1996 [14], is a widely used approach to isolate the causative mutations. Here genome fragments can be cloned into plasmids or mutations can be reintroduced into the genome individually. Mutations can then be tested for their individual contributions. An IME step also requires screening libraries and it is most effective if the entire library can be assessed. As the complexity of the manipulated system increases, so does the genotype/phenotype space that one wishes to explore (Fig. 1.2A). Construction of vast libraries to explore network, and genome space is possible, but they are only practically useful if they can be screened efficiently.

Due to the diversity of small molecule properties, methods for their detection are much less generalizable than genetic diversity generation and frequently are the bottleneck in strain engineering efforts. Molecules with colorimetric or fluorescent properties can enable visual evaluation or fluorescence-activated cell sorting (FACS), thus reaching significantly higher throughputs (Fig. 1.2B). For example, the production levels of antioxidants lycopene and β-carotene can be assessed by the shade of a bacterial colony on an agar plate. Molecules that can produce optical readouts can also be screened using droplet microfluidics platforms, which enables high-throughput screening for extracellular production levels. Most target molecules, however, are inconspicuous and the standard way to assess them is chromatography- or mass-spectrometry-based assays, which at best evaluate only hundreds to thousands of variants per day. High-throughput mass spectrometry with microdroplets has been proposed as a generalizable solution but has yet to be fully realized for strain engineering purposes.

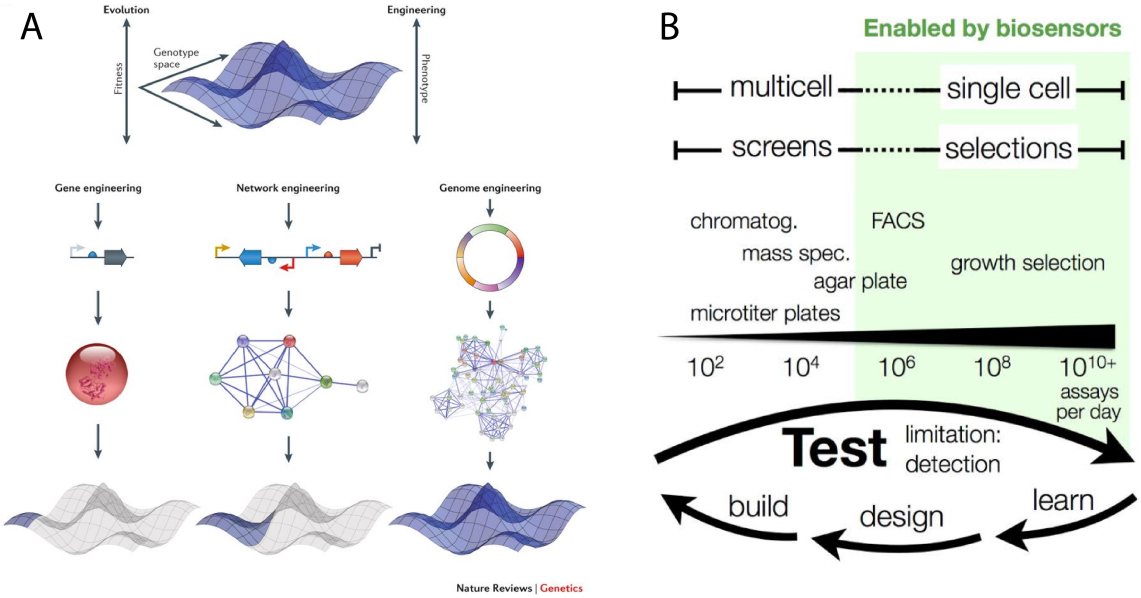

Figure 1.2: Strain engineering and screening. (A) Genotype/phenotype space explored by gene, network and genome engineering. Reproduced from [15]. (B) Throughputs of various metabolite detection schemes. Reproduced from [16].

1.1.2 Biosensors for high-throughput screening

Biosensors that can transduce inconspicuous molecule concentrations into conspicuous outputs such as fluorescence or growth advantages are an appealing approach to increasing screening throughputs, leading to significant research efforts in this area. Biosensor strategies are reviewed in [4, 16, 17]. Biosensors are often developed in the context of environmental contamination detection but most of these approaches are theoretically adaptable to strain engineering, although they may require modification of the assay dynamic range.

One approach is to convert the target molecule into a fluorescent compound via enzymatic and/or chemical reactions or to use chromogenic or fluorescent dyes to make the target molecule easily detectable. For example, Santos & Stephanopoulos developed a colorimetric assay for tyrosine production by *E. coli* by expressing a tyrosinase that converts tyrosine to the black pigment melanin [18].

Other approaches utilize biosensing machinery that responds to the presence of the target molecule in a dose-dependent manner. This machinery may consist of proteins, nucleic acids, or whole cells and may be located within the production strain itself, or outside of the cell. Protein strategies include transcription factors (either native or engineered), G-protein-coupled receptors, and fluorescent proteins with added ligand binding properties. Binding of the target molecule triggers the protein to change conformation and produce a response, either downstream or in the protein itself, in the case of fluorescent proteins. For example, a transcription factor that is responsive to a target molecule can be configured such that it regulates a promoter controlling a fluorescent protein's expression. These protein-based methods are highly useful but generally require a dedicated protein engineering effort for each target molecule in order to achieve specificity and desirable dynamic range. They also require an existing ligand-binding domain that already has some binding capability for the target molecule, or at least a similarly structured molecule, from which to begin directed evolution and/or computational design efforts.

Nucleic acid strategies have also been explored. Properties of nucleic acids are much

easier to predict than those of proteins, and there has been much work into small molecule aptamer design and directed evolution [19]. The relative ease of producing nucleic acids with desirable binding properties makes them promising as generalizable biosensor strategies. Riboswitches have been designed that produce downstream transcriptional responses upon binding the molecule of interest and have been used as screens and selections in enzyme directed evolution endeavors. RNA Spinach aptamers is another theoretically generalizable strategy, in which the ligand-binding causes the aptamer to bind a dye and produce a fluorescent output [20].

For the whole-cell approach, whole-cell auxotrophic biosensors are valuable tools for accurately determining concentrations of metabolites in biological samples [21–27]; however, they have not been widely employed for high-throughput strain screening endeavors. This may be due to their typically small dynamic ranges, at very low concentrations of the focal molecule (Fig. 1.3). A notable exception is an engineered mevalonate auxotroph that was used for screening libraries for mevalonate production [28]. In this work, the producer strain's supernatant was diluted so that it fell into the dynamic range of the biosensor.

In the work presented in this dissertation, we make use of cross-feeding co-cultures of microorganisms as a biosensing tool. Syntrophic co-cultures, consisting of auxotrophic strains that cross-feed their required metabolites, enabling co-growth, are prevalent in nature [29,30] and have been used historically by microbiologists as a tool to interrogate biochemical pathways [31,32] and for assessment of whether a specific metabolite is produced by a strain of interest [33]. Wintermute & Silver [34] and Mee *et al.* [35] have characterized co-growth of a variety of synthetic combinations of auxotrophic *E. coli* strains. While the strains have not been engineered to overproduce the amino acids that must be cross-fed, many of the pairs can support co-growth (Fig. 1.4).

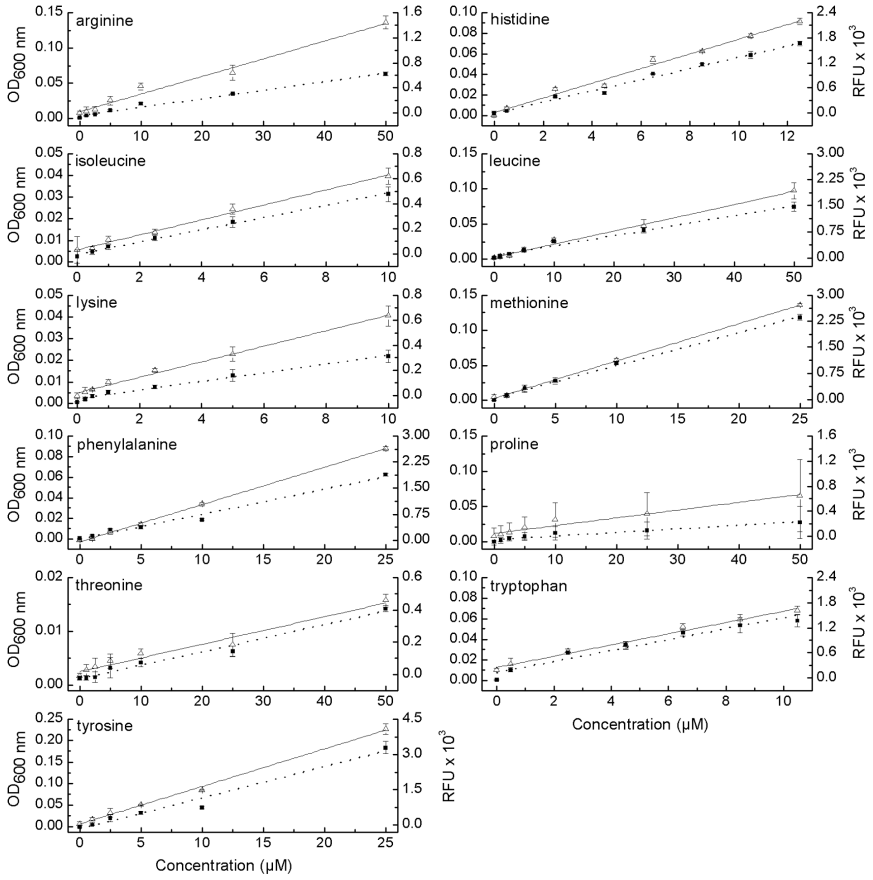

Figure 1.3: Characterization of various amino acid auxotrophic strains. Reproduced from [25].

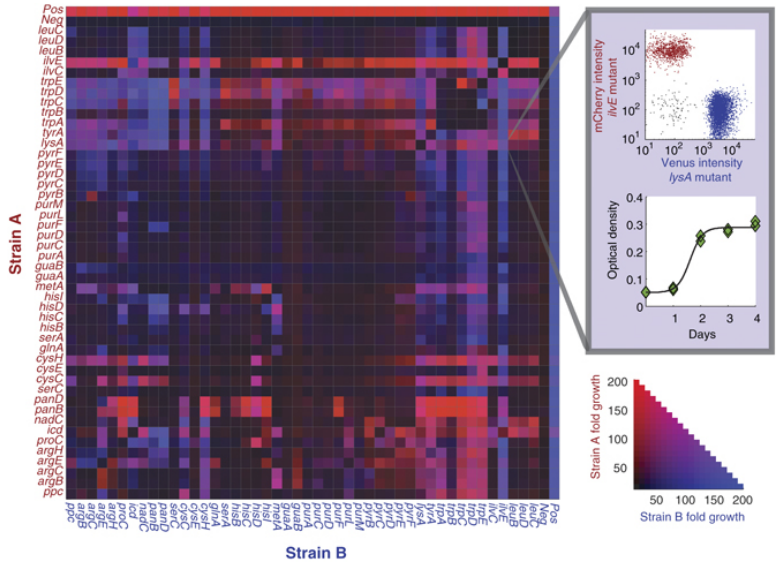

Figure 1.4: Growth of various combinations of auxotrophic strains. Reproduced from [34].

1.2 Approaches for consolidated bioprocessing (CBP) of lignocellulosic biomass

1.2.1 Motivation for CBP

Microbial processing of lignocellulosic material, an abundant and under-utilized carbon source, into liquid fuels is a promising alternative to petroleum-based fuels, but production costs using current technologies remain prohibitively high. Furthermore, the variable nature of biomass across space and time means that processes need to be optimized for region and season [36]. Cellulosic biofuel processes typically involve the following steps: (i) thermochemical pretreatment to break down the biomass into its three major components (cellulose, hemicellulose and lignin), (ii) enzymatic saccharification of cellulose and hemicellulose to hexose and pentose sugars, (iii) microbial fermentation of soluble sugars to generate fuel, and (iv) downstream processing to separate and concentrate the fuel [37]. Consolidated bioprocessing (CBP), in which enzyme production, enzymatic hydrolysis, and microbial fermentation occur in a single bioreactor (Fig. 1.5A), is believed to be the configuration of lowest cost, provided a suitable microbe or combination of microbes can be developed (Fig. 1.5B) [38,39].

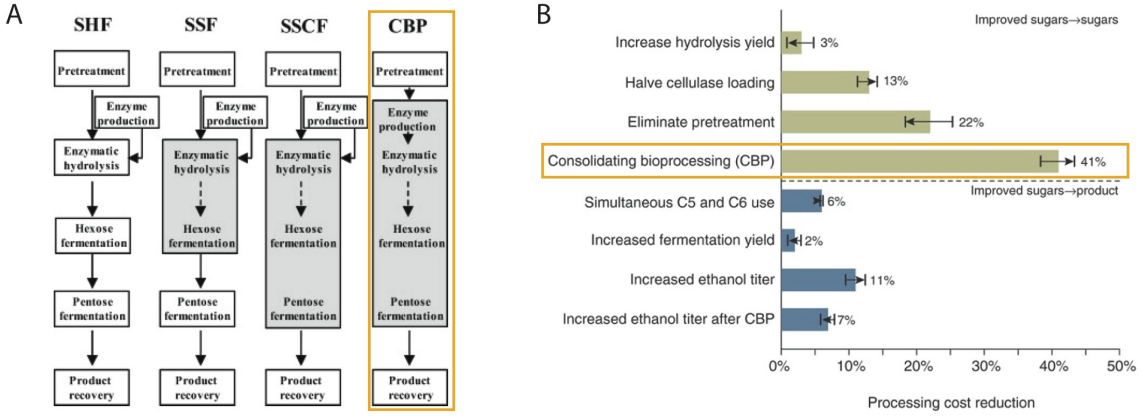

Figure 1.5: Cost savings of process consolidation. (A) Summary of process integration schemes for cellulosic biofuel generation. Each box represents a bioreactor (not to scale). SHF, separate hydrolysis and fermentation; SSF, simultaneous saccharification and fermentation; SSCF, simultaneous saccharification and co-fermentation; CBP, consolidated bioprocessing. Adapted from [36]. (B) Estimated cost reduction for various technological advances. Values represent the average of two scenarios: 2,205 and 5,000 dry tons feedstock/day. Adapted from [40].

The required traits for such a microbe are cellulase expression and secretion, biofuel production, solvent tolerance, lignin tolerance, and elimination of carbon catabolite repression that inhibits metabolism of other sugars in the presence of glucose [36]. Much effort toward CBP has been directed toward integrating all functionalities into a single organism, either starting from a cellulolytic organism and engineering for high fuel yield and tolerance or starting with a productive fermentation strain and adding cellulolytic capabilities. However, neither strategy has yet to result in an organism with sufficiently high yields and productivities for commercial viability [39]. An alternative approach is to split the work among multiple specialist organisms that can stably co-exist with each other. This approach reduces the metabolic load and number of functionalities that must be optimized in one organism.

1.2.2 Microbial consortium approaches for CBP

Natural microbial consortia capable of efficient (50-90%) lignocellulose degradation have been identified. These consortia largely extract all the available energy since accumulation of molecules that can serve as fuels invites exploitation by an additional consortia member. Most of the resulting products (e.g., organic acids, CO_2, CH_4) of naturally occurring consortia are not directly utilizable as liquid fuel [41]. Recent progress in the investigation of natural cellulose-degrading consortia includes identification of a thermophilic community capable of producing 2 g/L ethanol from 7 g/L cellulose. Upon optimizing the population composition, the ethanol titer was increased to 2.5 g/L, 78% of the theoretical yield [42]. However, natural consortia tend to involve large numbers of minimally characterized species, making them difficult to adapt to an industrial application aimed at producing a specific product. The challenge is thus to develop well-defined microbial consortia in which the electron transfer cascade stops short of completion, with some energy captured in fuel molecules [41, 43].

Although today's bioprocessing industry is dominated by monoculture systems, there

is growing interest in using microbial consortia for applications including biofuel production, bioremediation, lipid production, and biopolymer production. Advances in engineering and analysis methods are increasing the viability of this type of technology. New tools, reviewed in [44] and [45], include high-throughput screening methods for evaluating population composition [46], hydrogel encapsulation [47, 48], and techniques for culture condition optimization [49, 50].

Various approaches exist for splitting the work of lignocellulosic biofuel production among different microorganisms. One approach that has been taken is engineering specialist strains to secrete cellulolytic enzymes that act synergistically. Tsai *et al.* [51] engineered a four-strain *S. cerevisiae* consortium to produce a mini-cellulosome that enabled growth and ethanol production on phosphoric acid swollen cellulose. Each of the strains produces one recombinant protein - three cellulases with docking tags, and a scaffold. When the population composition of the consortium was optimized via inoculation ratio, it produced ethanol at yields of up to 0.475 g/g (93% of theoretical) and titer of 1.87 g/L ethanol.

Another approach is to combine a cellulolytic specialist that produces the enzymes for saccharification and a fermentation specialist that produces the fuel molecule from the soluble saccharides. Several proof of concept studies have demonstrated the feasibility of this approach. For example, Zuroff *et al.* [43] have constructed an anaerobic consortium of the cellulolytic *Clostridium phytofermentans* and fermentative *S. cerevisiae*. The yeast protected the *C. phytofermentans* from oxygen and symbiosis could be induced by controlling the volumetric transport rate of oxygen. This system reached 22 g/L ethanol from 100 g/L microcrystalline cellulose but required externally produced endoglucanase. Brethauer & Studer have used a multi-species biofilm membrane (MBM) setup to co-culture *Trichoderma reesei*, an aerobic cellulolytic fungus, *Scheffersomyces stipitis*, a natural pentose fermenter under microaerophilic conditions, and *S. cerevisiae*, an efficient hexose fermenter under anaerobic conditions [52]. The MBM enabled the species to self-assemble into layers, each receiving the appropriate oxygen content. This system achieved 67% of the theoretical yield on pretreated

wheat straw with added β-glucosidase enzyme.

Minty *et al.* [53] have also taken a saccharolytic/fermentation division of labor approach, developing a stable dual member aerobic consortium for conversion of cellulosic biomass into isobutanol. Isobutanol has a higher energy content, lower vapor pressure, and lower hygroscopicity than the "first generation" biofuel, ethanol. Unlike ethanol, which must be blended with gasoline, isobutanol is compatible with current engines and infrastructure [54].

The cellulolytic specialist of this system is *T. reesei* strain RUT-C30, a strain developed to hypersecrete cellulase enzymes that is widely used industrially and academically [55]. The fermentation specialist is one of several *E. coli* strains engineered by the Liao group to produce isobutanol from glucose via the combination of the branched-chain amino acid biosynthesis pathway with two heterologous Ehrlich pathway steps. The strain NV3 was created for isobutanol production by random mutagenesis and selection on norvaline, a valine/leucine analog, the presence of which requires the cells to increase branched-chain amino acid production for survival [11]. NV3r1 is a derivative of NV3 in which a mutation to the gene encoding RpoS, an important stationary phase transcriptional regulator which helps the cell survive stress, was repaired. Under ideal conditions these strains can produce isobutanol with 59% (NV3) or 76% (NV3r1) yield of the theoretical maximum from glucose, the latter being one of the highest reported for candidate next-generation biofuels [11].

The co-culture (with *E. coli* strain NV3) achieved up to 1.88 g/L isobutanol and yields up to 62% of the theoretical maximum on ammonia fiber expansion (AFEX) pretreated corn stover in minimal medium without any enzyme supplementation [53]. In an investigation of potential limiting factors, it was found that plasmid loss is quite significant and likely a limiting factor of culture performance (Fig 1.6).

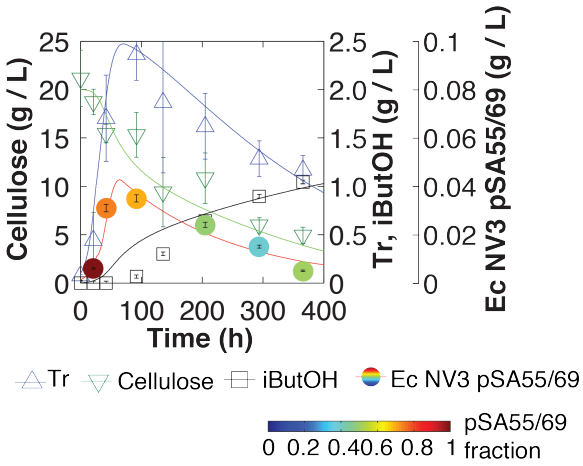

Figure 1.6: Plasmid loss over time in *T. reesei/E. coli* cocultures. Reproduced from [53].

1.3 Chromosomal integration and expression optimization in *Escherichia coli*

1.3.1 Benefits of pathway integration

Although many metabolic engineering efforts are conducted using plasmids due to the ease of manipulation, for industrial production strains it is highly desirable to have pathway or recombinant protein genes stably integrated into the genome. Chromosomal integration eliminates the need for the use of antibiotics for plasmid maintenance, which is desirable for cost reasons and also to avoid the spread of antibiotic resistance. However, with current methods, it is difficult to predict genomic expression levels of synthetic constructs. Techniques for chromosomal integration and expression level optimization are reviewed in [56]. Generally, chromosomal expression is lower than desired. Methods to increase expression include the use of a strong promoter [57,58] or increasing copy number [59–61]. Chromosomal location choice has also been explored since expression level at different sites can vary significantly [61,62]. More work is needed to elucidate the best strategies for choosing integration location and optimizing expression level from the genome.

1.3.2 Isobutanol pathway integration

Some work has already been done to develop strains with chromosomally integrated isobutanol pathways. Akita *et al.* [63] integrated the isobutanol pathway into the genome under xylose inducible promoters, with genes in several different genomic locations, and used the strains to produce isobutanol from Japanese cedar hydrolysate without the addition of antibiotics or exogenous inducer molecules. These strains reached 66% of the theoretical yield on glucose/xylose and 14% on hydrolysate. Bassalo *et al.* [64] have also demonstrated a chromosomally integrated isobutanol pathway, with the whole pathway in one location, but production level was low (2.2 g/L isobutanol from 85 g/L glucose). Further optimization of expression levels can presumably increase yields and titers.

1.4 Dissertation overview

This dissertation presents the development of a new high-throughput screening platform utilizing cross-feeding auxotrophic biosensors. It then describes the utilization of the screening method, as well as other strategies, to develop chromosomally integrated isobutanol production strains. It also describes strategies to improve the *T. reesei/E. coli* consolidated bioprocess developed by Minty *et al.* [53]. The chapters are as follows:

Chapter 1: Background and motivation: This chapter discusses the motivation for novel high-throughput screening methods, consolidated bioprocessing for lignocellulosic biofuel production, and chromosomal integration of heterologous genes.

Chapter 2: Development of Syntrophic Amplification of Production Phenotype (SnoCAP) screening platform: This chapter describes the development and testing of a screening strategy based on a cross-feeding metabolic circuit. It demonstrates implementation in microtiter plates.

Chapter 3: Higher-throughput formats of SnoCAP: This chapter describes the implementation of SnoCAP in a colony screening assay and a microdroplet co-cultivation and

sorting assay. It also details the screening of a chemically mutagenized library to find strains with improved 2-ketoisovalerate production based on genomic expression rather than pathway overexpression on a plasmid.

Chapter 4: Development of *E. coli* strains with chromosomally integrated isobutanol pathway: This chapter describes the use of chemically inducible chromosomal evolution [59] to produce strains with multiple copies of the *kivd* and *adhA*. It also discusses the construction and screening of a library in which *alsS* is integrated into random sites in the genome via transposon in order to achieve varying expression levels.

Chapter 5: Strategies for improving fungal-bacterial co-culture for more efficient consolidated bioprocessing: This chapter describes efforts to improve the *T. reesei/E. coli* by deletion of competing pathways in the *E. coli* strain and preliminary work toward adaptive evolution of *T. reesei/E. coli* co-cultures.

Chapter 6: Concluding remarks and perspectives: This chapter concludes the dissertation and describes potential future directions.

Chapter 2: Development of Syntrophic Amplification of Production phenotype (SnoCAP) screening platform

2.1 Introduction

Advances in genome engineering design and construction technologies have enabled rapid generation of diverse microbial strains that efficiently explore the genotype space [65]. Characterizing these strains, however, is often a bottleneck in the design-build-test cycle of synthetic biology. In particular, concerning the development of strains for production of many target molecules, increasingly large and complex strain libraries can be created, yet the throughput of screening for identifying top-performing variants is limited, sometimes lagging several orders of magnitude behind the construction phase. Traditionally, for molecules that lack chromogenic or fluorescent properties, metabolic engineers must rely on chromatography or mass spectrometry quantification. Automation can help to increase the throughput of these assays, but they have high capital investment and space requirements. Biosensor-based high-throughput screenings seek to address this challenge by converting target molecule production level into a conspicuous phenotype such as growth or fluorescence, either within the production cell itself or in a partner strain. The latter approach enables the sensing of extracellular secretion levels and reduces the interference between the production and sensing functionalities. The sensing machinery generally consists of proteins, nucleic acid molecules or whole cells that respond to the target molecule in a dose-dependent manner and produce a detectable read-out [15, 17, 66].

One class of whole-cell biosensor consists of auxotrophic strains which are unable to

produce an essential metabolite and whose growth characteristics, therefore, change in response to changing concentrations of the target molecule in the surrounding environment. Auxotrophic microbial strains have been identified or constructed and utilized as biosensors for a variety of metabolites including amino acids [21–25], vitamins [26], and hormones [27]. Pfleger *et al.* applied auxotrophic biosensors to high-throughput screening of production strain libraries by developing a fluorescent mevalonate auxotroph whose growth reports on production strain performance [28]. Tepper & Shlomi [67] have established a computational framework to predict gene deletions that can be used to produce auxotrophic strains for use as biosensors. For *Escherichia coli*, for instance, they predict 53 molecules for which auxotrophic biosensor strains could be created. Furthermore, for molecules for which no auxotrophic biosensor is available, they present a strategy to engineer the producer strain by gene knockout so that production of the target molecule is coupled to that of a proxy metabolite, for which a biosensor does exist.

Despite the array of auxotrophic biosensors available, a limitation in applying them to high-throughput screening is that they generally have narrow dynamic ranges, confined to low concentrations of the focal molecules. Thus, although these molecules are essential for growth, they do not directly confer a selective advantage if the strains are producing more than small quantities of the molecule. This small dynamic range limitation can sometimes be overcome by dilution of the samples [28, 68], but this makes screening more cumbersome, lowering throughput.

Here, we describe the utilization of auxotrophic strains in a cross-feeding circuit for Syntrophic Co-culture Amplification of Production phenotype (SnoCAP) that enables high-throughput screening of production strains via colocalization with a partner strain (Fig. 2.1). One strain, the "sensor", is auxotrophic for the target molecule; its ability to grow depends upon the amount of target molecule excreted by the other strain, the "secretor." The secretor is auxotrophic for an orthogonal molecule supplied by the sensor strain. In model microbial systems, it has been shown that changes in the secretion or uptake characteristics

of either partner of a cross-feeding pair determine the resulting composition of the co-culture as well as its overall co-growth rate [69,70]. We predict that a secretor strain with improved production rate will lead to faster growth and an increased final sensor-to-secretor ratio (Fig. 2.1A). The strategy requires compartmentalization of secretors of unique genotype with the sensor strain and we have implemented this in several formats (Fig. 2.1B). In this chapter, we use confinement in wells of 96-well microtiter plates. Higher throughput formats (colony screening on agar plates and microdroplet co-cultivation and sorting) are explored in Chapter 3.

The majority of the work presented in this chapter has been submitted for publication as a manuscript titled "Syntrophic co-culture amplification of production phenotype for high-throughput screening of microbial strain libraries." Authors: T. E. Saleski, A. R. Kerner, M. T. Chung, C. M. Jackman, A. Khaasbaatar, K. Kurabayashi, X. N. Lin.

2.2 Results

2.2.1 Model-based prediction of amplification of production improvement via metabolic cross-feeding circuits

Kerner *et al.* [69] presented an ODE model of a cross-feeding co-culture in its exponential growth phase, assuming constant secretion and uptake parameters and Monod kinetics for growth on a limiting nutrient (i.e., the amino acid for which the strain is auxotrophic). In this model, the co-culture reaches a pseudo-steady-state in which the two strains have the same growth rate (μ, unit: 1/hr), which depends on each auxotroph's secretion rate (α_{Sec}, α_{Sens}; unit: mmol/gDM-hr) of its shared metabolite and the per cell growth requirements for the cross-fed metabolites (β_{Sec}, β_{Sens}; unit: mmol/gDM). Additionally, a steady population composition ratio (r) between the number of each cell type (N_{Sec}, N_{Sens}) is reached.

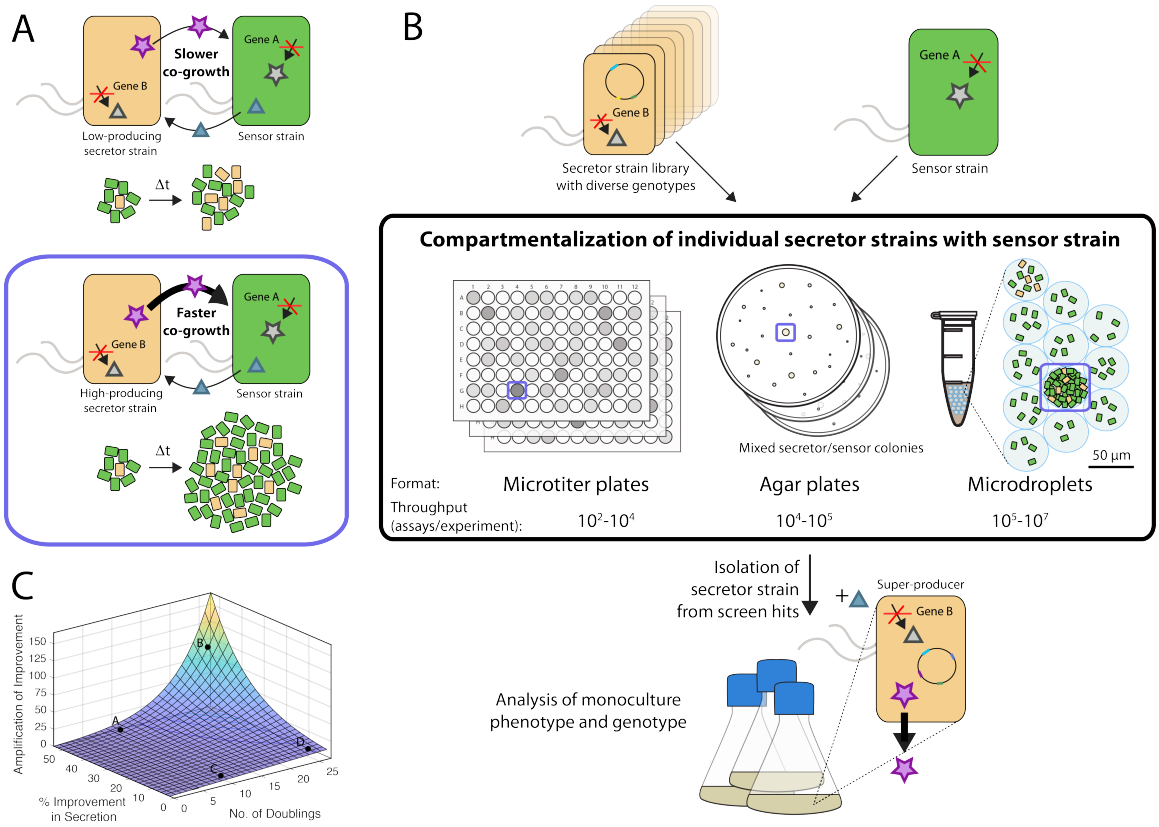

Figure 2.1: Overall schematic of screening strategy. (A) Production improvement leads to increased co-growth and increased final sensor-to-secretor ratio. Δt represents a change in time. The purple star represents the molecule that one wishes the secretor strain to overproduce. The blue triangle represents a secondary cross-fed metabolite. (B) Expected amplification of secretion improvement by co-culture growth. A production strain with 50% improvement compared to a base strain will lead to a co-culture with 4.7 times as many cells as that of the base strain (370% more) after a time corresponding to 10 doublings of the base co-culture. This is a 7.5-fold amplification of the percentage improvement (point A). After 25 doublings of the base strain co-culture, this amplification will rise to 96-fold (point B). For a smaller improvement of 5%, there will be 1.2 times as many cells in the improved secretion co-culture after 10 base doublings, a 3.7-fold amplification (point C). After 25 doublings of the base co-culture, this will increase to a 10.7-fold amplification (point D). (C) Screening implementation formats explored in this study: separating secretor cells of unique genotype by compartmentalization in wells of microplates, spatial separation on agar plates, or confinement in microdroplets. After growth, secretor strains are isolated from well-grown co-cultures for further analysis in monocultures.

Mathematically, these properties are given by:

$$\mu = \sqrt{\alpha_{Sec}\alpha_{Sens}/(\beta_{Sec}\beta_{Sens})} \text{ and } r = \sqrt{\alpha_{Sec}\beta_{Sec}/(\alpha_{Sens}\beta_{Sens})}.$$

Thus, if the secretion rate of the secretor (α_{Sec}) increases and all else remains unchanged, we expect the co-culture growth rate to increase and the percentage of the final population that is the sensor to increase. If we have a base-level secretor strain with secretion rate α_{Base} and an improved secretor with secretion rate $\alpha_{Base}(1+x)$ and we grow each secretor with the sensor strain, the exponential growth of the co-culture quickly amplifies even moderate differences in production level (Fig. 2.1C). After an amount of time corresponding to n doublings of the base strain co-culture, the improved secretor's co-culture will have $2^{n(\sqrt{1+x}-1)}$ times as many cells as the base strain's co-culture.

At some high enough $\alpha_{Sec}N_{Sec}$ such that the target molecule is no longer limiting for the sensor strain, this model breaks down. There is then one-directional feeding where the sensor can grow without further growth of the secretor strain. In this case, the maximum growth rate may not be increased for a higher producing strain, but the time to reach this critical value of $\alpha_{Sec}N_{Sec}$ will vary for secretor strains with different α_{Sec} production rates. We, therefore, expect that, over some range of secretion levels, improvements in secretor strain production will lead to detectable changes in co-culture growth and composition.

2.2.2 Development of two model systems: 2-ketoisovalerate and tryptophan

We first demonstrated the SnoCAP screening framework using 2-ketoisovalerate (2-KIV) as a target molecule. 2-KIV is a precursor of the branched-chain amino acids valine and leucine. By overexpression of an alpha-ketoisovalerate decarboxylase (Kdc) and an alcohol dehydrogenase (Adh) in *E. coli*, 2-KIV can be converted into the drop-in biofuel isobutanol. Additional overexpression of three enzymes that catalyze the conversion of pyruvate to 2-KIV leads to substantially improved isobutanol production [10] (Fig. 2.2A, B).

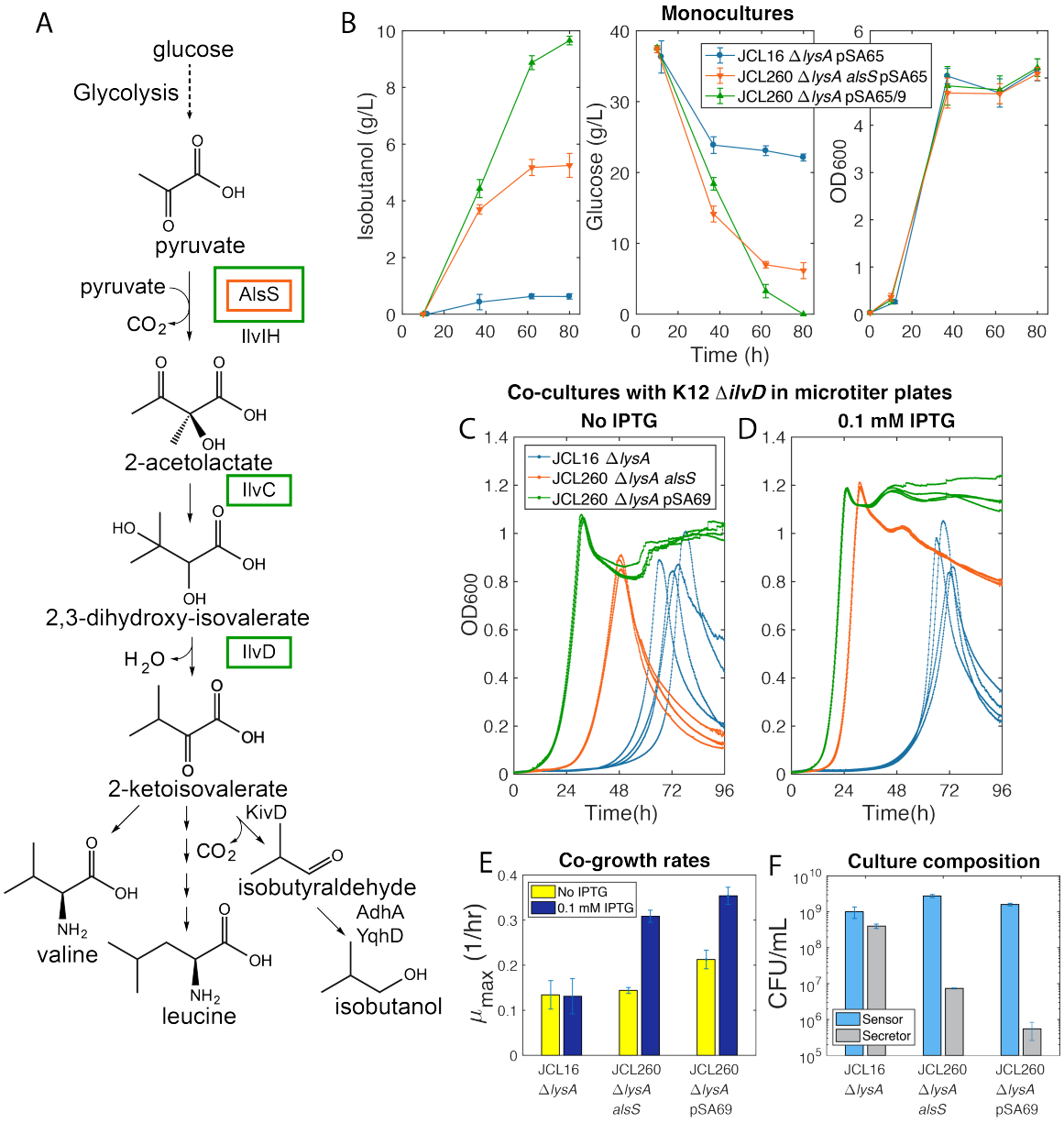

Figure 2.2: Co-culture growth characteristics correlate with monoculture production performance. (A) 2-KIV and isobutanol production pathway. Genes that are overexpressed in the production strains are boxed in the color corresponding to the strain in the legend in (B). (B) Monoculture production characteristics of three different secretor strains carrying pSA65 ($P_L$lacO1::*kivd–adhA*). Isobutanol, glucose and cell growth profiles are shown. Error bars represent the standard deviation of four biological replicates. Growth profiles of co-cultures of the same three secretor strains (not carrying pSA65) with sensor strain K12 Δ*ilvD::kan* in a microplate with (C) and without (D) IPTG-induced expression of the *alsS/ilvCD* genes. Three replicate wells are shown for each strain pair, plotted in the same color. (E) Co-culture growth rates. Error bars represent the standard deviations between at least 10 biological replicates, combined from two separate experiments. (F) Population composition of microplate cultures in early stationary phase. Error bars represent standard deviations from two biological replicates with two technical replicates each.

We chose a Δ*ilvD* auxotroph as the sensor strain. IlvD, dihydroxy-acid dehydratase, catalyzes the conversion of 2,3-dihydroxy-isovalerate into 2-KIV. IlvD is also part of the isoleucine biosynthesis pathway, catalyzing the conversion of 2,3-dihydroxy-3-methylvalerate into 2-keto-3-methylvalerate. We therefore supplemented the co-cultures with an excess of isoleucine in order to eliminate effects from variation in isoleucine cross-feeding levels. When grown with excess isoleucine and varying levels of 2-KIV, the *ilvD* auxotroph's growth rate and maximum cell density increase in response to increasing 2-KIV over a certain range (Fig. 2.3A, B).

For a given *E. coli* auxotroph, a variety of cross-feeding partner auxotroph options are generally available, exhibiting a range of co-culture growth rates [34, 35]. We tested a panel of potential partner auxotrophic strains for their ability to cross-feed with K12 Δ*ilvD* in a minimal medium supplemented with isoleucine. Of the partner auxotrophs tested (Δ*hisD*, Δ*leuB*, Δ*lysA*, Δ*pheA*, Δ*ppC*, Δ*trpB*, and Δ*tyrA*, each in strain BW25113), Δ*lysA* (a lysine auxotroph) and Δ*pheA* (a phenylalanine auxotroph) showed consistent growth (Fig. 2.4). Both showed considerable lag phases (∼2 and 5 days, respectively) and lower maximum optical densities than monocultures, leaving room for improved secretion to boost co-culture growth. We chose to proceed with lysine as the secondary cross-fed molecule. Selecting a different auxotroph, including ones that have no growth with the base production-level strain, may be a useful strategy to adjust the dynamic range of the screening.

We tested the growth properties in 96-well microplates of sensor strain K12 Δ*ilvD* in co-culture with secretor strains of several different secretion levels. Genotypes of these strains are listed in Table S1. JCL16 Δ*lysA* is the base strain, with a low production level. JCL260 Δ*lysA* *alsS*, which has six gene deletions to direct flux through the isobutanol pathway and a single copy of *alsS*, under an IPTG-inducible promotor, integrated into the genome, represents an intermediate-production-level strain. JCL260 Δ*lysA* pSA69, which has the same gene deletions and carries a plasmid for overexpression of *alsS* and *ilvCD* under an IPTG-inducible promotor, represents a high-production-level strain. When transformed with

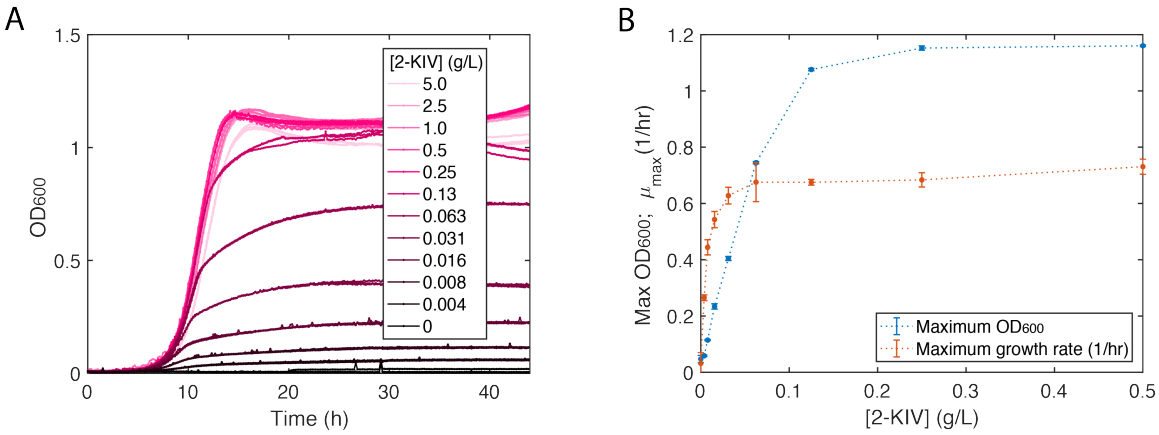

Figure 2.3: Sensitivity of K12 Δ*ilvD* sensor growth characteristics to 2-KIV concentration in monoculture. (A) Growth profiles of the K12 Δ*ilvD* sensor strain. The medium contains an excess of isoleucine. Three replicate wells are plotted in the same color for each 2-KIV concentration. (B) Maximum OD_{600} and growth rate from the growth profiles in (A).

pSA65, which carries *kivD/adhA*, and grown in monoculture fermentations, these strains have differing levels of isobutanol production and glucose consumption, and similar growth profiles (Fig. 2.2B). We inoculated these strains (lacking pSA65) in co-culture with the sensor strain at a 1:1 initial ratio in minimal medium with isoleucine. Monoculture inoculation of any of these strains or the sensor strain in this medium produces no detectable growth. In the co-cultures (Fig. 2.2C), growth order and growth rate (Fig. 2.2E) increases with increasing strain production level. Addition of IPTG results in a higher growth rate for co-cultures with each of the strains carrying an IPTG-inducible operon (either *alsS* only or *alsS-ilvCD*), and no difference for base strain JCL16 (Fig. 2.2D, E). This indicates that expression of

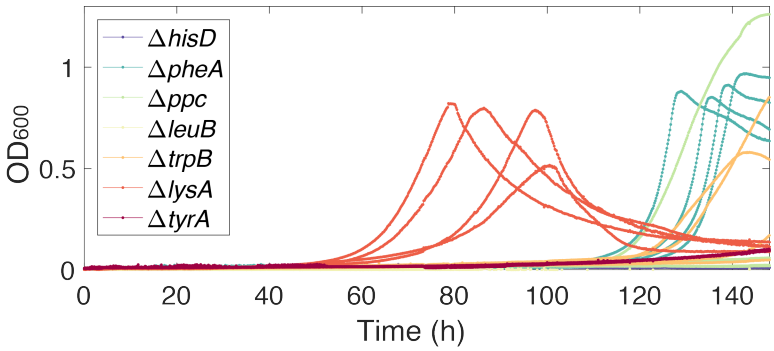

Figure 2.4: Co-culture growth profiles of various Keio strains with K12 Δ*ilvD*. Biological replicates are plotted in the same color (n = 4). Legend denotes the gene knockout of the Keio strain.

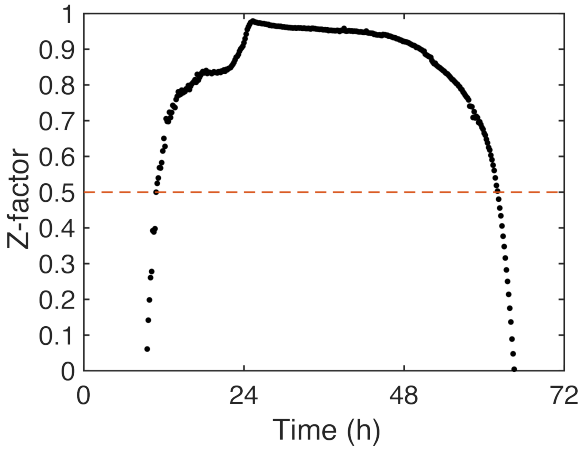

Figure 2.5: Z-factors over time for the microplate assay. Z-factors were calculated from the OD_{600} readings of co-cultures containing secretor strain JCL16 $\Delta lysA$ as negative control and JCL260 $\Delta lysA$ pSA69 as positive control (8 replicates each; medium containing 0.1 mM IPTG and no norvaline). The Z-factor is defined as $Z = 1-(3\sigma_p+3\sigma_n)/(|\mu_p-\mu_n|)$, where μ and σ are the mean and standard deviation of the positive (p) and negative (n) controls.

the genes leading to 2-KIV translates into increased growth in the co-culture setting. We determined the composition of the co-cultures when in early stationary phase by differential plating (see Materials and methods). As predicted, the sensor-to-secretor strain ratio rises with increasing production level of the secretor strain (Fig. 2.2F).

To evaluate the suitability of the cross-feeding co-culture growth assay for high-throughput screening, we calculated the Z-factor, a parameter that reports on a combination of the signal dynamic range and assay precision [71]. The Z-factor is defined as $Z = 1-(3\sigma_p+3\sigma_n)/(|\mu_p-\mu_n|)$, where μ and σ are the mean and standard deviation of the positive (p) and negative (n) controls and values between 0.5 and 1.0 are considered indicative of an excellent assay. We considered co-cultures containing secretor JCL16 $\Delta lysA$ as the negative control and JCL260 $\Delta lysA$ pSA69 as the positive control and found that the Z-factor was in the "excellent" range from hours 11 to 62 of the assay period (Fig. 2.5). This analysis indicates that the assay can be quite effective at identifying improved variants at times before the parent strain has produced co-growth.

As a second test case of the SnoCAP screening framework, we examined tryptophan-producing strains. For this implementation, we chose BW25113 $\Delta trpB$, which lacks the

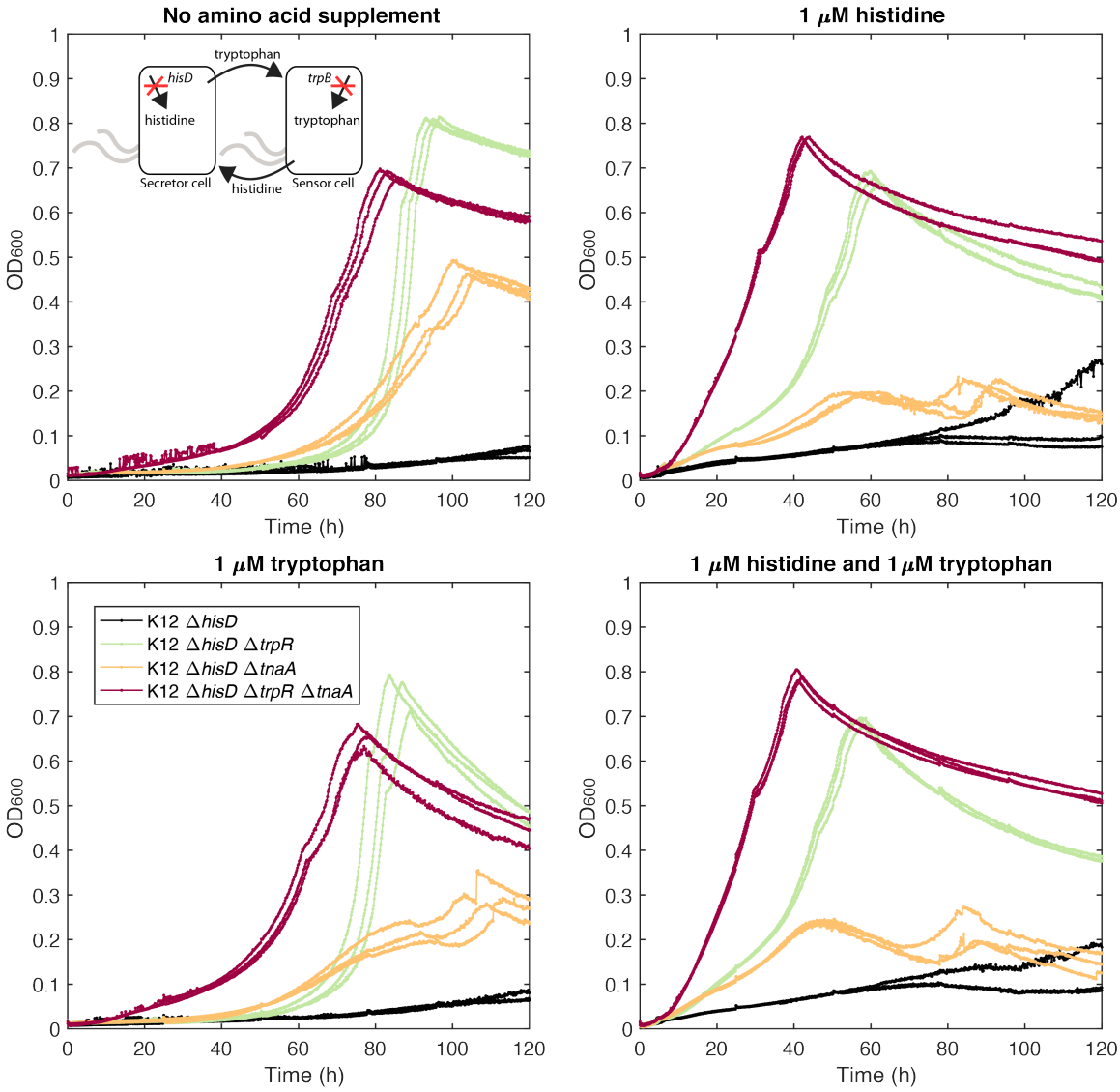

Figure 2.6: Implementation of the SnoCAP screening method for tryptophan production. Plots show co-culture growth in 96-well microplates with different initial amino acid supplementation. Three biological replicates are shown for each strain, plotted in the same color. The inset shows the scheme of tryptophan/histidine cross-feeding.

catalytic subunit of tryptophan synthase, as the sensor strain. We selected histidine as the secondary cross-fed molecule (Fig. 2.6, inset), based on work showing that growth of tryptophan/histidine cross-feeding co-cultures is affected by overproduction of tryptophan [72]. For overproduction strains, we deleted the *trpR* and *tnaA* genes, both individually and sequentially in the same strain. Deletion of these genes, which encode a repressor of the trp operon and a tryptophanase, respectively, are typical early steps in the engineering of tryptophan production strains (e.g., [73, 74]). In liquid co-cultures in microplates, all three modified strains had increased co-culture growth compared to the base strain. However, the growth took several days to be observable. Since these deletions are only first steps in tryptophan strain engineering, it may be useful that these strains' co-growth is slow, since it leaves substantial room for improvement with higher production-level strains. Nevertheless, we were also interested in determining whether we could decrease assay time while still maintaining the detectable differences in growth phenotype between the higher production-level strains and base strain. We added low levels of histidine and tryptophan to jump-start the co-cultures and found that addition of histidine produced the desired effect (Fig. 2.6).

Biphasic growth was observed in many of these cultures (Fig. 2.6). It is not immediately obvious what leads to this phenomenon, but one possibility is that tryptophan has started to accumulate in the medium and this changes the dynamics of the culture. Analysis of the culture composition by fluorescent labeling or by differential plating may provide insights into what is occurring.

2.2.3 Increasing the dynamic range by utilization of an inhibitory analog of the target molecule

One common issue in previous biosensor-based screening is the limited dynamic range of production level. We hypothesized that the addition of an inhibitory analog of the target molecule could expand the dynamic range of screening significantly and increase differences of co-culture growth in the SnoCAP framework. Analog selection, in which bacteria are grown

on an inhibitory analog of a metabolite in order to select for overproduction of that metabolite, is a useful strategy for strain development and has been used to identify mutations that overexpress pathway genes or decrease feedback inhibition from target molecules [75, 76]. A computational framework has recently been developed to identify candidate metabolite analogs for use in strain improvement [77].

Norvaline, a toxic analog of valine and leucine, has been used as a selection agent for increased flux through the valine [75] and isobutanol [11] pathways. Norvaline produces a growth defect that can be partially recovered by the addition of either leucine or valine alone or fully recovered by addition of both leucine and valine (Fig. 2.8A). We first verified that the sensor strain's growth remains responsive to increasing 2-KIV in the presence of norvaline (Fig. 2.7A, B). We then added norvaline to the co-cultures and found that its addition magnifies differences between the strains at the higher end of the production spectrum (Fig. 2.8B). For low levels of norvaline, the co-cultures with JCL260 Δ*lysA* pSA69 as the secretor strain are not affected by the norvaline, while those with JCL260 Δ*lysA alsS* exhibit an increasing lag phase time with increasing norvaline concentration. For higher levels of norvaline, such as 1.0 g/L, co-cultures with JCL260 Δ*lysA* pSA69 show lengthening lag phase as well. Thus, norvaline can be used to expand the dynamic range and increase the production threshold below which secretor strains cannot support co-culture growth.

It is also interesting to note that, because the co-culture growth characteristics are determined primarily by production levels, the SnoCAP screening framework can avoid or mitigate the issue of metabolic burden on cell growth caused by over-expression of synthetic gene constructs, compared to direct screening with monocultures. For instance, although pSA69 increases 2-KIV production, we observed that when its expression is induced with IPTG, the growth of the strain in norvaline is decreased (Fig. 2.8B).

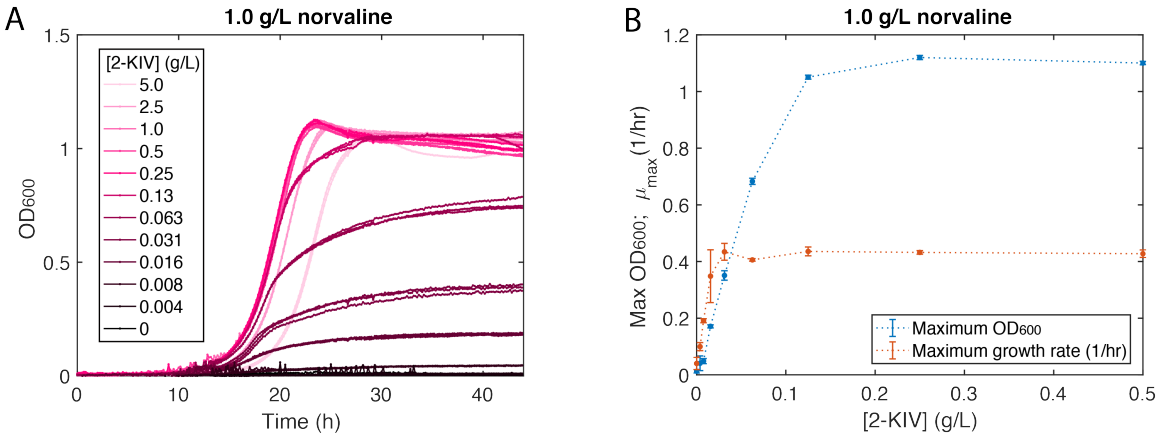

Figure 2.7: Sensitivity of K12 Δ*ilvD* sensor growth characteristics to 2-KIV concentration in monoculture with norvaline. (A) Growth profiles with 1.0 g/L norvaline. The medium contains an excess of isoleucine. Three replicate wells are plotted in the same color for each 2-KIV concentration. (B) Maximum OD_{600} and growth rate from the growth profiles in (A).

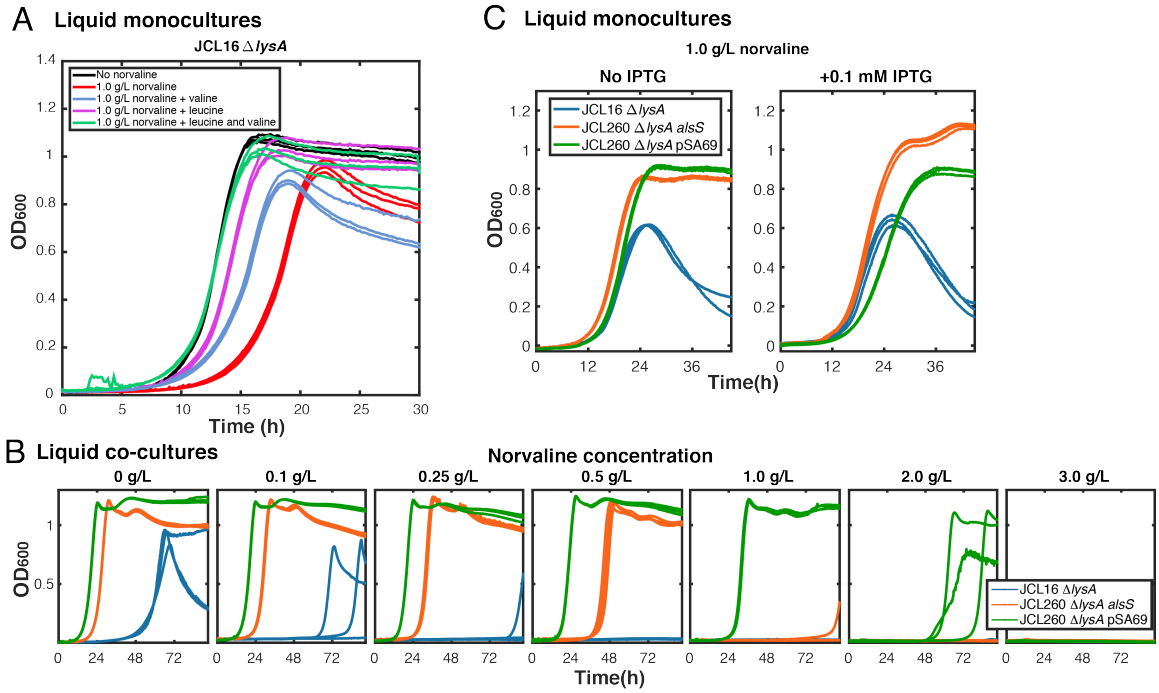

Figure 2.8: Norvaline as a tool to expand assay dynamic range. (A) Growth profiles of JCL16 Δ*lysA* supplemented with 3 mM lysine and 3 mM isoleucine with and without 1.0 g/L norvaline, 3 mM valine, and 3 mM leucine. Three biological replicates of each culture are plotted in the same color. (B) Liquid co-cultures with various levels of norvaline. (C) Liquid monocultures on 1.0 g/L norvaline, demonstrating that norvaline is not an effective selection tool for these strains in monoculture.

2.2.4 Application of an intermediate-sensor assisted push-pull strategy to screen for isobutanol production

Intermediate-sensor assisted push-pull strategy has been proposed as a general method for screening for production of molecules that lack a direct sensor but for which a sensor is available for an intermediate, and has been successfully demonstrated for deoxyviolacein with a tryptophan biosensor [78]. This approach involves using the intermediate sensor to screen for increased production of the intermediate and then improving the conversion of intermediate to final product by screening for decreased readout from the biosensor. We investigated whether SnoCAP can be employed in such a manner, i.e., whether increased activity of a pathway converting the cross-fed molecule to a target molecule will decrease the co-culture growth.

Using the 2-KIV system, we compared co-culture growth with secretor strain JCL260 Δ*lysA* pSA69, a strain in which the pyruvate to 2-KIV part of the pathway performs well, to that of secretor strain JCL260 Δ*lysA* pSA65/9, which additionally carries the 2-KIV to isobutanol part of the pathway. We observed a significant decrease in co-culture growth rate and increase in lag time for the pSA65 carrying strain (Fig. 2.9A). We also examined intermediate levels of Kivd/AdhA expression in strains that contain copies of these genes integrated into the genome. These strains were generated using chemically inducible chromosomal evolution (CIChE), which enables the copy number of a construct of interest to be adjusted by changing the concentration of a lethal chemical (e.g., chloramphenicol, to which the resistance can be rendered by the integrated construct in a dose-dependent manner) in the medium [59]. Using this method we obtained strains JCL260 Δ*lysA* cm 20 and JCL260 Δ*lysA* cm 80. These strains have differing levels of isobutanol production (Fig. 2.9B). We saw that with the lower levels of Kdc/Adh expression there is still a decrease in co-culture growth but to a lower degree than with the high expression from pSA65 (Fig. 2.9C).

Because the addition of pSA65 adds metabolic burden and a toxic product, it was not clear how much of the decrease in growth was indeed due to 2-KIV being channeled away

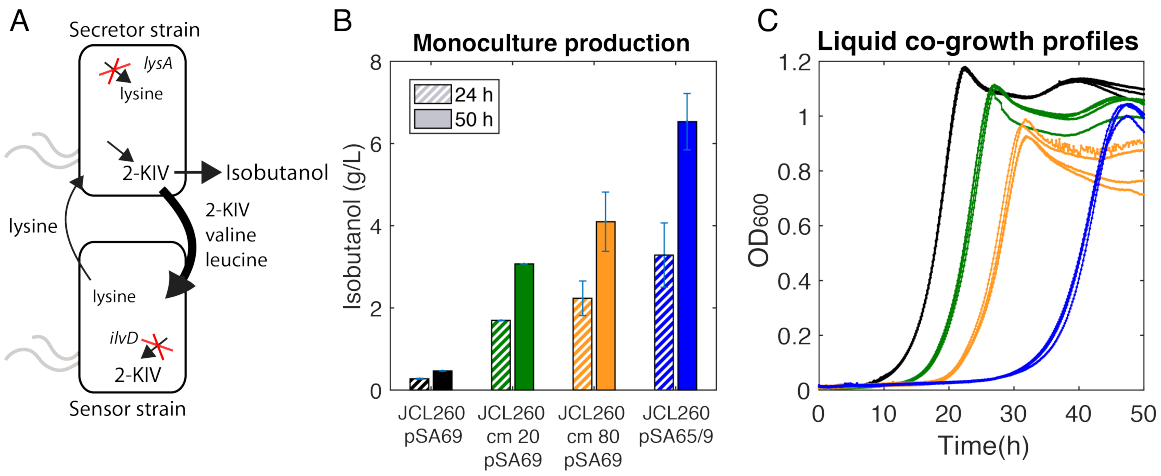

Figure 2.9: Screening for conversion of 2-KIV into isobutanol. (A) Schematic of sensing intermediate. (B) Production levels of the four strains with differing levels of *adhA/kivd* expression in M9 minimal medium, 36 g/L glucose. (C) Liquid co-culture growth profiles of secretor strains carrying varying levels of *adhA/kivd* expression.

from cross-feeding. We, therefore, tested these co-cultures under the condition of excess leucine and valine instead of isoleucine. This condition leads to a co-culture that must cross-feed isoleucine or intermediates for the production of isoleucine (Fig. 2.10A). If defects in growth were due mainly to factors other than 2-KIV channeling, such as a metabolic burden or product toxicity, then we would expect them to be observed under the leucine/valine supplementation as well. Under the leucine/valine-supplemented conditions, we observed some basal growth of the sensor strain monoculture, but the co-cultures have significantly more growth. For both secretor strains tested (JCL260 Δ*lysA* *alsS* and JCL260 Δ*lysA* pSA69), the version carrying pSA65 has reduced growth compared to that without pSA65 under the isoleucine-supplemented condition (Fig. 2.10C) and has similar growth to that without pSA65 under the leucine/valine-supplemented condition (Fig. 2.10B). These results support the prediction that the co-growth characteristics will be determined primarily by the production level of the cross-fed molecules.

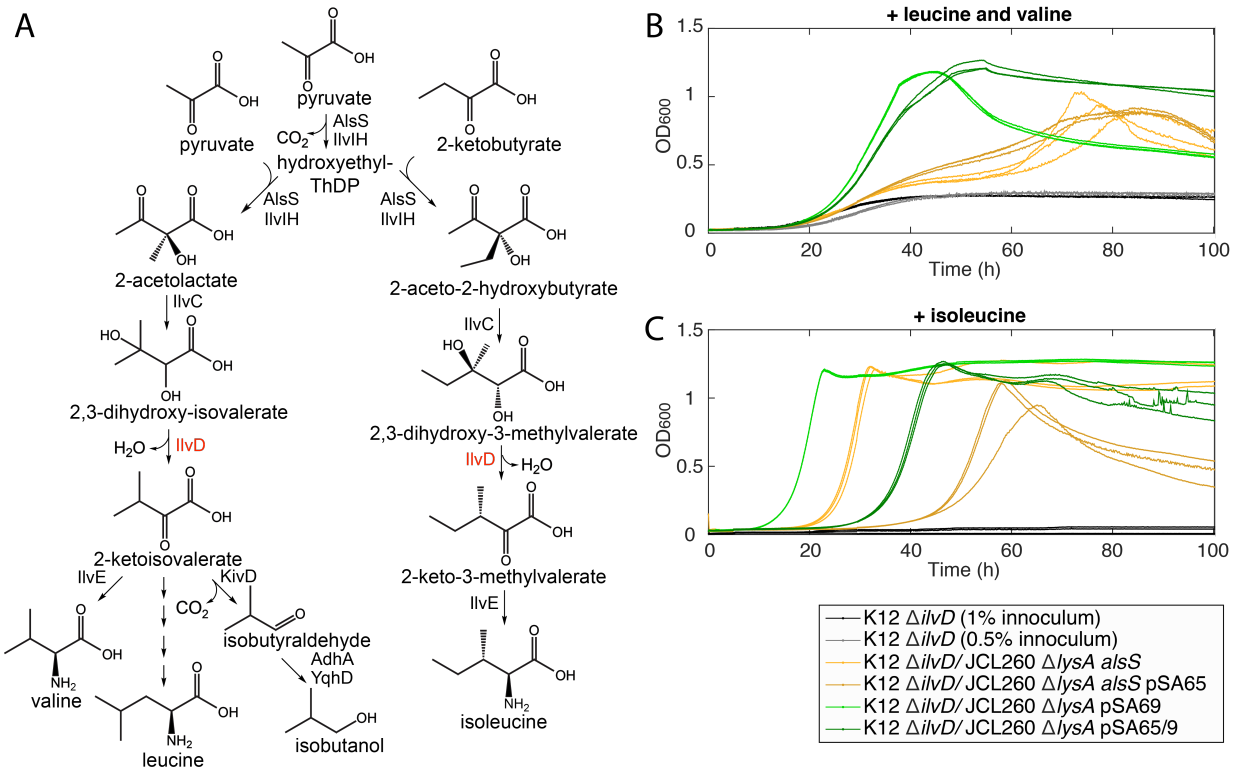

Figure 2.10: Investigation of the cause of the reduced co-culture growth in strains with increased alpha-ketoisovalerate decarboxylase and alcohol dehydrogenase activity. (A) Branched-chain amino acid and isobutanol pathways. Co-culture and sensor strain monoculture growth profiles in liquid medium supplemented with either 3 mM of both leucine and valine (B) or 3 mM isoleucine (C).

2.3 Discussion and conclusion

In this chapter, we have developed the SnoCAP framework for converting inconspicuous production phenotypes into growth phenotypes, facilitating high-throughput screening of production strain libraries via compartmentalization of cross-feeding production and sensor strains. The use of metabolite analogs makes the assay dynamic range highly tunable without requiring any genetic modifications. The tunability could be further improved by making genetic modifications to the sensor strain, such as adjusting the production rate of the secondary cross-fed molecule. We also demonstrate that the assay timing can be adjusted by kick-starting the culture with low-concentrations of the cross-fed metabolites. Since assay readout is cell density or cell fluorescence, no costly assay reagents are necessary. We have shown the utility of this method for screening strain libraries and identified a strain that

overproduced 2-KIV without a plasmid, which is desirable since plasmids require antibiotics for maintenance, and even then may still be lost under non-ideal conditions [53].

As proofs of concept, we explored SnoCAP for screening production of 2-KIV, which can be converted into the drop-in biofuel isobutanol, and the amino acid tryptophan. We expect that the SnoCAP screening framework can be applied to a wide variety of industrially relevant target molecules for which auxotrophic strains exist or can be constructed. Amino acid precursors that can be converted into other alcohols of interest (Fig. 1.1A) would also be interesting to explore. We also expect that this technology can be applied to various other microbial species. Synthetic cross-feeding consortia have been examined in yeast [79], and *E. coli* has shown the ability to cross-feed with other species, such as *Acinetobacter baylyi* [72] or Salmonella species [33, 80]. The strategy could also be extended to screening for overproduction of other compounds of interest that a strain with a diverse metabolism, such as *Pseudomonas putida*, can utilize as a carbon source but the secretor strain cannot. In this case, a carbon source would be supplied to the secretor that the sensor cannot utilize. Since many microbes have naturally occurring auxotrophies, there are likely cases where the secondary auxotrophy does not even require genetic modification of the secretor strain.

2.4 Materials and methods

2.4.1 Strains and plasmids

Strains and plasmids used are listed in Tables 2.1 and 2.2, respectively. JCL16, JCL260, pSA65 and pSA69 [10, 81] were provided by James Liao, UCLA. Keio strains were obtained from the *E. coli* genetic stock center (CGSC, http://cgsc2.biology.yale.edu/). The pTGD plasmid [59] was provided by Keith Tyo, Northwestern University.

2.4.2 Gene deletions and insertions

Oligonucleotides were ordered from Integrated DNA Technologies (Coralville, IA) and are listed in Table 2.3.

Gene deletions were constructed by P1 phage transduction as previously described [82,83]. Keio strains were used as the donor strains and LB agar with 50 μg/mL kanamycin was used as the selective medium. Transductants were purified from residual P1 phage by isolation streaking on LB agar supplemented with 0.8 mM sodium citrate and 50 μg/mL kanamycin and verified by colony PCR as described previously [83]. When required, the FRT-flanked kanamycin resistance gene used for selection was removed by transformation with a temperature-conditional plasmid, pCP20, expressing FLP-recombinase from a thermoinducible promoter.

To integrate the *alsS* gene into the site between *yghX* and *gpr*, the region of pSA69 containing *kan* and $P_L$lacO1-*alsS* was amplified using primers alsS_yghX_int_for and alsS_yghX_int_rev, which contain 50 bp overhangs that add homology to the E. coli genome between *yghX* and *gpr* (location selected based on [63]). The PCR product was digested with both DpnI and SpeI for 12 h to degrade plasmid DNA. The remaining linear construct was integrated into JCL260 Δ*lysA*::FRT harboring pSIM6 by λ-Red recombineering, following published protocols [84].

The *adhA* and *kivd* genes were amplified from the pSA65 plasmid using primers pstI_adhA_kivd_for and mluI_adhA_kivd_rev. The product was DpnI digested to degrade residual template plasmid and cloned into the pTGD plasmid using restriction enzymes PstI and MluI. pTGD provides a chloramphenicol resistance cassette and flanking 1 kb homology regions to enable CIChE. The CIChE construct was then amplified from the resulting pTGD-*adhA-kivd* using aslB_integr_for and aslB_integr_rev, which add 40 bp regions of homology to the *aslB* locus of the *E. coli* genome. After DpnI digestion, the linear construct was integrated into NV3r1 by λ-Red recombineering using pSIM6 [84] with selection on LB plates with 20 μg/mL chloramphenicol. A resulting integrant (named NV3r1 cm 20) was verified

by PCR and Sanger sequencing. Subsequently, CIChE was performed by growing NV3r1 cm 20 to saturation and passaging into in successively higher concentrations of chloramphenicol (cells were passed 1% v/v and antibiotic concentration was doubled in each passage). To construct JCL260 cm 20 Δ*lysA* pSA69 and JCL260 cm 80 Δ*lysA* pSA69 we prepared P1 lysates from NV3r1 cm 20 and NV3r1 cm 80 and transduced them into JC260 Δ*lysA*::FRT pSA69, selecting on LB plates with 10 μg/mL tetracycline, 50 μg/mL kanamycin and the corresponding concentration of chloramphenicol (20 or 80 μg/mL). Resulting colonies were isolation streaked twice on LB agar with the same antibiotics and 0.8 mM sodium citrate.

2.4.3 Cell preparation for SnoCAP screening

Strains were maintained as glycerol stocks at -80 °C and streaked from cryostocks on LB agar plates with the appropriate antibiotics. Colonies were then picked into liquid LB Lennox with appropriate antibiotics and grown to stationary phase (16-18 h, 37 °C, 250 rpm). For microtiter plate assay, the stationary phase cells were harvested in 1 mL aliquots by centrifugation at 12,000 g for 1 min. The cells were washed twice with 1 mL of 1x M9 salts without amino acids and resuspended to an optical density corresponding to $\sim 10^9$ CFU/mL based on OD_{600} measurement in a VersaMax microplate reader (Molecular Devices).

2.4.4 Media

M9IPG, consisting of M9 salts (47.8 mM Na_2HPO_4, 22.0 mM KH_2PO_4, 8.55 mM NaCl, 9.35 mM NH_4Cl, 1 mM $MgSO_4$, 0.3 mM $CaCl_2$), micronutrients (2.91 nM $(NH_4)_2MoO_4$, 401.1 nM H_3BO_3, 30.3 nM $CoCl_2$, 9.61 nM $CuSO_4$, 51.4 nM $MnCl_2$, 6.1 nM $ZnSO_4$, 0.01 mM $FeSO_4$), thiamine HCl (3.32 μM) and dextrose (D-glucose) at the stated concentrations, was used as the base medium. When specified, 5 g/L yeast extract was added to the medium. Antibiotics were used at the following concentrations: ampicillin, 100 μg/mL; kanamycin, 50 μg/mL; tetracycline, 10 μg/mL; chloramphenicol, 20 or 80 μg/mL. All amino acids were the enantiopure L-isomer, except for norvaline, which was a racemic form (Thermo Fisher

Scientific).

2.4.5 Co-growth assay in microplates

M9IPG with 20 g/L glucose, 3 mM isoleucine, and 50 μg/mL kanamycin was used for all microplates in the 2-KIV screen. No other antibiotics were added. When noted, the medium contained IPTG at 0.1 mM concentration. M9IPG with 4 g/L glucose and no antibiotics were used for all microplates in the tryptophan screening system. When noted, cultures were supplemented with the stated concentrations of tryptophan and histidine. Cells were prepared as described above and then inoculated into the medium. For monocultures, the cells were inoculated 1:100 by volume unless otherwise specified. For co-cultures, each strain was inoculated 1:200 by volume. Cultures were vortexed and then distributed into a 96-well clear microplate (Brand), 200 μL per well. Microplate lids were coated with a solution of 0.5% Triton X-100 in 20% ethanol to reduce condensation and lids were fastened with tape. Microplates were incubated at 37 °C, with shaking in a VersaMax plate reader (Molecular Devices), with absorbance readings at 600 nm taken every 10 min. μmax was calculated via linear regression of natural log of OD_{600} values (after subtracting blank values) vs. time; regression was performed over the time intervals corresponding to early exponential growth phase.

2.4.6 Cultivation for isobutanol production assay

Production assessment was performed similarly to previous studies (e.g., [10]). Overnight cultures in LB with appropriate antibiotics were diluted 1:100 v/v into 10 mL of M9IPG with 36 g/L glucose, 100 μg/mL ampicillin, and 50 μg/mL kanamycin in a 125 mL baffled, unvented polypropylene flasks. The medium was supplemented with either 5 g/L yeast extract or 3 mM lysine. Cells were grown to early exponential phase at 37 °C, 250 rpm, followed by addition of 0.1 mM IPTG. Flasks were then sealed with parafilm and incubated at 30 °C, 250 rpm.

2.4.7 Metabolite detection

Isobutanol and glucose concentrations were assessed by applying filtered culture broth to a Shimadzu high-performance liquid chromatograph (HPLC; model DGU 20A3R) equipped with an autosampler, Phenomenex Rezex ROA Organic Acid H+ (8%) guard and analytical columns (5 mM H_2SO_4; 0.6 ml/min; column temperature, 60 ^oC), and a refractive index detector.

Table 2.1: Strains employed in the studies described in Chapter 2

Strain	**Relevant genotype**	**Reference**
JCL16	BW25113/F′[*traD36*, *proAB+*, *lacIq* ZΔM15 (TetR)]	[10]
JCL260	JCL16 Δ*ldhA* Δ*frd* Δ*adhE* Δ*pta* Δ*fnr* Δ*pflB*	[10]
JCL16 Δ*lysA*	JCL16 Δ*lysA::kan*	This study
JCL260 Δ*lysA*	JCL260 Δ*lysA::kan*	This study
JCL260 Δ*lysA alsS*	JCL260 Δ*lysA*::FRT with $P_L$lacO1::*alsS* integrated between *yghX* and *gpr*	This study
JCL260 Δ*lysA* pSA69	JCL260 Δ*lysA*::FRT pSA69	This study
JCL260 cm X	JCL260 Δ*lysA*::FRT with *adhA-kivd* CIChE construct integrated in the *yihG* site, maintained both on plates and in liquid medium with X mg/mL chloramphenicol	This study
K12 Δ*ilvD*	K12 Δ*ilvD*::kan	This study
Δ*trpB* Keio strain	BW25113 Δ*trpB::kan* (JW1253-1)	CGSC
K12 Δ*hisD*	K12 MG1655 Δ*hisD::kan*	This study
K12 Δ*hisD* Δ*trpR*	K12 MG1655 Δ*hisD*::FRT Δ*trpR::kan*	This study
K12 Δ*hisD* Δ*tnaA*	K12 MG1655 Δ*hisD*::FRT Δ*tnaA::kan*	This study
K12 Δ*hisD* Δ*trpR* Δ*tna*	K12 MG1655 Δ*hisD*::FRT Δ*trpR*::FRT Δ*tnaA::kan*	This study
Δ*hisD* Keio strain	BW25113 Δ*hisD::kan* (JW2002-1)	CGSC
Δ*ilvD* Keio strain	BW25113 Δ*ilvD::kan* (JW5605-1)	CGSC
Δ*ilvE* Keio strain	BW25113 Δ*ilvE::kan* (JW5606-1)	CGSC
Δ*lysA* Keio strain	BW25113 Δ*lysA::kan* (JW2806-1)	CGSC
Δ*ppc* Keio strain	BW25113 Δ*ppc::kan* (JW3928-1)	CGSC
Δ*tnaA* Keio strain	BW25113 Δ*tnaA::kan* (JW3686-7)	CGSC
Δ*trpB* Keio strain	BW25113 Δ*trpB::kan* (JW1253-1)	CGSC
Δ*trpR* Keio strain	BW25113 Δ*trpR::kan* (JW4356-2)	CGSC
Δ*tyrA* Keio strain	BW25113 Δ*tyrA::kan* (JW2581-1)	CGSC

Table 2.2: Plasmids used in the studies described in Chapter 2

Plasmid	**Relevant genotype**	**Reference**
pSA65	ColE1 ori; AmpR; $P_L$lacO1::*kivd–adhA*	[81]
pSA69	p15A ori; KanR; $P_L$lacO1::*alsS–ilvC–ilvD*	[10]
pSIM6	Red expression plasmid; ampicillin-resistant	[84]
pTGD	CIChE plasmid; chloramphenicol- and ampicillin-resistant	[59]
pTGD-*adhA-kivd*	pTGD with $P_L$lacO1::*kivd–adhA* cloned in between the 1 kb homology regions, using enzymes PstI and MluI	This study

Table 2.3: Primers and oligos used in the studies described in Chapter 2

Name	**Sequence 5' → 3'**
p15a_for	TCTGACGCTCAAATCAGTGG
p15a_rev	AGGCGTGGAATGAGACAAAC
alsS_yghX_int_for	AATTTCGAAACAATGTTTCTAGTTTAGCGA TTCGCCAGCGCGTATCCCGTACCGAGCG TTCTGAACAAAT
alsS_yghX_int_rev	GCATAAGCACGTATTTTTGCCCAGTTTTTC GTCACTCTGTGAGCCAGACTGGTGATTCC TCGTCGACCTA
alsS_int_chk_front_for	ACCTCTCCTTTCCACCGTTC
alsS_int_chk_front_rev	TCGCCTTCTTGACGAGTTCT
alsS_int_chk_back_for	GGGGAACTCATGAAAACGAA
alsS_int_chk_back_rev	GAGATTTTCCCGTGAGCGTA
pstI_adhA_kivd_for	CTACTGCAGAATTGTGAGCGGATAACAAT
mluI_adhA_kivd_rev	AACTACGCGTACAACAGATAAAACGAAAGG
aslB_integr_for	ATGCGTCAGCATCGCATCCGGCAAAGGCAG ATC TCAGCGACGAGGCAGCAGATCAATTCG
aslB_integr_rev	CCACCACGCGCGCAGATTAAATCTGACTAAG CCGGCGCTAGCTACGGCGTTTCACTTCTG

Chapter 3: Higher-throughput formats of SnoCAP

3.1 Introduction

A microplate-based screen's throughput can be extended by miniaturization and automation, but generally becomes limited by space and setup time to no more than $\sim 10^4$ assays per experiment. Automation of the plate handling also requires costly specialized equipment. We were therefore interested in exploring other formats of compartmentalization of individual secretor cells with the sensor strain. To this end, we investigated implementation as a colony screening assay and as a microfluidic droplet assay.

The use of water-in-oil droplets as miniature bioreactors has long been an attractive approach and dates back to at least the 1950s [85]. The development of droplet microfluidics has led to high-throughput techniques for generation and manipulation of monodisperse droplets. This has enabled high-throughput screening for a variety of applications, with low space and reagent requirements. Use of droplet microfluidics for microbiology applications, including biotechnology, is reviewed in [86]. Droplet sorting, either via microfluidic sorting devices [87] or commercial flow cytometry, has proved effective at screening enzyme variant libraries for directed evolution. It has also been applied to screening whole cell libraries for production level. With current droplet sorting technology, the highest throughput of sorting is achieved using fluorescence signal. Strategies for coupling strain production to fluorescence for droplet screening have included encapsulation with oxidase enzymes and horseradish peroxidase [88], encapsulation with a sensor strain with a transcription-factor sensor [89], encapsulation with Spinach aptamers [90], inherently fluorescent products [91], and staining with fluorescent dyes [92].

To increase the throughput of SnoCAP, we have explored higher throughput compartmentalization strategies for the SnoCAP screening strategy presented in Chapter 1: spatial separation on agar plates and encapsulation in water-in-oil microdroplets. While microplate assays reach throughputs of 10^2-10^4 assays/experiment, these formats can reach orders of higher throughputs (10^2-10^4 assays/experiment on agar plates, and 10^4-10^5 assays/experiment in microdroplets). The higher throughput formats also offer other advantages, such as smaller space and capital cost requirements and shorter incubation times. We applied the agar plate format of the 2-KIV screening to a library of chemically mutagenized JCL260 Δ*lysA* *alsS* and identified a strain that can produce a 5-fold higher titer of isobutanol than the parent strain. We sequenced the genome of this strain and identified potential mutations that may contribute to the improved phenotype. We reconstructed a mutation in the substrate binding loop of the *aceK* gene and determined that it led to a small but statistically significant improvement in isobutanol production. We then implemented SnoCAP by co-cultivation in microfluidically generated droplets. We encapsulated model libraries consisting of low percentages of the high-producing strain and demonstrated that we could isolate the rare high-producers by fluorescent-activated droplet sorting (FADS).

The majority of this work has been submitted for publication, in the same manuscript as Chapter 2.

3.2 Results

3.2.1 Implementation of co-culture screening in an agar plate format

We implemented SnoCAP as a colony-screening assay by spreading ∼10^7 CFU of the sensor strain (enough to form a lawn were the cells able to grow in monoculture) and ∼100 CFU of secretor strain on agar medium in 10 cm diameter Petri dishes. This produced mixed colonies, each originating from a single secretor cell. We verified that these colonies were

indeed mixtures of secretor and sensor by streaking them on LB plates with IPTG/X-gal and observing both blue (sensor) and white (secretor) colonies. We also observed no growth on monoculture plates containing only one strain, demonstrating that colonies only form when both sensor and secretor are present. We compared mixed colony formation between the base strain (JCL16 Δ*lysA*) and the high-producing strain (JCL260 Δ*lysA* pSA69). JCL16 colonies appeared later and ultimately formed flatter, more translucent mixed colonies that

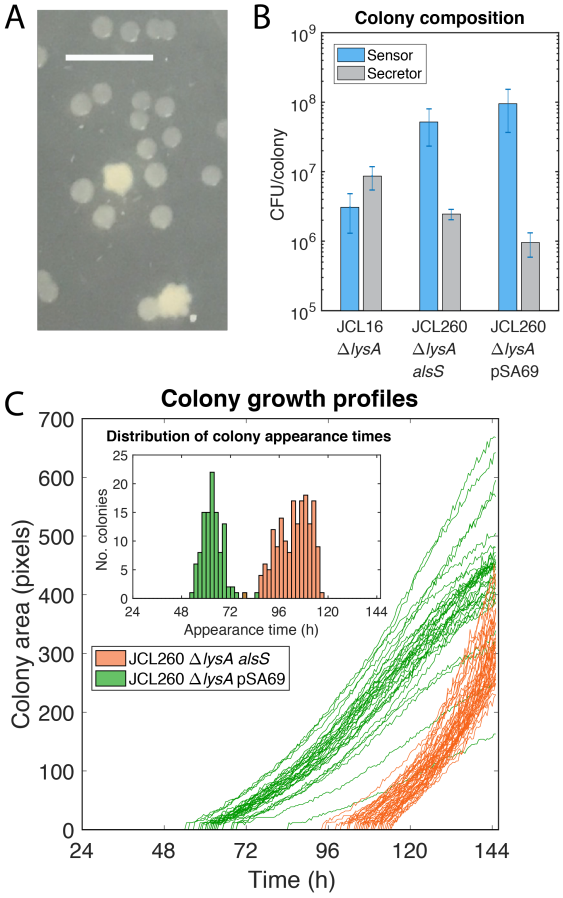

Figure 3.1: 2-ketoisovalerate screening implementation on agar plates. (A) Mixture of JCL16 Δ*lysA* and JCL260 Δ*lysA* pSA69 mixed colonies after 6 days at 37 °C. More opaque white colonies contain JCL260 Δ*lysA* pSA69, while more translucent colonies are JCL16 Δ*lysA*. Scale bar represents 0.5 cm. (B) Composition of colonies after 7 days incubation at 35 °C. Error bars represent the standard deviations of 4 colony replicates, with 4 technical replicates each. (C) Colony area over time for the plates. Growth curves are shown from 55 colonies on a JCL260 Δ*lysA* *alsS* plate and 48 colonies on a JCL260 Δ*lysA* pSA69 plate. The inset shows a histogram of colony appearance times, combined from three plates of each secretor strain (126 total colonies from JCL260 Δ*lysA* pSA69 plates and 110 colonies from JCL260 Δ*lysA* *alsS* plates).

are easily distinguishable from the taller, opaque JCL260 Δ*lysA* pSA69 mixed colonies, even

once both have grown to a substantial footprint area (Fig. 3.1A). Higher-production-level secretors also produced colonies with increased sensor-to-secretor ratio (Fig. 3.1B). When the composition of the mixed colonies on plates with norvaline was assessed, it was found that on 0.5 g/L norvaline, JCL16 Δ*lysA* produced colonies that were majority secretor, while JCL260 Δ*lysA* *alsS* and JCL260 Δ*lysA* pSA69 produced ones that were majority sensor. On 1.0 g/L norvaline, both JCL16 Δ*lysA* and JCL260 Δ*lysA* *alsS* produced colonies that were mainly secretor cells, while JCL260 Δ*lysA* pSA69 still produced colonies that were majority sensor. We next compared the growth of the intermediate level strain (JCL260 Δ*lysA* *alsS*) and the high production strain (JCL260 Δ*lysA* pSA69). Here, both strains eventually formed large opaque colonies, but JCL260 Δ*lysA* *alsS* colonies appeared later. We implemented the ScanLag technique for automated imaging and colony growth profile analysis developed by Levin-Reisman *et al.* [93,94], to observe colony lag time and growth dynamics. This inexpensive method uses photo scanners to image the Petri dishes periodically during growth and a MATLAB-based application that aligns the images and returns colony growth phenotype information. The analysis revealed that JCL260 Δ*lysA* *alsS* mixed colonies appear later and then grow to eventually reach similar colony size as JCL260 Δ*lysA* pSA69 mixed colonies (Fig. 3.1C).

For the tryptophan system, we compared the Δ*trpR* strain and the base strain in the agar plate-screening assay. Similar to the microplate, colonies were slow to develop on plates that were not supplemented with any amino acids, but after two weeks we observed colonies on the Δ*trpR* secretor co-culture plates and none on the base secretor co-culture plates, and none on any monoculture plates. Addition of small initial amounts of histidine produced a significant decrease in time required for colony visibility while maintaining clear differences in the growth characteristics between the two secretor strains (Fig. 3.3A, B). In this case, both the base and improved secretor strains formed colonies, but the base strain colonies were flat, translucent, and easily distinguishable from the taller, opaque Δ*trpR* secretor colonies.

For the 2-KIV, we tested the addition of norvaline to the agar plates and found that it

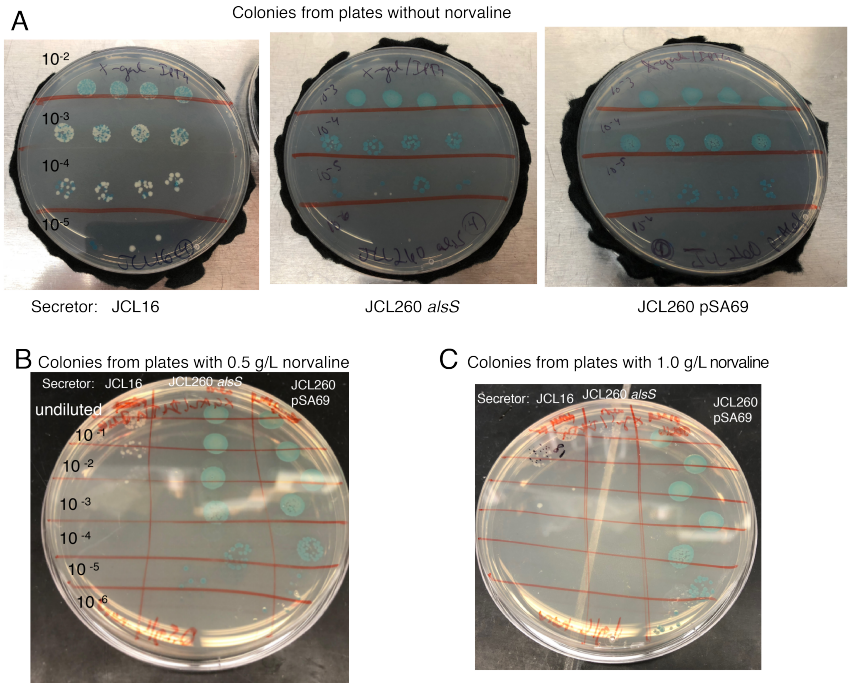

Figure 3.2: Population composition determination of mixed colonies by plating dilutions of the colony on LB plates with IPTG/X-gal. Blue colonies are sensor cells; white colonies are secretor cells. Spots from serially diluted colonies from plates with no (A), 0.5 g/L (B) and 1.0 g/L (C) norvaline are shown. Spots of the dilutions were also plated on LB plates with tetracycline to determine secretor strain CFU concentration compared to total CFU concentration.

was also effective in this format in widening the difference in growth between co-cultures containing the intermediate- and high-producing secretor strains (Fig. 3.4). At 0.5 g/L norvaline, only the highest producing strain (JCL260 Δ*lysA* pSA69) formed opaque colonies, while the intermediate-producer (JCL260 Δ*lysA alsS*) developed smaller, translucent colonies that were not visible in the scanner images.

We also tested the push-pull arrangement on agar plates. We observed similar trends to those in the microplate, with the mixed colonies appearing later for secretor strains with higher Kdc/Adh expression levels (3.5D, E).

3.2.2 Screening model libraries in the agar plate format

To evaluate the ability of the agar plate screening assay to identify rare higher-producing strains, we tested model libraries consisting of the intermediate-producing strain spiked with

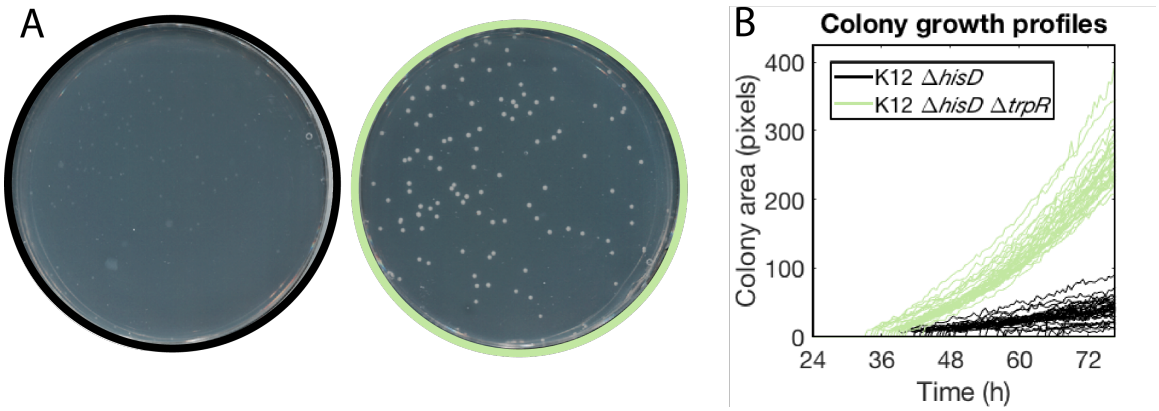

Figure 3.3: Tryptophan screening implementation on agar plates. Images after 72 h incubation (A) and colony growth profiles (B) of agar plate co-cultures comparing secretor strains K12 Δ*hisD* and K12 Δ*hisD* Δ*trpR*, with 5 μM initial histidine.

a small percentage of the high-producing strain. On each plate, we observed the development of many small translucent colonies as well as a smaller number of large opaque colonies. We isolated the secretor from the large colonies and investigated the strain identity.

In one experiment, we spread a square 24.5 cm plate with 102 ± 25 LB-CFU of JCL260 pSA69 and 18,200 ± 500 of JCL260 Δ*lysA* *alsS* (based on plating the inocula on LB plates; SD, n = 2). This amounts to a library consisting of 0.6% high secretor, 99% medium secretor strain. After 7 days, 25 large colonies were apparent and are numbered in yellow in Fig. S5. 17 intermediate-sized colonies were also observed. The four that were further investigated are numbered in blue (Fig. 3.6)A).

Numbered colonies, as well as 4 small colonies, were streaked on LB plates with tetracycline and kanamycin to isolate secretor strain. A single colony from each of these plates was assayed by PCR using primer sets alsS_int_chk_front_for/alsS_int_chk_front_rev, which produce a short band only if *alsS* is integrated, and alsS_int_chk_front_for/alsS_int_chk_back_rev, which produce a short band only if *alsS* is not integrated (Fig. 3.6D). 24 of the 25 large colonies were identified as JCL260 Δ*lysA* pSA69. Colony 13, the only large colony which was identified as containing JCL260 Δ*lysA* *alsS*, was a particularly large colony, so we suspected it could have merged with surrounding small colonies, leading to a mixture of secretor strains. We repeated the streak out of Colony 13 on an LB tetracycline plate, picking cells

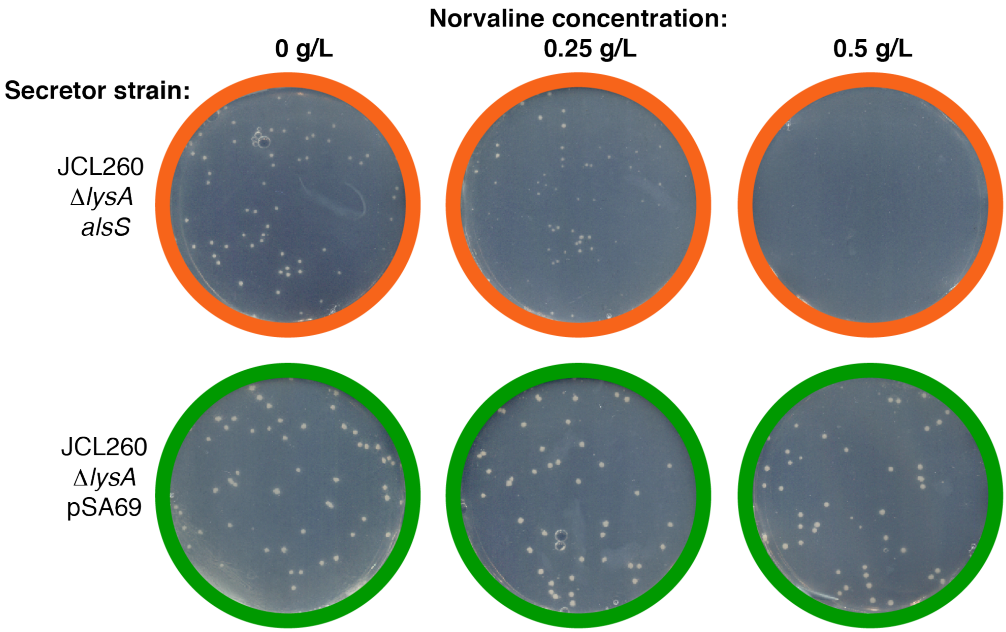

Figure 3.4: Norvaline addition to agar plates for 2-KIV screening. Agar plate co-cultures containing various concentrations of norvaline, after 168 h incubation at 35 °C. At 0.5 g/L norvaline, JCL260 Δ*lysA* *alsS* plate has no visible colonies.

from the very center of the colony. When we repeated the PCR assay on 10 colonies from this streak, we found that 9 of 10 were JCL260 Δ*lysA* pSA69 and only one was JCL260 Δ*lysA* *alsS*.

The four small colonies that were chosen were all identified as JCL260 Δ*lysA* *alsS*, as expected. Of the four intermediate-sized colonies that were investigated, three were JCL260 Δ*lysA* pSA69 and one (colony 30) was JCL260 Δ*lysA* *alsS*. Colony 30 is quite close to colony 6, so it is possible that diffusion between colonies in close proximity is occurring, which should be taken into consideration when selecting colonies.

A 0.1% model library was also tested. We plated 9.1 ± 1.6 LB-CFU of JCL260 Δ*lysA* pSA69 and 78,500 ± 70 LB-CFU (SD, n = 2) of JCL260 Δ*lysA* *alsS* on each of two square plates. After six days incubation at 37 °C, one plate had developed one large colony and the other developed two. All three large colonies were found to contain JCL260 Δ*lysA* pSA69 as the secretor strain.

We concluded that, when an appropriate level of norvaline is utilized for the range of

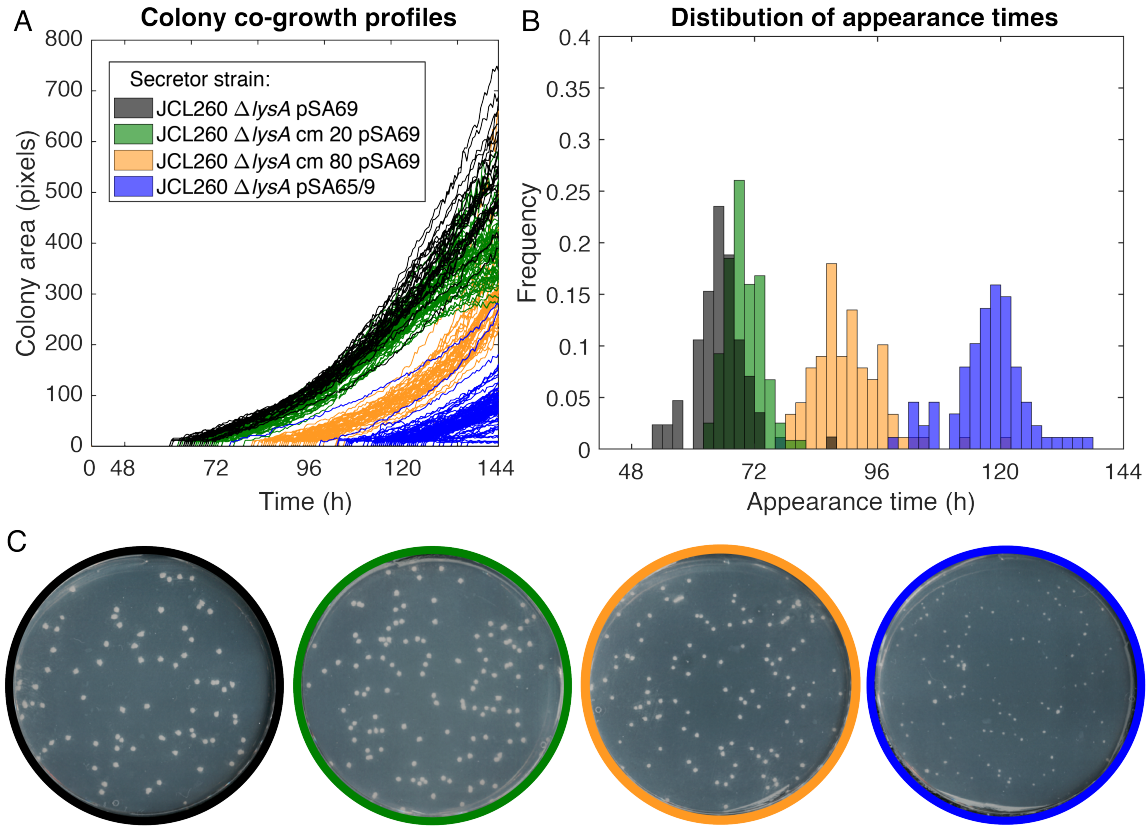

Figure 3.5: Push-pull screening in agar plate format. Colony growth profiles (A) and colony appearance time histogram (B) for agar plate assay comparing the four secretor strains. Plot in (A) shows profiles for the unmerged colonies from one plate of each secretor strain (29, 50, 34, 84 for JCL260 Δ*lysA* pSA69, JCL260 Δ*lysA* cm 20 pSA69, JCL260 Δ*lysA* cm 80 pSA69, JCL260 Δ*lysA* pSA65/9, respectively) (the JCL260 Δ*lysA* pSA65/9 plates have more unmerged colonies because colony size is smaller). The histogram in (B) shows appearance times for the unmerged colonies from a combination of two plates from each strain except for JCL260 Δ*lysA* pSA65/69 for which colonies from one plate are included since smaller colony size resulted in reduced merging and more analyzable colonies per plate. In total, this amounts to 85 from JCL260 Δ*lysA* pSA69, 119 from JCL260 Δ*lysA* cm 20, 89 from JCL260 Δ*lysA* cm 80, and 84 from JCL260 Δ*lysA* pSA65/9. (C) Plate images after 144 h incubation. Legend in (A) also applies to (B) and (C).

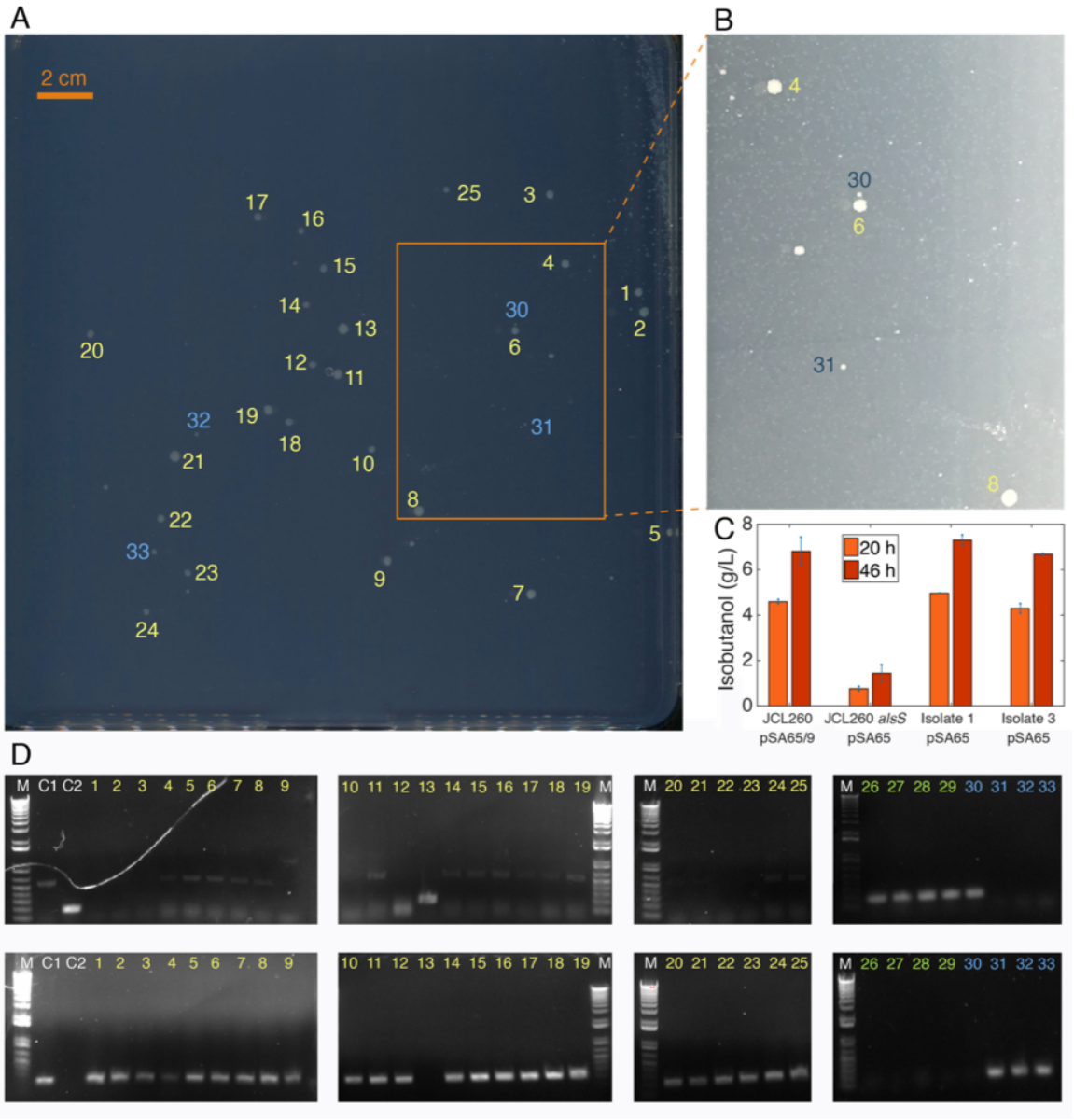

Figure 3.6: Model library on an agar plate. (A) 0.6% library plate after incubation at 37 °C for 7 days, imaged from below on photo scanner. Colonies that were further examined are numbered. (B) A close-up of the plate, photographed from above. Many small colonies are visible surrounding the large colonies. (C) Production levels of secretor strain isolated from colonies 1 and 3 in the screen after transformation with pSA65, in comparison with JCL260 Δ*lysA* pSA65/9 and JCL260 Δ*lysA* *alsS* pSA65. Error bars represent standard deviation between two biological replicates. (D) Verification that large colonies in the model library are JCL260 Δ*lysA* pSA69 and not JCL260 Δ*lysA* *alsS*. Gel electrophoresis of PCR reactions with either primers alsS_int_chk_front_for and alsS_int_chk_front_rev (top) or alsS_int_chk_front_for and alsS_int_chk_back_rev (bottom). M is Invitrogen 1 Kb plus ladder, C1 is JCL260 Δ*lysA* pSA69, C2 is JCL260 Δ*lysA* *alsS*. 1-25 are secretor strain colonies isolated from the 25 largest mixed colonies. 26-29 are secretor strain isolated from four small colonies from the model library plate. 30-33 are intermediate-sized colonies. Numbers correspond to the numbering shown on (A), except for the small colonies which are not visible in that image.

strain improvement that is targeted, the screening can easily identify rare higher production level strains at frequencies as low as 0.1% (or lower if a larger number of plates are employed). Transformation with pSA65 and production testing of secretor cells isolated from the colonies confirmed that the strain is not adversely affected by the screening conditions (Fig. 3.6C).

3.2.3 Screening of a chemically-mutagenized strain library for improved plasmid-free isobutanol production

We next applied the SnoCAP screening method to strain development for higher 2-KIV production based on genomic modifications (rather than by pathway overexpression from the pSA69 plasmid). We introduced random mutations into the genome of JCL260 Δ*lysA* *alsS* by mutagenesis with N-Methyl-N'-nitro-N-nitrosoguanidine (NTG). NTG is a chemical mutagen that adds alkyl groups to the O^6 of guanines and, to a lesser extent, the O^4 of thymines, leading to mispairing during DNA replication and thus mutations. We plated the library on 24.5 cm square Petri dishes containing various concentrations of norvaline, along with an excess of sensor strain. 10^4 LB-CFU (cells that would form colonies in monoculture on an LB plate) of secretor strain library were spread on each plate. After seven days of incubation at 37 °C, we identified large opaque colonies among the large number of small translucent colonies. We streaked large colonies from the 1.0 g/L norvaline condition on LB plates with tetracycline to select for the secretor strains and then rescreened them in the microplate co-culture assay and selected 7 isolates that showed improved co-culture growth. We transformed these isolates with pSA65 and tested their isobutanol production. Due to reduced growth rates observed in some of the library isolates, we tested their production levels in M9IPG supplemented with 5 g/L yeast extract. Of the 7 isolates tested, one, which we call strain B1, showed significantly improved production compared to the base strain. After 73 h fermentation, B1 pSA65 produces 9.4 ± 0.4 g/L (SD, n = 3) isobutanol, representing 60% of the theoretical yield, compared to 1.8 ± 0.1 g/L and 16% theoretical yield by the parental strain JCL260 Δ*lysA* *alsS* (Fig. 3.7). It should be noted that JCL260

Δ*lysA* *alsS* performs less well under yeast extract supplemented conditions compared to minimal medium conditions, likely due to lower expression of the ilvCD genes when it is not necessary to produce all of its own amino acids. Nevertheless, B1 isolate performs superiorly even to JCL260 Δ*lysA* *alsS*'s production level in minimal medium (i.e., the production level presented in Fig. 2.2B). B1 performs nearly as well as JCL260 Δ*lysA* pSA69, but through a different mechanism than plasmid overexpression of the *ilvCD* genes.

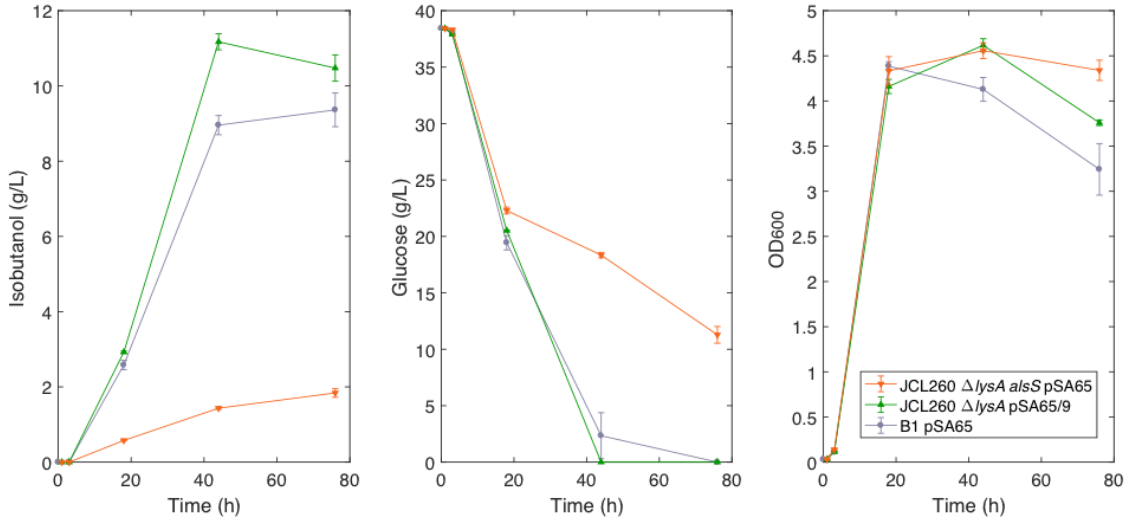

Figure 3.7: Production culture performance of mutagenesis library isolate B1 compared to parent strain JCL260 Δ*lysA* *alsS* and high-producing strain JCL260 Δ*lysA* pSA69. Error bars represent standard deviations of biological replicates (n = 3).

To investigate the mutations leading to this improved production, we sequenced the genome of B1 and parental strain with >100X coverage and identified 73 SNPs, 39 of which lead to amino acid substitutions or stop codon introductions, in B1. The latter set of mutations are listed in Table 3.1. Interestingly, most of the mutations in this strain are clustered in the 3.1-3.3 Mb and 3.8-4.6 Mb regions. There is a strong bias for G/C to A/T transitions, as has been previously reported for NTG-mutagenesis in *E. coli* (e.g., [95]).

One intriguing mutation occurs in the *aceK* gene, which encodes the isocitrate dehydrogenase kinase/phosphatase. This bifunctional enzyme controls the branch between the tricarboxylic acid (TCA) cycle and glyoxylate cycle by modification of isocitrate dehydroge-

nase. B1's mutation consists of a proline to serine substitution in residue 510, which is part of the substrate recognition loop [96]. We reintroduced this mutation into a *mutS*- version of the parental strain by single-stranded oligo recombination, calling this strain JCL260 Δ*lysA alsS aceK-mut*. When tested in liquid co-culture with the sensor strain, JCL260 Δ*lysA alsS aceK-mut* shows improved co-growth compared to the parental strain (Fig. 3.8A). After transformation with pSA65, we compared the production and found that this mutation does, indeed, lead to a modest increase in isobutanol production (Fig. 3.8B).

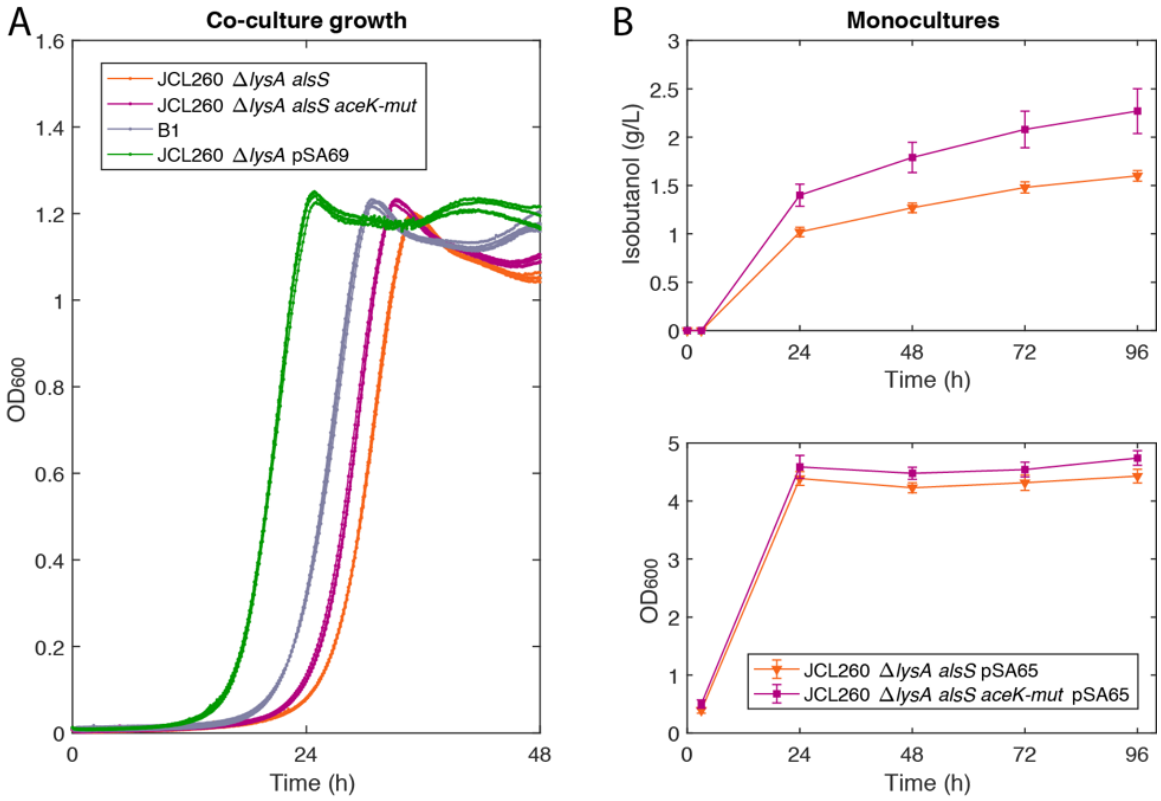

Figure 3.8: *aceK* mutant strain characterization. (A) Growth profiles of secretor strains co-cultured with K12 ilvD sensor strain in a 96-well microplate with 0.1 g/L norvaline (n = 4, replicates plotted in the same color). (B) Monoculture isobutanol production and growth profiles of *aceK* mutant strain and parental strain after transformation with pSA65. Error bars represent the standard deviation of biological replicates (n = 10). For isobutanol production at the 24 hr time point, two-tailed P-value < 0.0001.

AceK activity is required for growth on acetate [97]. We found that both B1 and JCL260 Δ*lysA alsS aceK-mut* were able to grow in an M9 medium with 20 g/L acetate as the sole carbon source (supplemented with lysine), which indicates that the mutation

does not completely abolish AceK activity. Knockout of *aceK* has previously been used to increase L-carnitine production [98].

Another mutation of potential interest occurs in the *ilvC* gene (part of the isobutanol pathway): a glycine to aspartate mutation at amino acid position 442. Although this is an unstructured region of the protein, it may still merit further investigation since the gene is part of the production pathway.

Table 3.1: SNP mutations identified in the genome of strain B1, which was derived by random mutagenesis and SnoCAP screening and has improved 2-KIV/isobutanol production

Mutation coordinate in BW25113 [99]	**Nucleotide substitution**	**Gene**	**Gene annotation**	**Amino acid substitution**
284584	C→T	*xynR*	CP4-6 prophage; DNA-binding transcriptional repressor XynR	G97E
1014968	C→T	*ompA*	outer membrane porin A	G181D
1101625	C→T	*ymdB*	2'-O-acetyl-ADP-ribose deacetylase, regulator of RNase III activity	A117V
2349193	C→T	*glpC*	anaerobic glycerol-3-phosphate dehydrogenase subunit C	A65V
3118229	C→T	*glcF*	glycolate dehydrogenase, putative iron-sulfur subunit	G197E
3119643	C→T	*glcE*	glycolate dehydrogenase, putative FAD-binding subunit	S80N
3158414	C→T	*parC*	dimer of DNA topoisomerase IV subunit A	D307N
3202308	C→T	*ttdT*	L-tartrate:succinate antiporter	A309V
3203862	C→T	*tsaD*	N6-L-threonylcarbamoyladenine synthase, TsaD subunit	G14D
3247697	C→T	*yhaK*	bicupin-related protein	A7V
3250346	C→T	*cyuP*	putative D/L-serine transporter	A342T
3252440	C→T	*tdcG*	L-serine deaminase III	G190D
3256586	C→T	*tdcD*	propionate kinase	S145N
3264955	C→T	*garK*	glycerate 2-kinase 1	A59T
3830769	C→T	*setC*	putative arabinose exporter	L153F
3837065	C→T	*adeQ*	adenine transporter	V29I
3952654	G→A	*ilvC*	ketol-acid reductoisomerase (NADP+)	G442D
3952951	T→C	*ppiC*	peptidyl-prolyl cis-trans isomerase C	T75A
3963542	C→T	*wecB*	UDP-N-acetylglucosamine 2-epimerase	A17V
3988970	G→A	*yigA*	conserved protein YigA	E10K
3992126	G→A	*uvrD*	ssDNA translocase and dsDNA helicase - DNA helicase II	E262K
4026380	C→T	*yigZ*	IMPACT family member YigZ	L177F
4132894	C→T	*frwC*	putative PTS enzyme IIC component FrwC	T146I
4195232	C→T	*purD*	phosphoribosylamine—glycine ligase	A210T
4210051	C→T	*aceK*	isocitrate dehydrogenase kinase / isocitrate dehydrogenase phosphatase	P510S
4245002	C→T	*plsB*	glycerol-3-phosphate 1-O-acyltransferase	G465S
4264125	C→T	*ssb*	ssDNA-binding protein	P25S
4269416	C→T	*ghxP*	guanine/hypoxanthine transporter GhxP	A337V
4313755	C→T	*phnD*	phosphonate ABC transporter periplasmic binding protein	A139T
4333393	C→T	*melB*	melibiose:H+/Na+/Li+ symporter	R70STOP
4334603	C→T	*melB*	melibiose:H+/Na+/Li+ symporter	A474V[a]
4338165	C→T	*dcuB*	anaerobic C4-dicarboxylate transporter DcuB	A133T
4382187	C→T	*glyY*	tRNA-Gly(GCC)	
4388964	C→T	*mutL*	DNA mismatch repair protein MutL	T579I
4430279	C→T	*ytfL*	putative inner membrane protein	D252N
4515057	C→T	*yjhI*	KpLE2 phage-like element; putative DNA-binding transcriptional regulator YjhI	W188STOP
4545803	C→T	*yjiC*	uncharacterized protein YjiC	G112D
4556915	C→T	*yjiN*	DUF445 domain-containing protein YjiN	G50D
4565276	C→T	*yjiV*	putative uncharacterized protein YjiV	S442F
4587974	C→T	*opgB*	phosphoglycerol transferase I	A429T
4627746	C→T	*creC*	sensory histidine kinase CreC	P412S

[a]This mutation is downstream of the introduced stop codon

3.2.4 Microdroplet co-cultivation and sorting implementation

To further increase the throughput of the SnoCAP screening framework, we next investigated compartmentalization by encapsulation in microfluidic water-in-oil droplets. These monodisperse droplets, with volumes in the picoliter to nanoliter range, provide miniaturized culture volumes that can be analyzed in a variety of ways, including by high-throughput automated sorting to isolate droplets containing the highest fluorescence signal. Cells are distributed according to a Poisson distribution, and cell density can be manipulated to ensure that initially i) all droplets contain several sensor cells, and ii) the majority of droplets contain either zero or one secretor cell.

To couple the co-growth output to fluorescence, we expressed fluorescent proteins in the strains. We first labeled the high secretor, JCL260 $\Delta lysA$ pSA69, with YFP, the sensor strain with mCherry, and observed co-culture growth in droplets (Fig. 3.9A). For sorting, we employed a version of the sensor strain carrying a plasmid encoding constitutively expressed mNeonGreen and screened by FADS for droplets with the highest fluorescence, corresponding to the largest number of sensor cells.

We observed that the sensor strain cells tend to become very long (sometimes >50 μm in length), while the high-secretor cells generally remain a normal length. This may be due to sensitivity to norvaline. At high norvaline concentrations we also sometimes observed what appeared to be physical associations between the cells (Fig. 3.10). We hypothesize this could be cross-feeding by intercellular nanotubes, as in [72]. Pande *et al.* found that this type of cross-feeding led to an exchange of fluorescent proteins between the two types of cells, so there is a possibility that that could be occurring in our system, although when both cells were differentially labeled we visually observed individual cells to show only one type of fluorescence strongly. This type of cross-feeding would allow more direct transfer of the cross-feed metabolites, possibly allowing better tolerance of the norvaline. These lengthened cells and physical associations were more prevalent when the medium was not freshly made the day of the experiment.

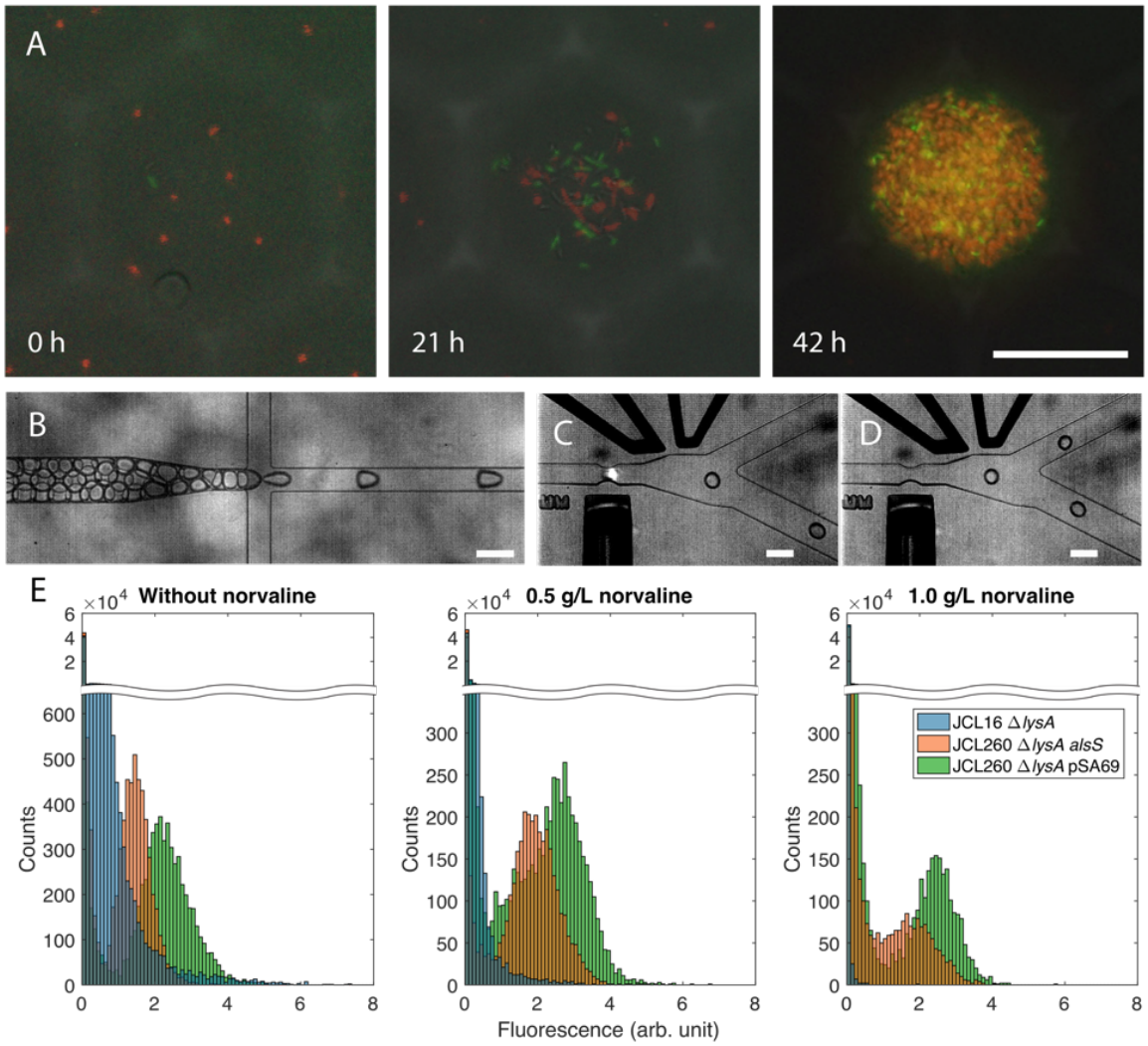

Figure 3.9: Screening for 2-KIV production implemented in microfluidic droplets. (A) Representative droplet images during co-growth of sensor strain expressing mCherry and JCL260 Δ*lysA* pSA69 expressing YFP on 1.0 g/L norvaline. Each image is an overlay of fluorescence and bright-field images, but cell locations do not align exactly due to cell movement between images. (B) Droplet reinjection and spacing. (C, D) Fluorescence-activated droplet sorting (FADS) to retrieve droplets with high levels of sensor strain growth. (E) Histograms comparing fluorescence signal from droplets containing co-cultures with different secretor strains (at an initial cell loading of 0.2 secretor and 5 sensor/droplet) on several concentrations of norvaline after a 30 h incubation. ∼50,000 droplets under each condition were analyzed. Scale bars: (A) 50 μm, (B-D) 100 μm.

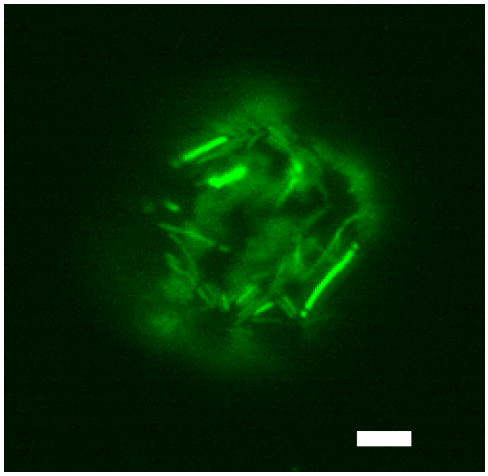

Figure 3.10: Droplet of diameter ∼55 μm containing medium with 1.5 g/L norvaline and a co-culture of K12 Δ*ilvD* pSAS31 sensor strain and unlabeled JCL260 Δ*lysA* pSA69 high-level secretor. In this droplet, the cells appear to physically associate. Image shows green fluorescence, which is indicative of the sensor cell. Variation in cell length and fluorescence intensity is visible between cells. Scale bar: 10 μm

We tested the ability of FADS to distinguish between two strains of differing production levels and to isolate the high secretor, JCL260 Δ*lysA* pSA69, when spiked at low percentages into a population of lower secretor cells. We identified conditions (i.e., an appropriate norvaline concentration) that enabled the desired level of separation between the two strains, encapsulated model libraries consisting of mixtures of the two secretor strains, incubated to allow co-growth with the sensor strain within the droplets, and then sorted to isolate the most fluorescent droplets. In order to monitor the collective growth in the droplets and determine what stage of growth most droplets are in, we incubated ∼100 μL of droplets in a microplate, incubated it in a microplate reader set to measure fluorescence every 15-20 min. We compared fluorescence of droplets incubating in 1.5 mL Eppendorf tubes compared to those incubating in the microplate reader and found that they were comparable (Fig 3.11). The cell suspension for encapsulation was prepared so that one would expect an average of 0.1 secretor cell/droplet if droplets were 50 μm in diameter and the same cell suspension was used for the three different droplet sizes, which may explain why the smaller droplets reach a lower total fluorescence than the larger droplets.

We mixed JCL260 Δ*lysA* pSA69 and JCL260 Δ*lysA* *alsS* *galK::cat* (*galK* deletion

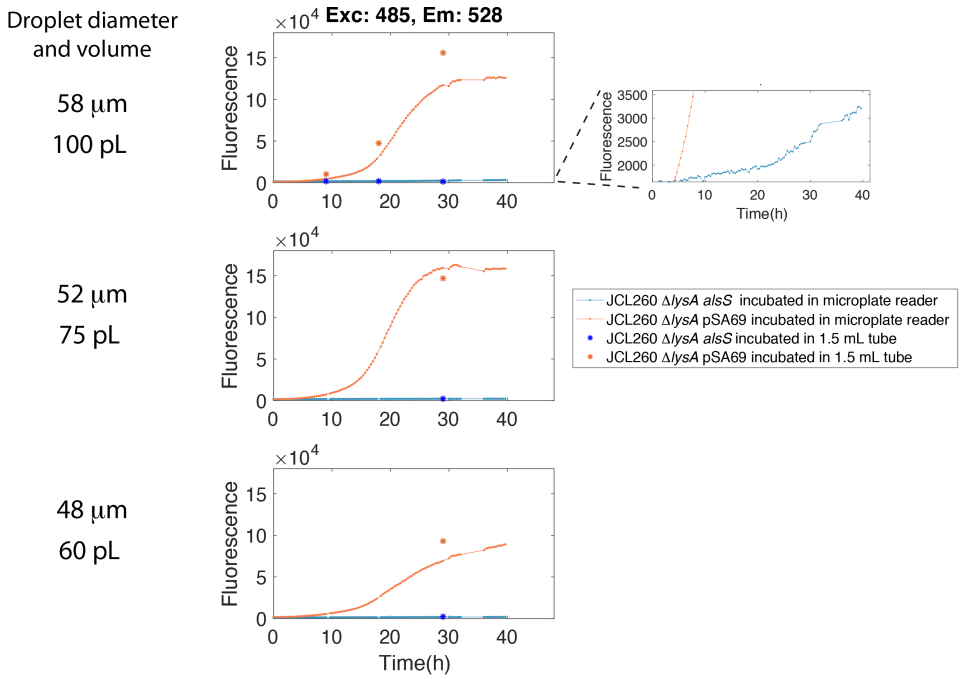

Figure 3.11: Collective droplet fluorescence profiles over time. Fluorescence readings in a microplate reader of droplets of various sizes, either incubated within the reader, taking measurements every 20 min or incubated in a 1.5 mL Eppendorf tube and transferred into a microplate for readings.

enabled color-based differentiation of the strains when plated on MacConkey agar with galactose, Fig. 3.12B) at a ratio of 1:100 and encapsulated them with sensor strain K12 Δ*ilvD::kan* pSAS31 on 1.75 g/L norvaline. We used cell densities such that the secretor cell loading was ∼0.1/droplet (cells are encapsulated according to Poisson distribution, with ∼90% containing no secretor) and sensor cell loading of ∼5/droplet. Following incubation to allow co-culture growth, the droplets were reinjected into a sorting device (Fig. 3.9B) and sorted based on fluorescence from excitation with a 450 nm laser (Fig. 3.9C,D). We selected sorting gate values based on comparisons of the signal profiles from mono-secretor control sets of droplets (Fig. 3.9E) and verified their effectiveness by observing that these values enabled bright droplets to be collected. For application to a library when a high-producing strain is not already available, the threshold can be selected to sort the top percentage of droplets at a desired stringency. After sorting, desired droplets were collected into a poly-dimethylsiloxane (PDMS) device with an elevated chamber to retain droplets while allowing

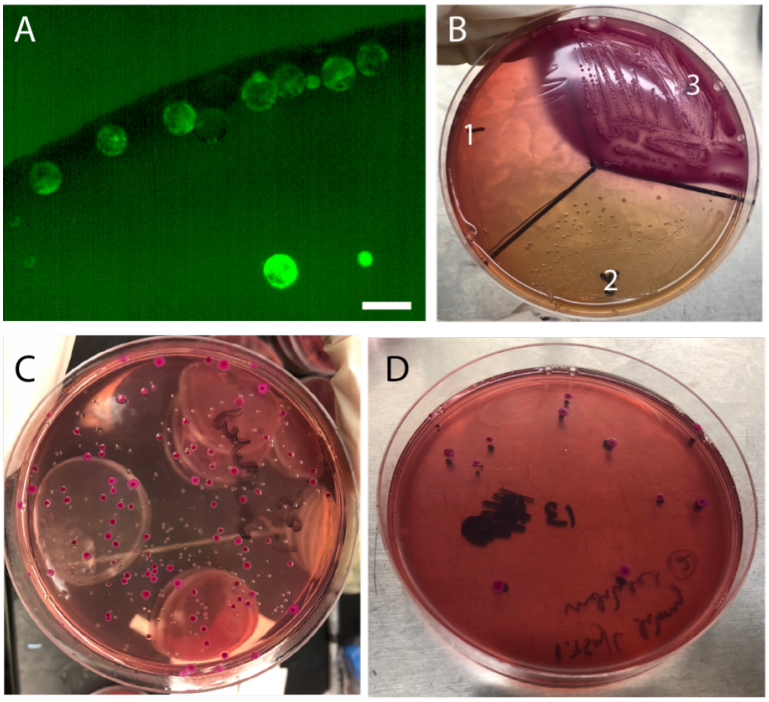

Figure 3.12: Retrieval of viable cells after droplet sorting. (A) Droplets collected after 25 min sorting of droplets encapsulating a mixture of JCL260 Δ*lysA* *alsS* and JCL260 Δ*lysA* pSA69 secretor strains at a 100:1 ratio. Scale bar: 100 μm. (B) Strain growth on MacConkey agar plates containing tetracycline. K12 Δ*ilvD* pSAS31 (1) does not grow. JCL260 Δ*lysA* *alsS* *galK::cat* (2) produces white colonies, and JCL260 Δ*lysA* pSA69 (3) produces purple colonies. (C) Colonies from plating unsorted droplets. (D) Colonies from plating the collected droplets that are pictured in (A).

oil to flow out. In a 25 min sorting period, we isolated 10 droplets, each of which showed significant sensor cell growth. Plating these droplets on a MacConkey agar plate with tetracycline (to prevent growth of the sensor strain) produced only purple colonies (>100-fold enrichment of the high-secretor) (Fig. 3.12D), whereas plating the unsorted droplets produced a mixture with 16% purple colonies (16-fold enrichment) (Fig. 3.12C). We assessed the productivity of purple colonies that resulted from sorted droplets after transformation with pSA65 and found that they maintained their productivity through the screening process (Fig. 3.13).

We noticed lower than expected cell recovery from the collected droplets. We also observed that the collected droplets begin to shrink after the collection device is disconnected from the sorting device and that shrinkage is less severe when larger numbers of droplets are

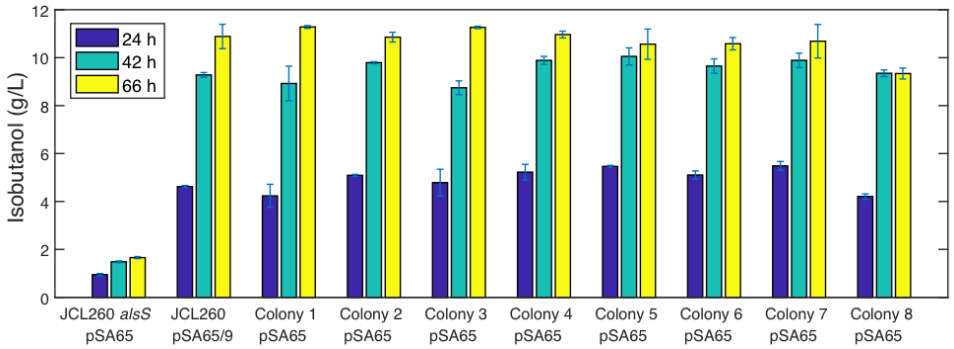

Figure 3.13: Productivity after retrieval from droplets. Isobutanol production of cells from eight randomly selected purple colonies isolated from sorting of model libraries and transformed with pSA65, compared to JCL260 Δ*lysA* *alsS* pSA65 and JCL260 Δ*lysA* pSA65/9. Error bars represent standard deviation of two biological replicates.

collected. We assessed whether viability is affected by the sorting process itself by running cell-containing droplets through the device either into the waste channel (electrode turned off) or into the collection channel (electrode turned on), collecting the droplets in Eppendorf tubes and assessing the cell viability. We found no significant difference between the droplets that had not been reinjected, the waste channel droplets, and the positive channel droplets ($1.7\text{x}10^4 \pm 3.4\text{x}10^3$, $2.0\text{x}10^4 \pm 3.2\text{x}10^3$, and $1.6\text{x}10^4 \pm 88$ CFU/μL droplets (SD, n = 2), respectively). We, therefore, conclude that droplet shrinkage is likely to cause incomplete cell recovery. Subsequently, as an alternative and direct way of examining sorting efficiency, we labeled the high secretor JCL260 Δ*lysA* pSA69 with mCherry by transformation with ampicillin-resistant pBT-proD-mCherry. To enable cultivation in ampicillin, lower secretor JCL16 Δ*lysA* was transformed with empty pTGD plasmid and K12 Δ*ilvD* Δ*galK::cfp-bla* pSAS31 (expressing both CFP and mNeongreen) was employed as the sensor strain. Thus, droplets could be sorted for green fluorescence and accuracy could be assessed based on whether droplets contained mCherry-expressing cells.

We encapsulated a mixture of JCL260 Δ*lysA* pSA69 and JCL16 Δ*lysA* at a 1:1,000 ratio on 1.0 g/L norvaline. In 30 min of sorting at a droplet reinjection rate of 2 μL/min (∼400,000 droplets total, or ∼40,000 containing a secretor strain), we retrieved five droplets in the collection device. All five contained substantial numbers of green fluorescent cells,

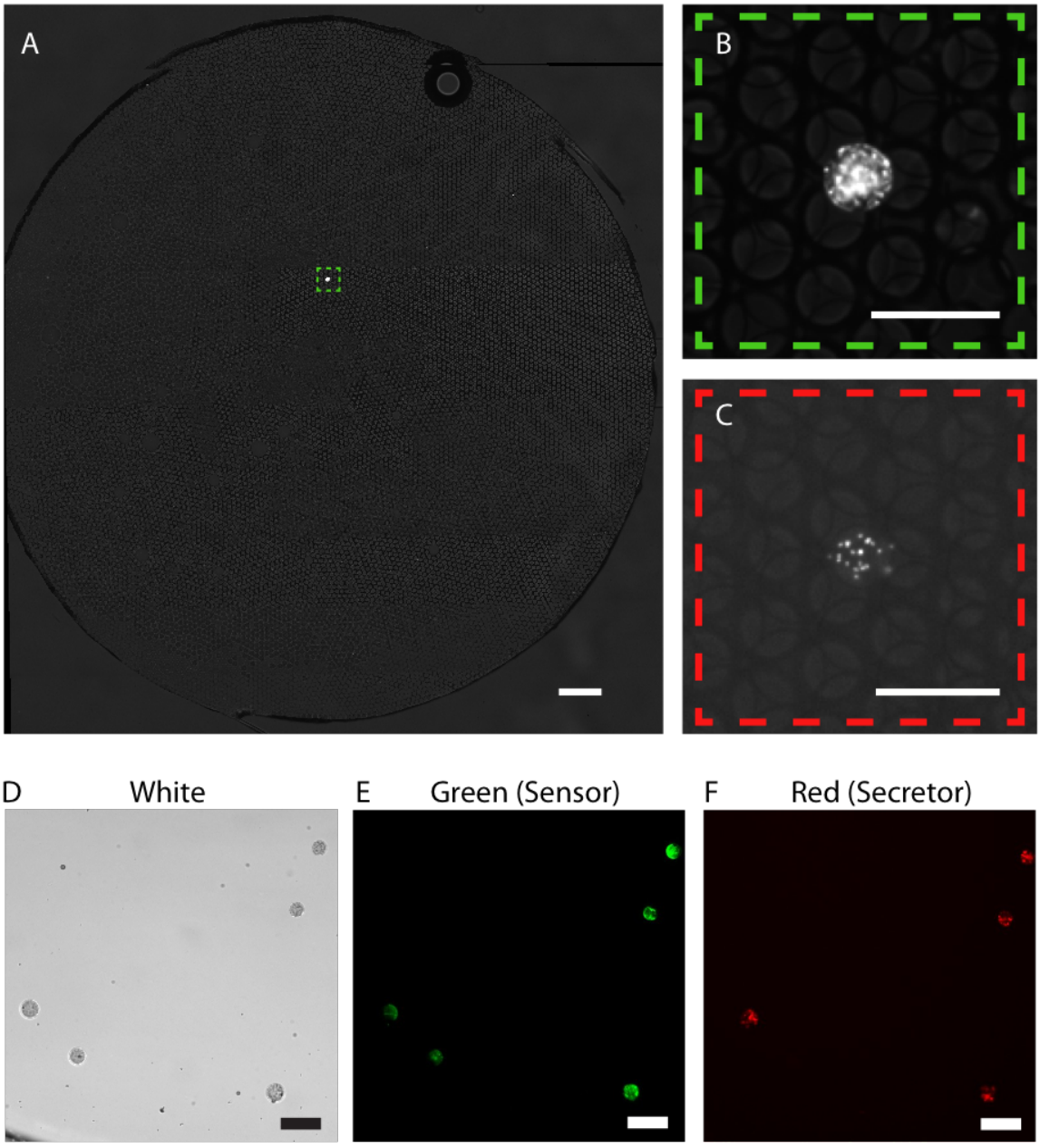

Figure 3.14: Droplets before and after fluorescence-activated droplet sorting (FADS). (A) ~90 pL droplets were generated containing sensor strain at a cell loading of ~5 sensor cells/droplet and secretor strains JCL16 Δ*lysA* pTGD and JCL260 Δ*lysA* pSA69/pBT1-proD-mCherry, mixed at a ratio of 1,000:1, with a total loading of ~0.1 secretor cell/droplet. After 36 h incubation, a sample of these droplets was examined. One droplet with substantial sensor cell growth is visible from among ~$5*10^4$ droplets (A, B). This droplet was identified as containing JCL260 Δ*lysA* pSA69/pBT1-proD-mCherry secretor strain by observation of cells expressing mCherry (C). After 30 min FADS of these droplets, five droplets were collected (D-F). All five showed growth of the sensor strain as determined by cells with green fluorescence (E) and four contained cells with red fluorescence, indicative of JCL260 Δ*lysA* pSA69/pBT1-proD-mCherry (F). Scale bars: (A) 500 μm, (B-F) 100 μm.

while four contained red fluorescent cells, indicative of the high secretor (true positive) (Fig. 3.12). This demonstrates the accuracy of the sorting system at high throughput (i.e., it is possible to sort a large number of droplets and isolate only those containing a high number of green cells) and the low biological false positive rate.

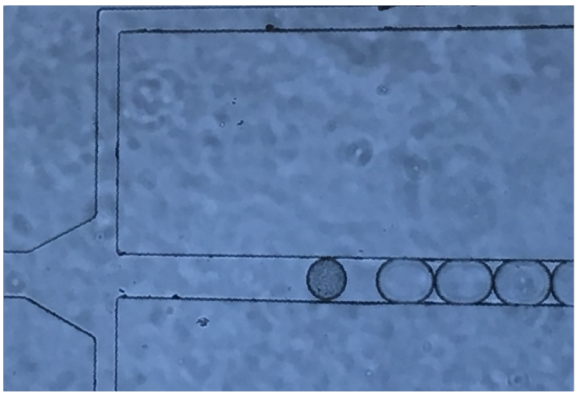

Figure 3.15: Droplets in droplet spacing device. The leftmost droplet displays significant growth.

We were also interested in whether we can isolate and retrieve cells from individual droplets. Each droplet will contain a single genotype of secretor, so it will be most efficient to isolate individual droplet after sorting and evaluate the cells from each one, as opposed to pooling the droplets and picking colonies that may have come from the same droplet. For this purpose, we utilized the droplet spacing device of [100]. This device incorporates a valve that, when partially closed, prevents droplets from flowing out of the device. It can be used to dispense individual droplets into wells of a microtiter plate. We encapsulated a mixture of JCL260 $\Delta lysA$ *alsS* $\Delta galK$ and JCL260 $\Delta lysA$ pSA69 in 125 μm diameter droplets and incubated to allow growth. We performed this experiment prior to optimizing the growth conditions, so not all of the encapsulated cells grew. Nevertheless, we identified some droplets with significant growth, such as the leftmost droplet in Fig. 3.15. Note that the droplets with significant cell growth are smaller than droplets without growth. This is likely due to osmosis occurring because of sugar consumption [101]. We used the spacing device to isolate six of the droplets that displayed good cell growth, depositing them into wells of a 96-well microplate. When these droplets were chemically destabilized and plated

on LB plates with tetracycline (to prevent growth of the sensor strain), four droplets yielded colonies and two droplets produced no colonies. The four that did produce colonies had 2, 13, 18, and 36 colonies, respectively. All colonies were purple, indicating they were JCL260 $\Delta lysA$ pSA69. Several colonies from each plate were also assayed by PCR to verify that they were indeed the high secretor strain. This technique can potentially be employed on the pools of droplets collected by FADS to isolate single droplets and reduce the amount of rescreening required. However, the spacing device is challenging to operate and it is easy to lose droplets, so it would be best to optimize the collection efficiency further or use when a large pool of droplet has been collected so that loss of droplets is not a problem.

3.3 Discussion and conclusion

We have demonstrated the implementation of SnoCAP as both a colony screening assay and utilized it to identify an efficient isobutanol production strain from a chemical mutagenesis library. We have also implemented the screening in microdroplets and demonstrated its ability to identify rare high-producers by testing model libraries. Each implementation formats has its own advantages and limitations. The microtiter plate assay of Chapter 1, although requiring the most space, avoids the issue of single cell variability and therefore provides the highest accuracy, with the tightest agreement of replicates, both between and within experiments. The agar plate format reaches high throughputs (up to $\sim 10^5$ assays per square meter of plate surface) and does not require specialized equipment. The high cell density possible in a colony enables a large number of doublings and hence a large degree of amplification of differences in production phenotype. The microdroplet format achieves ultrahigh throughputs ($\sim 10^6$ droplets/h, corresponding to $\sim 10^5$ assays/h, or $\sim 10^6$ assays/day, when a secretor loading of ~ 0.1/droplet is used) that can drastically reduce screening time and enable screening of libraries that are orders of magnitude larger. The small culture volume also significantly shortens the requisite incubation periods.

The droplet format can provide substantial time and cost savings. Agresti *et al.* [102]

estimate that to perform $5*10^7$ assays in microplates with liquid-handling robots would require two years and $10 million in pipette tip costs alone. At the sorting rates reported here, it would take about three weeks to evaluate $5*10^7$ assays and further optimization of the assay can lead to reductions in this time. The cost is also quite low, with the major cost being the optical setup (∼$10,000). Amortized over five years, this amounts to an equipment cost of $100 for a three week screening endeavor.

As with most high-throughput single cell tools, single cell variability decreases the precision of the agar plate and microdroplet formats. In the microdroplet assay, there is additional variability between replicates because the cells are encapsulated according to the Poisson distribution, and each droplet does not start with an identical number of sensor cells. One-to-one pairing of the cells (using techniques such as those presented in [103] or [104]) could be explored to achieve a more uniform sensor cell number per droplet. A more sophisticated incubation setup for improved aeration (such as that presented in [105]) may also improve the homogeneity of the conditions experienced by each droplet. The microdroplet sorting assay could also be further developed to enable even higher throughputs and to screen based on different sorting criteria. We have screened by sensor strain content, but secretor-to-sensor ratio may also be a useful quantity to examine. Advances in absorbance-activated droplet sorting (such as that reported in [106]) may also enable application to screening strains that are not easily made fluorescent.

An interesting future application of these high-throughput forms of SnoCAP would be to take an inverse metabolic engineering (IME) approach, reintroduce the B1 mutations by multiplex automated genome engineering (MAGE) or by cloning genomic fragments into a plasmid, and rescreen this library to elucidate the mechanism of production improvement. For certain applications, the microdroplet assay could also be combined with next-generation sequencing in order to gather large amounts of data about which members of a library perform well or poorly. For example, a ribosome binding site or promoter library could be generated and the results compared between cells from droplets that are positively sorted

and those that are negatively sorted.

3.4 Materials and methods

3.4.1 Strains and plasmids

Strains and plasmids used that were not already listed in Tables 2.1 and 2.2 are listed in Tables 3.2 and 3.3. pET-ara-mCherry was constructed by Jihyang Park [46] and pSAS31 by Scott Scholz [107]. The Δ*intC::yfp-cat* construct was initially obtained from the DS1-Y strain, from the Balaban group, Hebrew University of Israel. pBT1-proD-mCherry was a gift from Michael Lynch (Addgene plasmid #65823).

3.4.2 Gene deletions, insertions and modifications

Oligonucleotides are listed in Table 3.4.

To introduce the *aceK* gene mutation into JCL260 Δ*lysA alsS*, we first knocked out the *mutS* gene using single-stranded -Red recombineering using pSIM5 [84] and oligo mutS_STOP_JM to introduce premature stop codons. We then used single-stranded λ-Red recombineering with oligo aceK_510_mut_oligo to introduce the SNP into *aceK*. Both *mutS* and *aceK* were verified by allele-specific PCR and Sanger sequencing by the University of Michigan Sequencing Core (Ann Arbor, MI).

To knock out the *galK* gene, we amplified the *cat* gene from NV3r1 cm 20 gDNA using primers galK_cat_for and galK_cat_rev to add homology to the galK locus. The resulting linear construct was integrated by λ-Red recombineering with pSIM6 [84] into JCL260 Δ*lysA alsS*.

For fluorescent labeling of cells, K12 Δ*ilvD*::FRT was transformed with either pET-ara-mCherry or pSAS31. Secretor strains JCL260 Δ*lysA*::FRT *alsS* and JCL260 Δ*lysA*::FRT pSA69 were P1 transduced with Δ*intC::yfp-cat* P1 lysate in order to integrate the *yfp* gene into the *intC* locus.

3.4.3 Cell preparation for cross-feeding screening

For the agar plate assay, cells were prepared in the same manner as for the microtiter plate assay. For the droplet assay, the stationary phase cultures were subcultured in 10 mL culture volumes into exponential phase (∼3.5 h for JCL260-based strains and ∼2 h for K12-based strains). The cells were harvested by centrifugation at 5,000 g for 5 min, washed twice in 1x M9 salts, and resuspended in the culturing medium (M9IPG with norvaline at the specified concentration). Cell density was then determined based on an OD_{600} to CFU calibration. Cell inocula were kept at room temperature and each culture was combined just before droplet generation. Each inoculum culture was serially diluted and spot plated on LB plates with kanamycin to verify CFU concentration. After completing the experiments presented in this chapter, we determined that using stationary phase cells for the droplet assay is also effective and may provide better consistency in growth performance.

3.4.4 Agar plate assay

Setup Plates were either 10 cm round Petri dishes, or 24.5 cm square bioassay dishes. M9IPG with 5 g/L glucose, 3 mM isoleucine, 12 g/L agar, 0.1 mM IPTG and 50 μg/mL kanamycin was used for all plates in the 2-KIV screen. M9IPG with 4 g/L glucose, 12 g/L agar and no antibiotics was used for all microplates in the tryptophan screen. The prepared sensor cells were diluted 10^{-1}, secretor cells were serially diluted to the desired concentrations. On round plates, 100 μL of 10^{-1} diluted sensor cells and 200 μL 10^{-6} diluted secretor cells were spread with glass beads. On square plates, 1 mL of 10^{-1} diluted sensor cells and a combination of secretor cells totaling $\sim 10^4$ CFU were spread with glass beads. Plates were allowed to dry thoroughly before incubation to ensure good separation of colonies. 10^{-6} dilutions of secretor strains were also plated on LB plates to determine LB-CFU counts. The LB-CFU counts were used to determine model library percentages.

Plate scanning and analysis Automated scanning and ScanLag analysis was performed according to the method of [93, 94], using Epson Perfection V37 photo scanners. Custom holders were 3D-printed so that plates would be located in consistent locations between experiments. Plates were covered with sterile black felt prior to incubation and scanning. Scanners and plates were incubated at 35 °C to accommodate recommended scanner operating temperature. Images were taken every 30 min. for the duration of the incubation period. Following colony growth, images were aligned and colonies detected. Colonies that had merged by the end of the culture period were eliminated from the growth profile plots.

Colony composition determination Entire colonies were scooped off plate using inoculation loops and resuspended in M9 salts. The colony suspension was then serially diluted and plated in the same manner described for the microplate population composition determination.

Determination of strain identity Secretor strains were isolated from mixed colonies by streaking on LB plates with tetracycline. The resulting colonies were then assayed by colony PCR with primers alsS_int_chk_front_for and alsS_int_chk_front_rev, which produce a short band if *alsS* is integrated, and with alsS_int_chk_front_for and alsS_int_chk_back_rev, which produce a short band if *alsS* is not integrated.

NTG mutagenesis and mutant screening NTG mutagenesis was performed according to the method of 10. The parental strain (JCL260 Δ*lysA* *alsS*) was cultured overnight in LB medium and then diluted 1% (v/v) into 5 mL of fresh LB and grown at 37 °C to an OD_{600} of 0.5. Cells were pelleted by centrifugation at room temperature (5,000 rpm, 10 min) and washed twice with an equal volume 0.1 M Na citrate (pH 5.5) before resuspension in 0.1 M Na citrate (half the original volume). NTG was added to a final concentration of 50 μg/mL from a 1 mg/mL stock in 0.1 M Na citrate and incubated at 37 °C for 15 min. For the control experiment, an equal volume of 0.1 M Na citrate was added instead of NTG.

After incubation, cells were washed twice with the original volume of 0.1 M phosphate buffer (pH 7.1). The cells were then resuspended in 5 mL of LB and incubated at 37 °C overnight for outgrowth. Before outgrowth, a small volume of cells were diluted and plated for both the NTG-treated tube and the control tube on LB plates to determine the kill count for the experiment before outgrowth. After outgrowth, cells were prepared for the agar plate screening as described in Cell preparation for cross-feeding screening, and plated with sensor strain on 24.5 cm square plates containing 0.5, 0.75 or 1.0 g/L norvaline, and incubated at 37 °C for 7 days. Cells were also plated on LB plates to determine LB-CFU. A volume of secretor cells corresponding to $\sim 6*10^4$ LB-CFU was plated on each square plate. The largest colonies from the cross-feeding screening plates were streaked on LB plates with tetracycline to isolate the secretor. Colonies from these plates were then re-screened in the microplate format of the co-culture screening. The most promising strains from the re-screening were transformed with pSA65 and subsequently tested for isobutanol production performance.

Genomic sequencing The genomes of parental strain JCL260 Δ*lysA alsS* and mutant B1 strain were sequenced by the University of Michigan Sequencing Core (Ann Arbor, MI) by Illumina HiSeq-4000 using 0.5% of a lane for each strain. Genomic DNA of the strains was isolated using a Qiagen DNEasy Blood and Tissue kit. The libraries were prepared by the core, with 250 nt insert size and reads were 150 nt paired-end reads. SolexaQA++ (http://solexaqa.sourceforge.net/) was used to trim the reads. Reads were mapped to the reference genome (JCL260 Δ*lysA alsS*) using Bowtie 2 (http://bowtie-bio.sourceforge.net/bowtie2/), and SNPs were compared using SAMtools (http://samtools.sourceforge.net/).

3.4.5 Microdroplet assay

Microfluidic device fabrication Polydimethylsiloxane (PDMS) droplet generation and detection/sorting devices were fabricated using standard soft lithography methods. Briefly,

an SU-8 photoresist (MicroChem Corp.) master mold was first created on a Si wafer by photolithography. For the detection/sorting device, multiple coating and exposure steps were required to construct flow channels and optical fiber grooves with different heights (50 μm and 80 μm, respectively). The wafer was silanized with vapor phase trichloro(1H,1H,2H,2H-perfluorooctyl)silane (Sigma-Aldrich) and the PDMS precursor was poured onto the master mold and cured at 65 C overnight. The cured PDMS was peeled off the Si wafer, punched to form inlets/outlets, treated with oxygen plasma (Femto Scientific Inc.) for activation, and finally bonded to a glass slide to seal the device. For the droplet generation device, a 1.2 mm biopsy punch was used for both inlets and the oulet. For the detection/sorting device, a 1.2 mm punch was used for the droplet inlet, a 1.5 mm punch for the electrodes, and a 0.75 mm punch for the oil inlets and droplet outlets. The microelectrodes were created by flowing low melting Bi/In/Pb/Sn alloy (247 Solder) into the microchannels at 150 ^{o}C. The optical fiber (F-MCB-T-1FC, Newport Corp.) was manually embedded into the fiber groove. The sealed flow channel was flushed with trichloro(1H,1H,2H,2H-perfluorooctyl)silane at a concentration of 2% (v/v) in Novec HFE-7500 prior to use. For the collection device, a $\sim$1 mm PDMS membrane containing a channel of height 100 μm and width 1 mm was punched with an 8 mm diameter hole and bonded to another PDMS slab. An inlet and outlet were punched with a 0.75 mm punch and the PDMS was bonded to a glass slide.

Encapsulation and cultivation Cells were prepared as described above. Individual strain inocula were kept at room temperature (storage on ice decreases viability for co-culture growth) and were combined with each other immediately before each set of encapsulation. The strains were combined with additional medium to achieve a total secretor cell loading of $\lambda_{\text{Secretor}} = 0.1$ cell per droplet and sensor cell loading of $\lambda_{\text{Sensor}} = 5$ cells per droplet for 55 μm diameter droplets, or $\mu_{\text{Sensor}} = 15$ cells per droplet for $\sim$125 μm diameter droplets. M9IPG medium with 20 g/L glucose was supplemented with 50 μg/mL kanamycin, 0.1 mM IPTG, and norvaline at the specified concentration. For this assay, it is important that the

medium be freshly made on the day of the experiment, and the LB medium used for growing the inoculum cultures should also be fresh. The inoculated cell culture and Novec HFE-7500 fluorinated oil (3M) with 2% (w/w) PEG-PFPE amphiphilic block copolymer surfactant (Ran Biotechnologies, 008-FluoroSurfactant) were loaded into separate syringes (BD, 1 mL and 5 mL, respectively) and injected with 23 gauge needles through PTFE tubing (0.022" ID, Cole-Parmer) into a flow-focusing droplet generation device using syringe pumps (Kent Scientific) with flow rates of 10 μL/min and 45 μL/min, respectively. With a device with channel height of 50 μm and aqueous and oil channel widths of 25 μm, these flow rates produced droplets with diameter of $\sim$55 μm. The emulsion was collected in 1.5 mL Eppendorf tubes in 10 min aliquots ($\sim$150 μL of emulsion). Excess oil was removed, except for $\sim$100 μL, and the capped Eppendorf tubes were then incubated at 37 ^{o}C. Overall population-level growth characteristics were monitored by incubating 100 μL of droplets with 50 μL additional fluorinated oil with surfactant in a black clear-bottom microplate (Greiner), sealed with a Mylar plate sealer (Thermo Scientific), in a microplate reader (BioTek Synergy H1) at 37 ^{o}C, reading fluorescence every 15 min (excitation 485 nm, emission 528 nm). Droplets were incubated for between 30 and 40 h. An Olympus DP71 microscope was used to examine the cell growth.

Droplet sorting Following off-chip incubation, droplets were poured into a capped syringe (Global, 1 mL) and any remaining volume of the syringe was filled with fluorinated oil with 2% surfactant before inserting the syringe plunger. We reinjected the droplets into a droplet sorting device with height 50 μm and main channel width 55 μm using a syringe pump (KD Scientific). The syringe was connected to the sorting device via a 23-gauge needle and 0.022" ID PTFE tubing. Droplets of 55 μm diameter were reinjected into the sorting device at 1.5-2.5 μL/min, corresponding to $\sim$150-300 droplets/sec. Spacing oil with surfactant was also injected into the device by syringe pumps. Sorting was performed in a manner similar to that of [104], with the major difference being that droplets were generated using a droplet

generation device, incubated off-chip to allow cell growth, and then reinjected into a sorting device. The flow channel structure of the detection/sorting device was designed such that the negative signal droplets spontaneously enter the waste channel due to lower flow resistance, which results from a larger width and shorter length of the channel. The positive channel was designed to have a high enough resistance that smaller broken droplets (which may occur during reinjection) would also flow into the negative channel.

Cell retrieval and identity determination Collected droplets were examined by microscopy in the collection device. The secretor identity was either determined directly by fluorescence, or by retrieving the droplets and plating on MacConkey agar plates to identify cell type by colony color. For determination by fluorescence, the droplets were inspected by a Nikon Eclipse Ti-S microscope to detect sensor strain green fluorescence (excitation filter 470/40, emission filter 525/50) and red fluorescence from mCherry (excitation filter 560/55, emission filter 675/67). A droplet containing cells displaying red fluorescence under these filters indicated the presence of the secretor strain JCL260 Δ*lysA* pSA69/pBT1-proD-mCherry and was therefore considered a positive droplet. For determination by colony count, the pool of collected droplets was retrieved by inverting the collection device and injecting HFE-7500 oil into the device to flow the droplets into an Eppendorf tube. The droplets were then chemically destabilized with 1H,1H,2H,2H-perfluoro-1-octanol. LB medium was added and the aqueous portion was plated on MacConkey agar plates containing tetracycline and 1% galactose. Plates were incubated for 48 h at 37 °C. Sensor cells did not grow due to the tetracycline. JCL260 Δ*lysA* *alsS* Δ*galK* produced white colonies and JCL260 Δ*lysA* pSA69 produced purple colonies. Purple colonies (indicating JCL260 Δ*lysA* pSA69) were restreaked on MacConkey agar plates with tetracycline and kanamycin (growth verified maintenance of the pSA69 plasmid) and on MacConkey agar plates with tetracycline and chloramphenicol (lack of growth confirmed that the colony is not mixed with JCL260 Δ*lysA* *alsS* ΔgalK). Cells from randomly selected purple colonies were transformed with pSA65 and tested for

isobutanol production performance.

Table 3.2: Additional strains employed in the studies described in Chapter 3

Strain	**Relevant genotype**	**Reference**
JCL260 Δ*lysA alsS* Δ*galK*	JCL260 Δ*lysA alsS* Δ*galK::cat*	This study
JCL260 Δ*lysA yfp*	JCL260 Δ*lysA*::FRT Δ*intC::yfp-cat*	This study
JCL260 Δ*lysA alsS aceK-mut*	JCL260 Δ*lysA alsS mutS*- with AceK mutation P510S	This study
K12 Δ*ilvD* pSAS31	K12 Δ*ilvD*::FRT pSAS31	This study
K12 Δ*ilvD cfp* pSAS31	K12 Δ*ilvD*::FRT Δ*galK::cfp-bla* pSAS31	This study

Table 3.3: Additional plasmids used in the studies described in Chapter 3

Plasmid	**Relevant genotype**	**Reference**
pSIM5	Red expression plasmid; chloramphenicol-resistant	[84]
pET-ara-mCherry	Arabinose-inducible mCherry expression plasmid; kanamycin-resistant	[108]
pSAS31	Constitutively expressed mNeonGreen plasmid; kanamycin-resistant	[107]
pBT1-proD-mCherry	Constitutively expressed mCherry plasmid; ampicillin-resistant	Addgene plasmid #65823

Table 3.4: Primers and oligos employed in the studies described in Chapter 3

Name	**Sequence 5' → 3'**
mutS_STOP_JM	G*C*G*G*AACTGCTGTATGCAGAAGATTTT GCTGAAATGTCGTGATGATAAGGCCGTCG CGGCCTGCGCCGTCGCCCGCTGTGGGAG TTTGAA
mutS_mut_as_for	CAGAAGATTTTGCTGAAATGTCGTG
mutS_mut_as_rev	GGGTGATTTCCAGATTACGACG
mutS_seq_for	GATATCAGTTCCGGGCGTTT
mutS_seq_rev	GTTCTCGACGCCAAAACC
galK_cat_for	GTTTGCGCGCAGTCAGCGATATCCATTTTC GCGAATCCGGAGTGTAAGAACGTTGATCGG CACGTAAG
galK_cat_rev	CGGAAGAGCTGGTGCCTGCCGTACAGCAAG CTGTCGCTGAACAATATGAATTACGCCCCGCC CTGCCA
aceK_510_mut_oligo	G*G*C*G*CATAGCCAGTGGCGAAACTCTTCCG GGAAAACATCGCCCGACGAGACGCTGTACCA CGGTTCGCTGGCAAGTTCGTCTTCCGGATA
aceK_mut_as_for	CCGTGGTACAGCGTCTCGT
aceK_mut_as_rev	TCTGCCTTTGAGTTGGCTTT
aceK_seq_for	CGCGTCTTATCATGCCTACA
aceK_seq_ rev	TCTGCCTTTGAGTTGGCTTT
	* indicates phosphothoriated bond

Chapter 4: Chromosomal integration and optimization of isobutanol pathway in *E. coli*

4.1 Introduction

Chromosomal integration of recombinant genes is desirable compared to expression from plasmids due to increased stability and elimination of the need for antibiotics for ensuring plasmid maintenance. Antibiotics are costly, can lead to the spread of antibiotic resistance, and also place additional stress on the cells. Chromosomal integration requires optimization of expression level. Options for achieving optimal expression levels include variation of the integrated gene's copy number, genome position, promoter/ribosome binding site, or post-transcriptional elements.

One strategy for stable integration at adjustable copy numbers is Chemically Inducible Chromosomal Evolution (CIChE), developed by Tyo *et al.* [59]. In this method, the genes to be integrated, as well as a chloramphenicol resistance gene (*cat*) are flanked by matching 1 kb homology regions that are not homologous to the *E. coli* chromosome (in [59] they were taken from *Synechocystis* PCC6803). Once this construct is integrated into the genome, the strain is passaged into higher concentrations of chloramphenicol. Strains can use RecA-mediated recombination to produce multiple copies of the integrated construct, giving them a growth advantage in the high chloramphenicol concentrations. The *recA* gene is then be deleted, stabilizing the strain. The method is summarized in Fig. 4.1. CIChE was first demonstrated for producing the biopolymer poly-3-hydroxybutyrate, and it increased the stability ten-fold compared to a strain with two plasmids carrying the pathway. It has also

been utilized with triclosan as the selection agent [109].

Another possibility for modulating the expression level of an integrated gene is by the choice of its integration location. Different chromosomal positions are known to have different expression levels [107, 110, 111], but it is not well understood how culture condition affects these levels. It is also not easy to predict how an integrated construct will affect the expression at that site, and it can vary depending on what the construct is [62]. The integrated construct itself can also have effects on the surrounding genes, which needs to be considered. Thus, expression optimization by chromosomal position-dependent expression variation is an underexplored strategy.

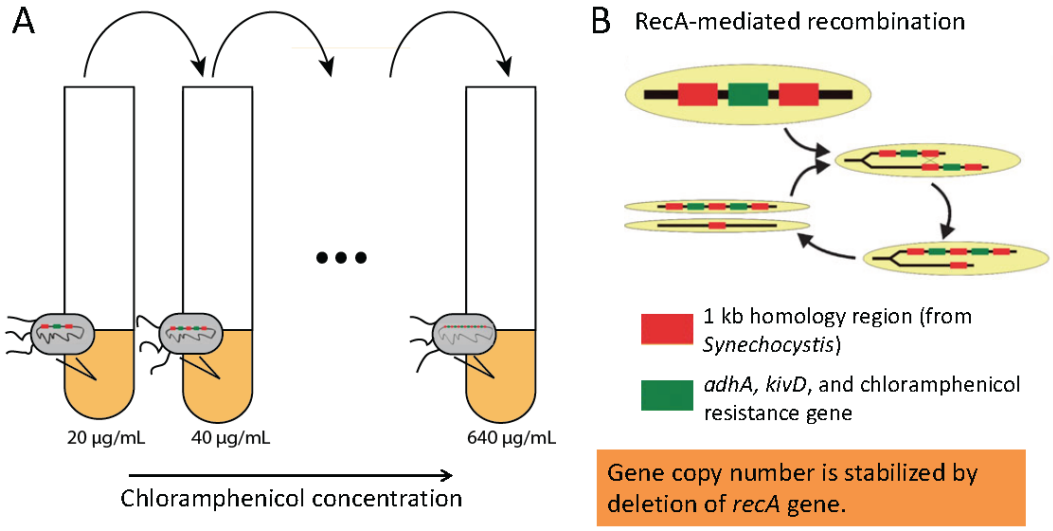

Figure 4.1: Overview of the CIChE method, applied for *kivd* and *adhA*. Passaging of cells into successively higher concentrations of chloramphenicol (A) selects for cells that have performed RecA-mediated recombination to increase the copy number of the genes of interest, which are flanked by 1 kb homology regions (B). This figure is adapted from [59].

4.2 Results

4.2.1 Investigation of plasmid burden

To investigate the effect of plasmid burden on the isobutanol production strains, we grew the NV3r1 strain with and without the isobutanol pathway genes in several different media with and without inducer (IPTG). We observed that when the strains were under

richer medium conditions, with yeast extract supplement, the growth rate was not affected by plasmids or antibiotics. Under minimal medium conditions, however, the strain with plasmids and antibiotics had severe growth defects. Therefore, integration of the pathway is expected to provide a benefit in terms of growth rate, leading to an increased production rate as well.

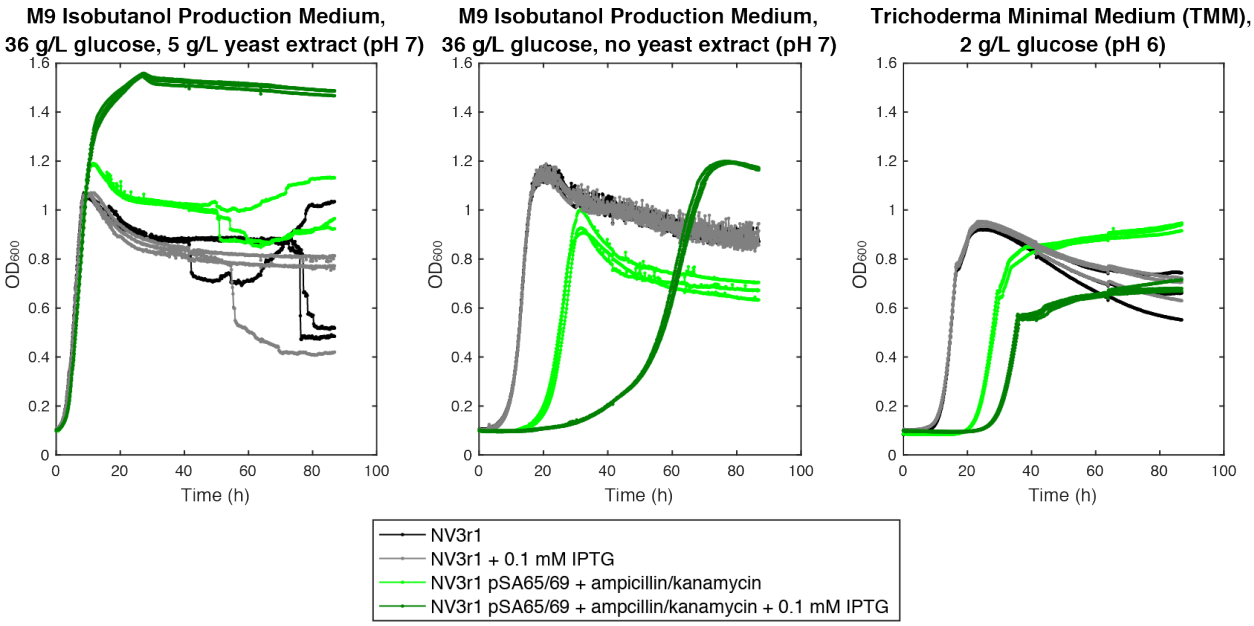

Figure 4.2: Growth profiles of NV3r1 with and without pSA65/9 plasmids, with and without IPTG induction, in several different media.

4.2.2 Integration and copy number optimization of *kivd/adhA* genes

To integrate and optimize copy number of the *kivd* and *adhA* genes we used CIChE. We integrated the $P_L$lacO1-*kivd-adhA* operon into the *aslB* locus of NV3r1 by λ-Red recombination, as described in the Materials and methods section of Chapter 2. The *aslB* site was chosen due to work showing that it is a highly expressed location [111]. We then passaged the resulting integrant in LB medium with successively doubled concentrations of chloramphenicol and then knocked out the *recA* gene by P1 transduction. The copy number of resulting strains was assessed by qPCR (Fig. 4.3A). The strains were also transformed with the pSA69 plasmid and tested for isobutanol production in monocultures in M9

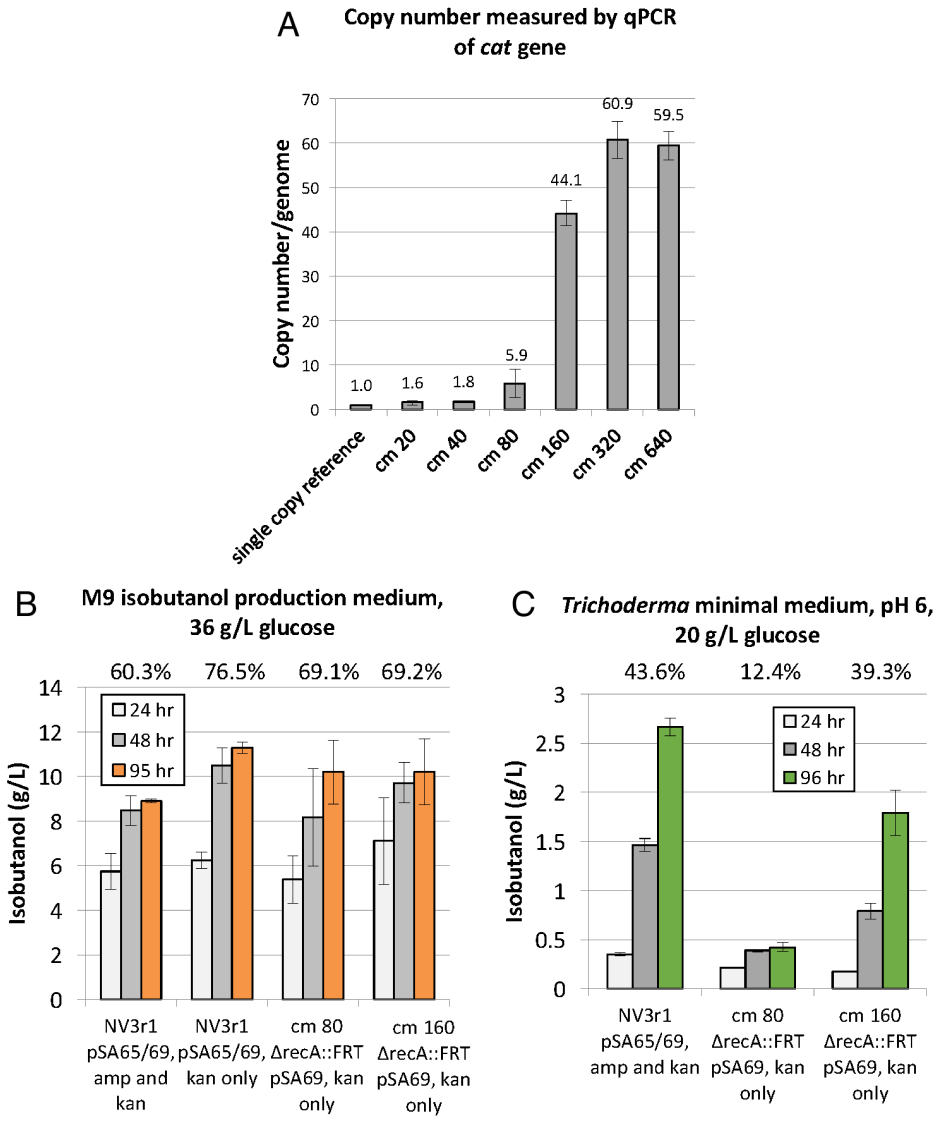

Figure 4.3: Characterization of integrated strains. (A) Construct copy number of strains after CIChE and *recA* deletion. Error bars represent the propagated standard deviation of three technical replicates of both *cat* gene PCR and *bioA* reference gene PCR. Isobutanol production after transformation with pSA69 in (A) M9IPG medium with 5 g/L yeast extract (B) and TMM (C). Labels indicate strain and which antibiotics were added to the culture. Percentages above the graphs represent the final percentage of the theoretical yield that was produced based on how much glucose was consumed. Error bars represent the standard deviation of three biological replicates.

medium with yeast extract (Fig. 4.3B) and in *Trichoderma* Minimal Medium (TMM), pH 6 (Fig. 4.3C). We found that the copy number increases with increasing copy number until it plateaus around 60 copies. In the more optimal M9IPG medium, with yeast extract and neutral pH, the strain with six copies performed similarly to the strain with 44 copies. In minimal medium, however, the strain with 44 copies much better than the one with six copies. This further highlights the usefulness of the strategy since the strains can be tuned to find the optimal expression for a given process condition.

4.2.3 Construction and screening of *alsS* transposon integration library

To integrate the *alsS* gene, we decided to explore position-dependent expression variation by creating a library in which the gene is randomly integrated into various sites in the genome (Fig. 4.4). We performed this integration using the transposon Tn5 to integrate copies of the *alsS* gene as well as a kanamycin resistance gene to use for selection. We integrated into JCL260 Δ*lysA* and produced a library of $3.4 \pm 1.1 * 10^4$ members. When library members were randomly selected, transformed with pSA65 and tested for isobutanol production, only one (of 12) colony produced significant amounts of isobutanol (data not shown). We therefore employed the SnoCAP screening strategy described in Chapters 2 & 3, in the microdroplet format. We encapsulated the library as well as several controls secretor strains along with the sensor strain K12 Δ*ilvD* Δ*galK::cfp-amp* pSAS31, and incubated to allow co-growth. The fluorescence profile of the droplets showed that the library had a higher percentage of fluorescent droplets compared to the parent strain (Fig 4.5A). We then sorted the droplets for those with high fluorescence, pooled the collected droplets, and plated them on plates with tetracycline to isolate secretor strain colonies. When colonies were rescreened in the microplate format of SnoCAP, the growth profiles demonstrated that we had removed the low performers (such as those visible in the randomly selected library members plot) (Fig 4.5B).

We also tried integrating *kan*-$P_L$lacO1-*alsS-ilvCD* into the genome using the same

method as described above. This construct it larger (∼6.3 kb) and produced a smaller library size ($6.0 \pm 2.7 * 10^3$ members). When screened by SnoCAP in droplets at various concentrations of norvaline, no significant growth was observed. We hypothesize that this apparent lack of productivity may be due to problems with the integration of such a large construct by the transposon method (i.e., although the cells acquire the antibiotic resistance they may not acquire all of the rest of the construct).

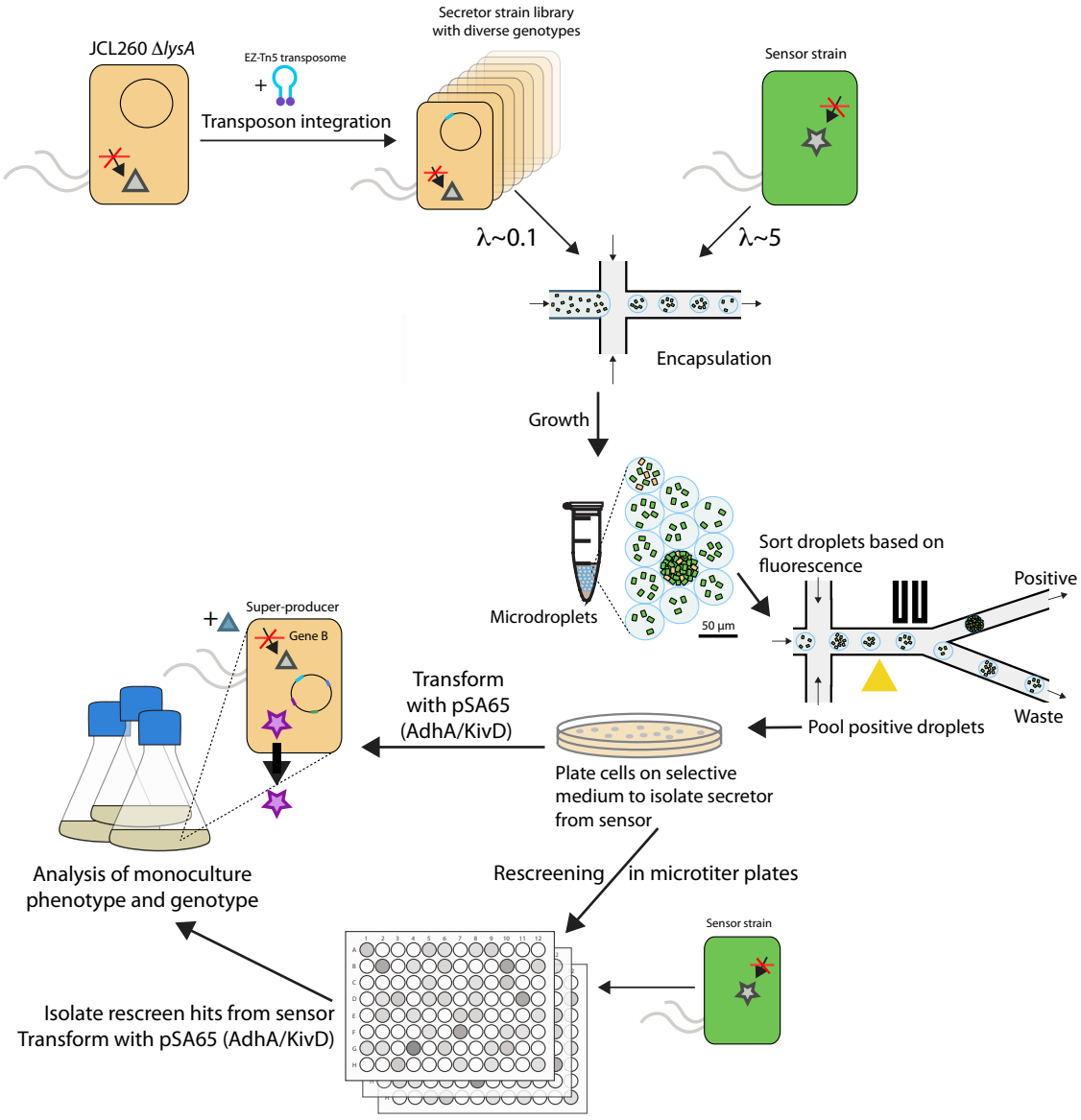

Figure 4.4: Overall schematic of transposon library construction and screening. A transposome containing *alsS* and the kanamycin-resistance gene is integrated into random genome positions in the parent strain JCL260 Δ*lysA*. The resulting library is screened by SnoCAP screening in the microdroplet format. Pools of sorted droplets are plated on LB plates with tetracycline to isolate the secretor strains.

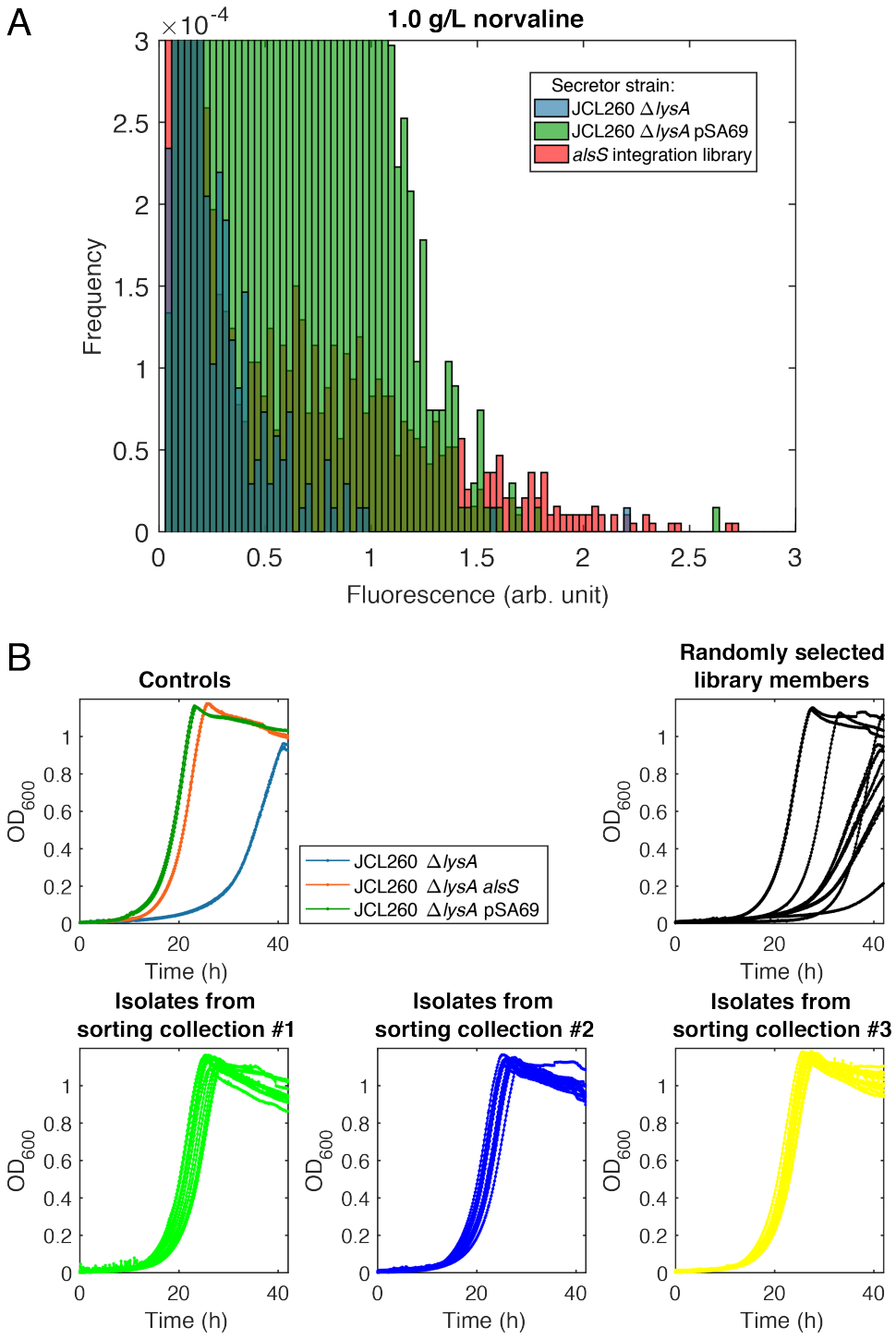

Figure 4.5: Screening of *alsS* integration library. (A) Histogram comparing fluorescence signal from droplets containing co-cultures with secretor strain JCL260 Δ*lysA*, JCL260 Δ*lysA* pSA69, and the *alsS* transposon integration library. K12 Δ*ilvD* Δ*galK::cfp-amp* pSAS31 was used as sensor strain, and the medium contained 1.0 g/L norvaline. ~70,000 of each of the control droplets (JCL260 Δ*lysA*, JCL260 Δ*lysA* pSA69) and ~190,000 of the library droplets were analyzed. (B) Co-growth profiles of library-isolates with and without screening, compared to controls, with sensor strain K12 Δ*ilvD*, without norvaline.

We transformed several of the fastest co-growing isolates (named d2, ds-1-4, ds-1-8, ds-1-9, ds-1-11, ds-2-7, ds-2-8, ds-3-1, ds-3-8) with pSA65 and tested isobutanol production. They displayed improved production compared to the parent strain (Fig. 4.6). Thus, the screening successfully enriched for the higher producers. It is interesting to note that, while in medium with yeast extract the isolates all have quite similar production levels, under minimal medium there is more variation. The production levels are also generally higher under the minimal medium conditions, likely due to higher expression of the *ilvCD* genes since under these conditions the strain must produce branched chain amino acids for growth.

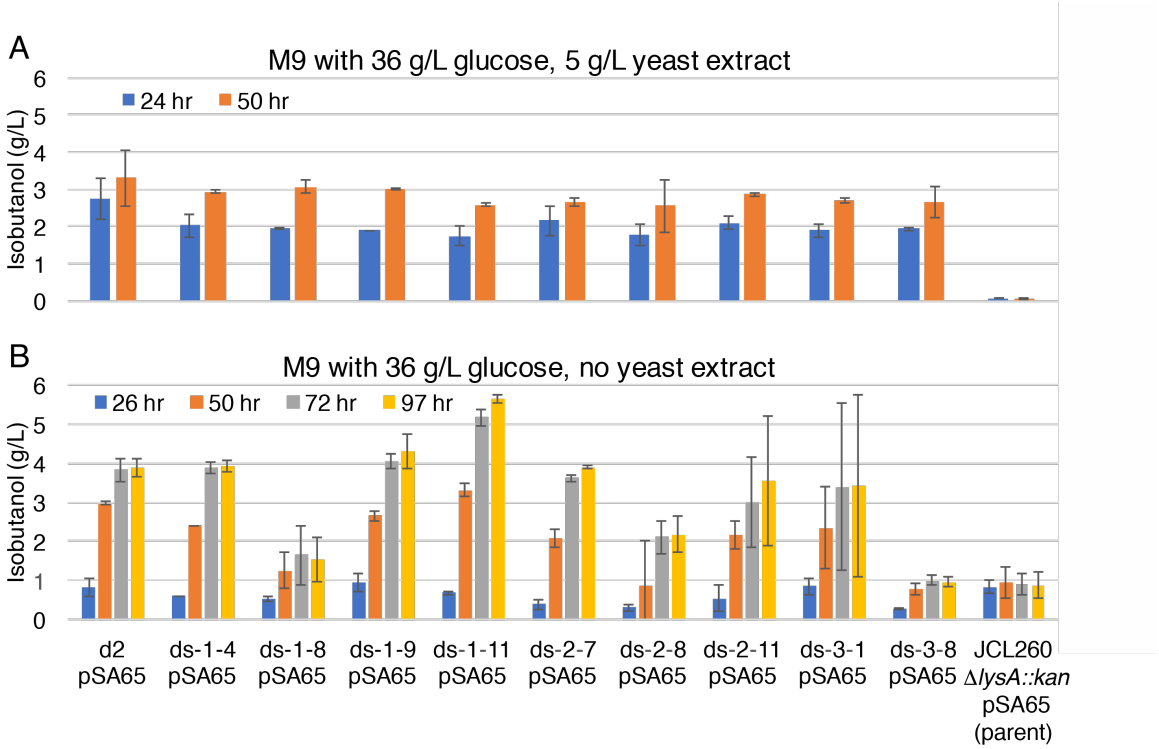

Figure 4.6: Isobutanol production of isolates from screening *alsS* integration library. Production testing was carried out after transformation with plasmid pSA65, in M9IPG medium with (A) and without (B) yeast extract.

4.3 Discussion and conclusion

This chapter has presented work on developing *E. coli* strains with the isobutanol pathway integrated into the genome. We utilized CIChE to integrate the *kivd* and *adhA*

genes at high copy numbers, generating strains with isobutanol production levels (with the pSA69 plasmid) that are comparable to the double plasmid strain. One consideration in using the CIChE method is that it requires the deletion of the *recA* gene to stabilize the strain. Under certain contexts this deletion may be deleterious to the cells since *recA* is important for DNA repair. If necessary, other methods of multiple copy integration, such as the flippase-based method of Gu *et al.* [60], can be used to circumvent this problem.

We also constructed and screened a library with *alsS* inserted into random locations in the genome. This *alsS* library has a relatively high percentage of strains that can grow in the droplets in 1.0 g/L norvaline (∼1 in 25 in one set of droplets and ∼1 in 100 in another set). While it is possible to find improved isobutanol-producing strains by randomly selecting colonies from this library without screening, the majority are low-producers. The screening successfully enriched for the higher producers. Future work can be done to multiplex integration of the *ilvCD* genes with the *alsS* integration. Because the *alsS* library has a relatively high percentage of good producers compared to the model libraries we tested, we expect the screening can be effective in screening this type of multiplexed library.

Future work should also be done to explore the mechanism by which the higher producers from the *alsS* integration library achieve their high-production phenotypes. Expression level analysis and genomic position determination by transposon footprinting are expected to provide information about the optimal expression level of this gene. It will also be interesting to see whether *alsS* expression level varies under minimal compared to yeast extract-supplemented conditions.

4.4 Materials and methods

4.4.1 Chemically Inducible Chromosomal Evolution (CIChE)

CIChE was performed in NV3r1 as described in Chapter 2. The CIChE construct was then amplified from the pTGD-*adhA-kivd* plasmid using aslB_integr_for and aslB_integr_rev

(see Table 2.3), which add 40 bp regions of homology to the *aslB* locus of the *E. coli* genome. The construct was integrated into the NV3r1 genome using λ-Red recombineering with pSIM6 [84] with selection on LB plates with 20 μg/mL chloramphenicol. A resulting integrant (named NV3r1 cm 20) was verified by PCR and Sanger sequencing. Subsequently, CIChE was performed by growing NV3r1 cm 20 to saturation and passaging into in successively higher concentrations of chloramphenicol (cells were passed 1% v/v, and antibiotic concentration was doubled in each passage). The *recA* gene was then deleted from cells at various final levels of chloramphenicol by P1 transduction from donor strain BW 26,547 Δ*recA::kan* Lambda *recA+*, which was obtained from the Coli Genetic Stock Center (CGSC). Transductants were selected on LB plates with 50 μg/mL kanamycin and the corresponding concentration of chloramphenicol.

4.4.2 qPCR for gene copy number determination

qPCR was performed with the same primers as [59], amplifying a portion of the *cat* gene and also of *bioA* as a single copy reference gene. Reactions were performed in 25 μL samples on an MJ Research (BioRad) Chromo4 thermocycler with SYBR Green qPCR Master Mix (Life Technologies). A strain with a single copy of *cat* integrated (NV3r1* *intC::yfp-cat*, from Chapter 4) was diluted and used to construct a standard curve.

4.4.3 Isobutanol production cultures

Isobutanol production cultures and metabolite measurements by HPLC was performed as described in Chapter 2.

4.4.4 Transposon integration of *kan-alsS*

kan-$P_L$lacO1-*alsS* was amplified by PCR from the pSA69 plasmid, adding the Tn5 mosaic ends. phosph_transp_kan_for (5Phos/CTGTCTCTTATACACATCTACCGAGCGTTC TGAACAAAT) and phosph_transp_alsS_rev (5Phos/CTGTCTCTTATACACATCTGGTG

ATTCCTCGTCGACCTA), where 5Phos/ represents a 5' phosphorylated primer. The product was then digested with both DpnI and SpeI enzymes (NEB) to digest the template plasmid, phosphorylated with T4 polynucleotide kinase (NEB) in case of lost phosphorylation during the previous steps, cleaned with a PCR clean-up kit (Qiagen) and eluted in TE buffer, and reacted with EZ-Tn5 transposase (Lucigen) according to the manufacturer's instructions. The resulting transposome was then electroporated into competent JCL260 Δ*lysA*::FRT. The cells were recovered for 1 hr with 1 mL SOC medium. Then 50 μL was used for dilution and plating on LB with kanamycin plates to assess library size, and the remaining ∼1050 μL of cells were grown to saturation in 100 mL LB with 50 μg/mL kanamycin to select for successful integrants. The cells were then frozen in 1 mL aliquots (resuspended in fresh LB with 25% glycerol) and later thawed, washed, and grown in LB with 50 μg/mL kanamycin to prepare them for screening.

4.4.5 SnoCAP screening

SnoCAP screening, in droplet and microplate formats, was employed as described in Chapters 3 and 2, respectively. Stationary phase cultures in LB were used as the inocula for both formats. The droplet collection device was soaked in a mixture of HFE-7500 oil and water for several days prior to use in order to improve droplet stability after collection.

Chapter 5: Strategies for improving fungal-bacterial co-culture for more efficient CBP

5.1 Introduction

Lignocellulosic biomass is an abundant and underutilized carbon source that does not compete with food supply. It is therefore a promising feedstock for sustainable production of liquid biofuels. The major challenge to cost-effective conversion of lignocellulosic biomass is that it must be broken down into soluble sugars before it can be fermented. Use of synthetic microbial consortia can help to address some of the limitations of monoculture bioprocessing, including high metabolic burden on the cells and the need to optimize multiple pathways or functions in the same species. Design of synthetic microbial consortia, however, brings the challenge of identifying or engineering strains that can perform optimally within a mixed population. Previous study has shown that a mixed culture of *Trichoderma reesei* and *Escherichia coli* can co-exist stably and directly convert cellulosic biomass to isobutanol, a promising biofuel candidate [53]. The production yield, titer, and rate require improvements for the process to be economically viable on an industrial scale. The increase in byproduct levels compared to monoculture fermentations and the instability of the plasmids encoding the isobutanol production pathway indicate that our *E. coli* strains are suboptimal for function in the co-culture and the hydrolysis rate is a major limiting factor in the process. In this chapter, we investigate approaches to improve co-culture performance. First, we knocked out competing pathways in an effort to decrease byproduct production and increase isobutanol, which did not prove successful. Second, we began work toward adaptive co-evolution of the

T. reesei/E. coli co-culture.

5.2 Results

5.2.1 Deletion of competing pathways

We observe a significant difference in the distribution of products generated by *E. coli* in co-culture with *T. reesei* compared to monoculture (Fig. 5.1). Additionally, we see plasmid loss from the *E. coli*, with only ∼50% of viable cells retaining both plasmids after 8 days [53]. These behaviors indicate that there is room for optimization of the *E. coli* strains to perform better under co-culture conditions. We also observe significant differences in the product distributions on the two types of substrate examined (microcrystalline cellulose (MCC) and AFEX-pretreated corn stover), indicating that the feedstock conditions greatly affect *E. coli* metabolism.

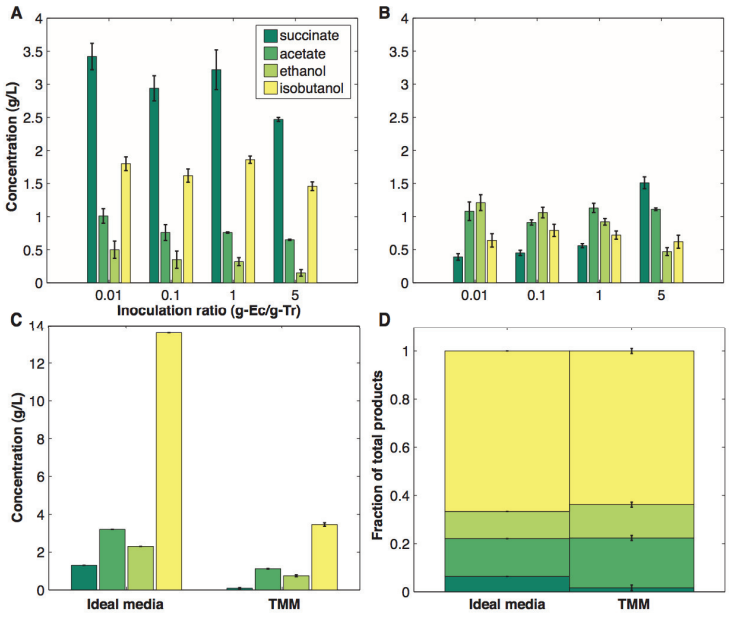

Figure 5.1: Major fermentation products of *T. reesei* RUT-C30/*E. coli* NV3 pSA55/69 co-cultures on AFEX pretreated corn stover (A) or microcrystalline cellulose (B), and of monocultures of *E. coli* NV3 pSA55/69 on glucose (C, D). Adapted from [53].

We first investigated targeted gene deletion in NV3 with the goal of reducing losses to side products and increasing isobutanol production. Knockouts of the *adhE*, *pta*, and

frdB were implemented with the goal of reduction in ethanol, acetate, and succinate production, respectively, These gene deletions were successful in reducing byproduct formation and increasing isobutanol production in monoculture (Fig. 5.2A). On MCC, they led to a reduction in ethanol and acetate and no significant change in succinate or isobutanol (data not shown). The knockout strain did not show any significant improvement compared to the base strain when co-cultured on AFEX-pretreated corn stover (Fig. 5.2B). Some of the lack of improvement in co-culture performance may be to reduced growth performance in the minimal, pH 6 medium used to grow the co-culture (Fig. 5.3). We also observed significant variability in our co-culture experiments, both between biological replicates in the same experiment and between experiments. This led us to pursue a directed evolution approach to develop strains that are better adapted to the co-culture environment.

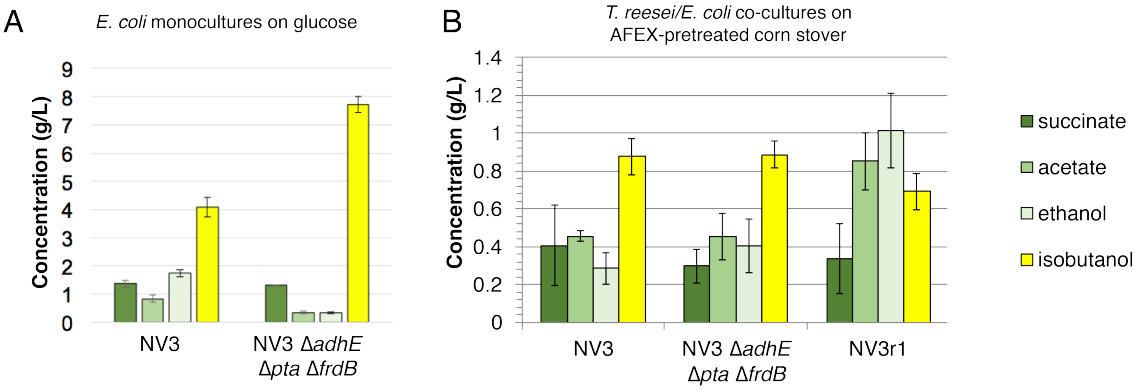

Figure 5.2: Major fermentation production of co-cultures with *E. coli* modified to knock out competing pathways. (A) *E. coli* monoculture fermentation products after 95 h fermentation in 36 g/L glucose M9 isobutanol production medium. Error bars represent standard deviation across two replicates. (B) Major fermentation products of *T. reesei/E. coli* co-cultures after 384 h on 20 g/L AFEX-pretreated corn stover. Error bars represent standard deviation across two replicates.

5.2.2 Toward adaptive evolution of *T. reesei/E. coli* co-cultures

We began our adaptive evolution strategy by growing mutagenized populations of *E. coli* wildtype K12 population strains and growing them in media made from *T. reesei/E. coli* co-culture supernatant supplemented with norvaline. This strategy was appealing since

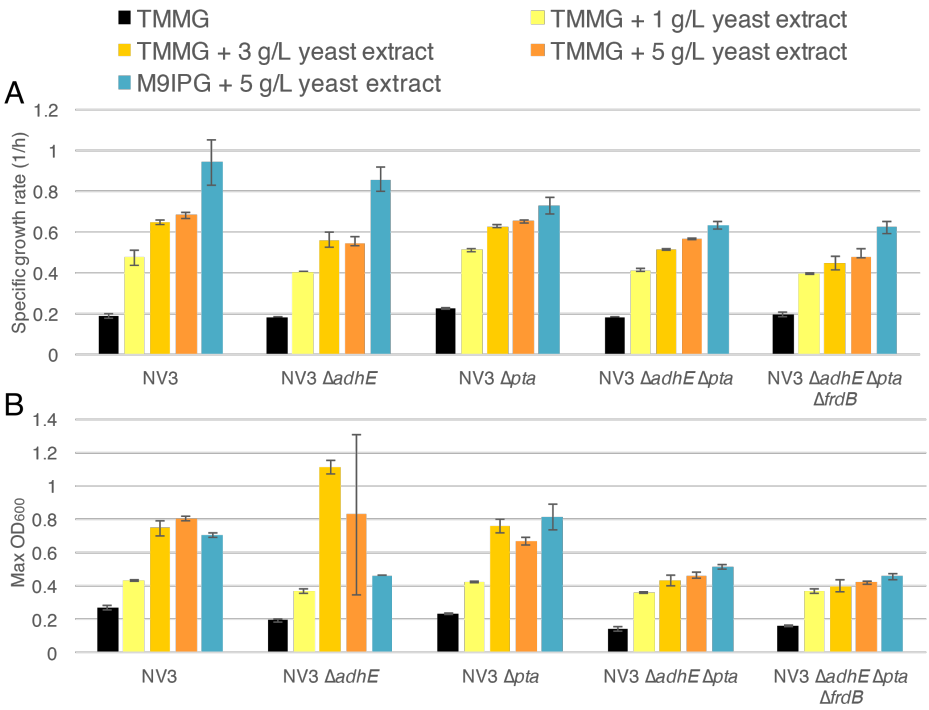

Figure 5.3: Growth properties of NV3 Δ*adhE*, NV3 Δ*pta*, NV3 Δ*adhE* Δ*pta*, Δ*adhE* Δ*pta* Δ*frdB* on *Trichoderma* minimal medium with glucose, pH 6 (TMMG) (A) or M9 isobutanol production medium (M9IPG) (B) with specified amounts of yeast extract in a 96-well microplate. Error bars represent standard deviation of three biological replicate wells.

growth is faster on glucose, allowing more generations of selection in a given time span, and because in monocultures *E. coli* growth can be screened using optical density. The resulting strains that showed improved growth under these conditions showed improved growth relative to K12 pSA65/69, which does not produce any isobutanol in co-culture with *T. reesei*, but strains growing on 12 g/L norvaline in supernatant only produced up to several hundred mg/L in co-culture, significantly less than the NV3(r1) strains.

We therefore undertook selection and screening in co-culture with *T. reesei*. Our *T. reesei* did show inhibition by norvaline (Data not shown). To screen for *E. coli* growth rate in co-culture, we have developed a fluorescence-based assay. The gene encoding yellow fluorescent protein (YFP) under a constitutively expressed promoter was integrated by P1 transduction into the genome of an NV3r1 isolate from the end of a co-culture with high isobutanol productivity (NV3r1*) and of wild-type K12. It is known that K12 has a higher

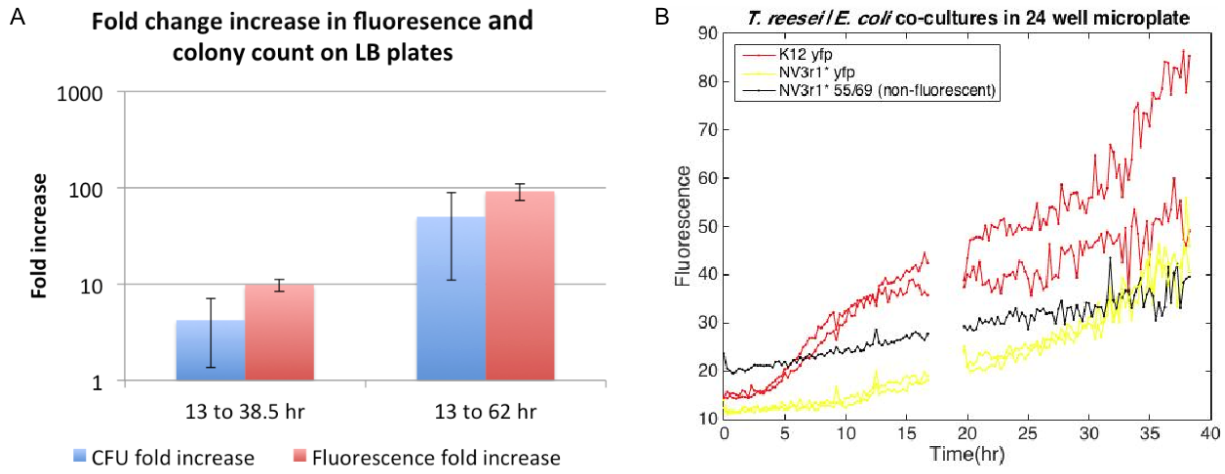

Figure 5.4: Use of fluorescence for monitoring *E. coli* growth in *T. reesei/E. coli* co-culture. (A) Fold increase in fluorescence (relative to a non-fluorescently labeled co-culture) and in viable *E. coli* cells. Error bars represent propagated standard deviations from two biological replicates. (B) Growth curve of *E. coli* K12 *yfp* and NV3r1* *yfp* in 24-well microplate measured by fluorescence (two replicates of each), along with a non-fluorescent control. The three-hour gap represents a period where the plate continued to be incubated and shaken, but the measurements were lost due to instrument error.

growth rate in co-culture with *T. reesei* than NV3 strains [53]. We therefore tested to see if by growing these strains in co-culture in a 24-well black microplate (600 μL per 3 mL well volume) and measuring the fluorescence (excitation: 485 nm, emission: 528 nm) we could observe the difference in growth rate. Flask cultures were conducted simultaneously and samples were periodically serially diluted and plated on LB agar plates to verify the difference in *E. coli* growth rates. Samples from the flask cultures were also read for fluorescence (in 96-well plates, so numbers are not directly comparable to the 24-well plate readings).

The data shown in Fig. 5.4A indicates that the culture's fluorescence is a reasonable quantitative indicator of *E. coli* count during the growth phase. The microplate growth curves (Fig. 5.4B) also show a clear growth advantage during the early hours for K12 *yfp* in comparison to NV3r1* *yfp*. Although microplate conditions clearly differ significantly from those of a flask in terms of aeration and shaking (later during the cultivation the microplate cultures begin to form clumps, disrupting fluorescence measurement), this experiment indicates that the microplate growth curve correlates with the flask growth rate in the co-culture

environment and that we can use relative microplate growth rates with norvaline to screen for the best growing isolates.

We mutagenized NV3r1* using N-Methyl-N'-nitro-N-nitrosoguanidine (NTG), following the procedure of [11]. We then inoculated co-cultures with approximately 7x10^8 *E. coli* cells and *T. reesei* RUT-C30 for an approximate initial ratio of 0.1 gDW *E. coli*:1 gDW RUT-C30. By taking samples and measuring the fluorescence in 96-well microplates, we observe growth in the norvaline-supplemented cultures (Fig. 5.5A). After 210 hours, when *E. coli* appeared to have reached stationary phase and fluorescence no longer increased, we subcultured into fresh media. After 218 hr, the second subculture was serially diluted and plated in on TMM 20 g/L MCC plates, TMM 20 g/L MCC plates with norvaline, and on LB plates. *E. coli* colonies were observed on the MCC cellulose plates and were isolated for transformation with the isobutanol pathway plasmids. However, upon testing in co-culture for isobutanol production, they did not show improvement compared to the parent. We hypothesize that using a strain with the isobutanol pathway integrated into the genome may allow for a better outcome of the adaptive evolution since the co-culture can be immediately testing, instead of isolating the *E. coli* into monoculture for plasmid transformation, which may be counter-selecting for the traits we desire for co-culture performance. It would also allow for co-evolution of the two strains since the population can be collectively tested without isolating monocultures.

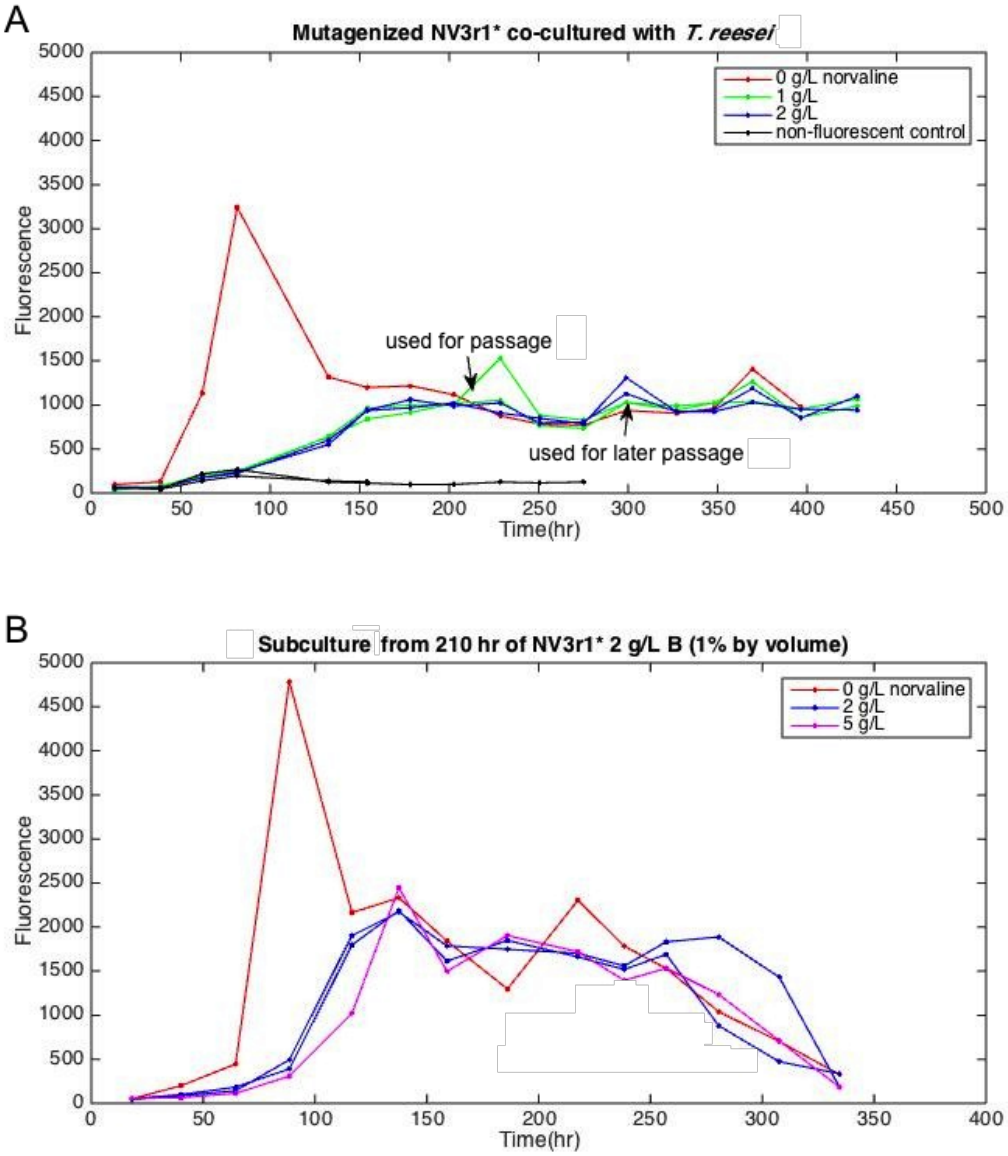

Figure 5.5: Fluorescence of co-cultures with mutagenized *E. coli* population. (A) Fluorescence of NTG-mutagenized NV3r1* populations grown in co-culture with *T. reesei* on 0, 1, or 2 g/L norvaline in 50 mL volume. Samples were read in a 96-well black microplate (200 μL per well). (B) Fluorescence of subculture #1 flasks, inoculated 1% by volume from a NV3r1* *yfp* culture on 2 g/L norvaline at 210 hr.

5.2.3 Investigation of β-glucosidase addition to co-cultures

On the hydrolysis side, we hypothesized that β-glucosidase might be a limiting factor in performance since it is commonly found to be a bottleneck in hydrolysis [112]. To test this, we added commercial β-glucosidase to co-cultures. While this addition led to a maintained increase in β-glucosidase activity over the culture period, it did not boost isobutanol production (Fig. 5.6). Sugar liberation may not be the limiting factor affecting fermentation performance.

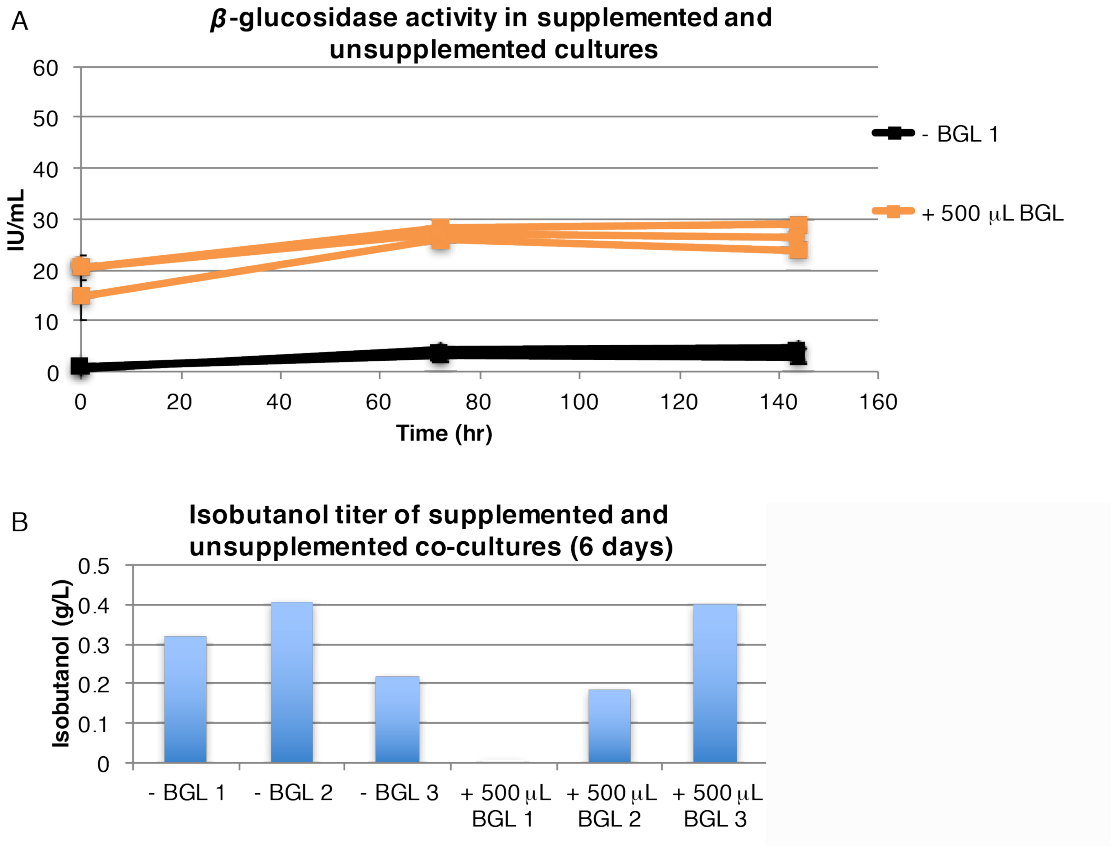

Figure 5.6: *T. reesei/E. coli* NV3r1* co-cultures with and without β-glucosidase supplementation. (A) β-glucosidase activity over time. (B) Isobutanol titers after 6 days of culturing. (A) and (B) show data for the same set of cultures (three replicates of each condition).

5.3 Discussion and conclusion

In this chapter, we have examined potential strategies for improving *T. reesei/E. coli* co-culture isobutanol production performance. The strategy for eliminating competing pathways proved ineffective at increasing co-culture production levels. Under the low glucose, low pH conditions of the co-culture, the *E. coli* cells are presumably in a very different state than in high glucose, ideal monoculture conditions. Due to the complexity of the conditions, we expect an evolutionary strategy may be a more effective approach. We demonstrated that fluorescence could be used to monitor *E. coli* growth while it is co-cultured with RUT-C30. We expect that this can be an effective strategy when employed with an *E. coli* strain that contains a stably integrated isobutanol pathway. Utilizing SnoCAP to screen for strains that produce well under co-culture conditions also appears to be a promising approach.

5.4 Materials and methods

5.4.1 Strains and plasmids

Strains and plasmids used in the work presented in this chapter are listed in Tables 5.1 and 5.2.

5.4.2 Gene deletions

Gene deletions were constructed by P1 phage transduction as previously described in [82, 83] and in Chapter 2. Keio strains were used as the donor strains and LB agar with 50 μg/mL kanamycin was used as the selective medium. Transductants were purified from residual P1 phage by isolation streaking on LB agar supplemented with 0.8 mM sodium citrate and 50 μg/mL kanamycin. The FRT-flanked kanamycin resistance gene was using the pCP20 plasmid to express FLP-recombinase.

5.4.3 Co-cultures

Medium preparation and culture setup were performed as in [53]. Cultures were of 10 mL volume in 125 mL polypropylene baffled flasks. IPTG (an added concentration of 0.1 mM) was added to cultures every 3-5 days. Glucose and fermentation products were assessed by HPLC, as described in Chapter 2. For the evolution experiments, co-cultures were cultured in 600 μL volume in 24-well black, clear-bottom microplates in an SpectraMax M5 plate reader (Molecular Devices) at 30 °C, with shaking, on TMM, buffered with 0.1 M maleate-NaOH, pH 6, 20 g/L MCC.

5.4.4 *E. coli* monoculture growth characterization

Cells were grown overnight in LB, washed twice with PBS, resuspended to an optical density corresponding to $\sim 10^9$ CFU/mL, and inoculated 1:100 into the specified medium. Cultures were vortexed and then distributed into a 96-well clear microplate (Brand), 200 μL per well. The microplate lid was coated with a solution of 0.5% Triton X-100 in 20% ethanol to reduce condensation and lids were fastened with tape. The microplate was incubated at 30 °C, with shaking in a VersaMax plate reader (Molecular Devices), with absorbance readings at 600 nm taken every 10 min. μmax was calculated via linear regression of natural log of OD_{600} values (after subtracting blank values) vs. time.

5.4.5 β-glucosidase activity

β-glucosidase activity was measured by reaction on 4-nitrophenyl β-D-glucopyranoside (pNPG) to produce a colored product, as previously described [113]. In a 96-well microplate, 25 μL of appropriately diluted sample was mixed with 25 μL ultrapure water and μL 200 mM pH 4.8 citrate buffer and incubated at 50 °C for 5 minutes. 25 μL 10 mM pNPG was then added and the plate was incubated 10 min at 50 °C. 100 μL of NaOH-glycine buffer (0.4 M, pH 10.8) was added to stop the reactions and absorbance was measured at 405 nm. One IU is defined as the amount of enzyme required to release 1 μmole of pNP per minute

(extinction coefficient = 18,000 1/(M*cm).

Table 5.1: Strains employed in the studies described in Chapter 5

Strain	**Notes**	**Reference**
NV3	Norvaline resistant mutant strain	[11]
NV3r1	NV3 with *rpoS* mutation repaired	[11]
NV3r1 Δ*adhE* Δ*pta* Δ*frdB*	NV3r1 Δ*adhE*::FRT Δ*pta*::FRT Δ*frdB*::FRT	This study
NV3r1*	Possibly mutated version of NV3r1, isolated from a *T. reesei*/*E. coli* co-culture	This study
NV3r1* *yfp*	NV3r1* Δ*intC::yfp-cat*	This study
K12 *yfp*	K12 Δ*intC::yfp-cat*	This study
RUT-C30	*Trichoderma reesei* strain engineered for hypersecretion of cellulases	ATCC 56765; [55]

Table 5.2: Plasmids used in the studies described in Chapter 5

Plasmid	**Relevant genotype**	**Reference**
pSA55	ColE1 ori; AmpR; $P_L$lacO1::*kivd–ADH2*	[10]
pSA65	ColE1 ori; AmpR; $P_L$lacO1::*kivd–adhA*	[81]
pSA69	p15A ori; KanR; $P_L$lacO1::*alsS–ilvC–ilvD*	[81]

Chapter 6: Concluding remarks and future directions

6.1 Summary

This dissertation has demonstrated the use of microbial cross-feeding systems as a tool for high-throughput screening of strain libraries, as well as explored strategies for optimizing strains for co-culture-enabled consolidated processing of lignocellulosic biomass. In the first part of the work, we have developed a screening framework based on cross-feeding auxotrophs, Syntrophic Co-culture Amplification of Production phenotype (SnoCAP). This framework has unique features that are not possible in a monoculture system. The cross-feeding configuration maintains the target molecule as limiting, thus producing a wider and more tunable dynamic range than that of a monocultured auxotrophic biosensor. Additionally, the exponential growth of the culture amplifies small differences in production level between strains into much more significant differences in growth phenotype of the co-cultures. We demonstrate three formats of compartmentalizing individual secretor genotypes with sensor strain: confinement in wells of microtiter plates, spatial separation as colonies on agar plates, and encapsulation in water-in-oil microdroplets. We also demonstrated that the dynamic range of the assay may be tuned by the addition of an inhibitory analog and that, by a push-pull strategy, SnoCAP can be extended to screening for secondary metabolites that are several steps removed from a primary metabolite. As demonstrations of SnoCAP's utility, we screened model libraries consisting of mixtures of two production level strains. We also employed the agar plate format of SnoCAP to screen a chemically mutagenized li-

brary and identified an efficient isobutanol producing strain. This strain nearly matches the production levels of the double plasmid strain, but it carries only one plasmid. The pyruvate to isobutanol conversion in this strain is accomplished by genomic mutations rather than by overexpression of the pathway genes.

In the second part of this work, we considered strategies for improving a synthetic fungal-bacterial co-culture for consolidated bioprocessing of lignocellulosic biomass into the biofuel isobutanol. One approach we took was integration of the isobutanol pathway genes into the bacterial genome, which is desirable for stability and cost-reduction. We first integrated the two genes responsible for the conversion of 2-ketoisovalerate (2-KIV) into isobutanol and increased their copy number by chemically inducible chromosomal evolution (CIChE), achieving similar production levels to the strain containing these genes on a plasmid. We next developed a method of optimizing expression level by random integration into genome positions with varying expression levels followed by SnoCAP screening. From a library of *alsS* integrants, we identified strains with significantly higher production levels than the parent strain, although not as high as the strain with the plasmid. This is to be expected since the plasmid overexpresses three genes that convert pyruvate to 2-KIV and we only integrated one of them. We also developed a strategy for adaptive laboratory evolution of the fungal-bacterial co-culture that we expect can be useful toward developing a more robust and productive process.

6.2 Future directions

6.2.1 Future directions for the SnoCAP platform

There is increasing interest in production of many chemicals from renewable resources and the number of compounds that are commercially produced using microorganisms is on the rise. As discussed in Chapter 2 Discussion and conclusion, we expect SnoCAP can be readily applied to a variety of target compounds, including many for which other types of

biosensors do not currently exist. Lee *et al.* [114] have recently compiled a comprehensive bio-based chemicals map of compounds that can be produced by biological methods or a combination of biological and non-environmentally harmful chemical methods from the sugars present in lignocellulosic biomass. They also include a list of those that are being produced on commercial scale and the best reported titers. Many of the compounds in this map have not yet been explored as metabolic engineering targets and future work on producing them can potentially benefit from SnoCAP screening.

We expect that constructing an auxotrophic sensor strain and identifying an appropriate partner auxotrophy will in some cases require far less work than development of a protein- or nucleic acid-based sensor and also offers the benefits of signal amplification and a tunable dynamic range. To speed the development of a new syntrophic pair-base screen, it may be useful to apply predictive modeling for identifying appropriate complementary auxotrophies. Wintermute & Silver took a flux-balance analysis (FBA) modeling approach to find shadow prices representing the benefit to a strain from consuming its required nutrient, b^A, and the cost, p^B in terms of lost growth rate from secreting metabolite A. They defined a cross-feeding strain's cooperation efficiency as $\epsilon_B = b^A/p^B$ and found that this was predictive of the co-culture behavior (Fig. 6.1). This approach can potentially be useful in choosing an appropriate secondary cross-fed molecule for a given target production molecule. The best initial level of cross-feeding with which to begin a SnoCAP screening will likely depend on the target molecule and how much improvement is sought. For example, if a large degree of improvement is expected then it will be beneficial to begin with a strain that has a low level of initial co-growth with the parent strain.

A key challenge of metabolic engineering is the lack of predictive power. To better inform computational models, we need to collect more data. High-throughput screening methods, such as SnoCAP, combined with high-throughput sequencing to characterize library members of different production levels can potentially help to reveal underlying principles that can guide future metabolic engineering efforts.

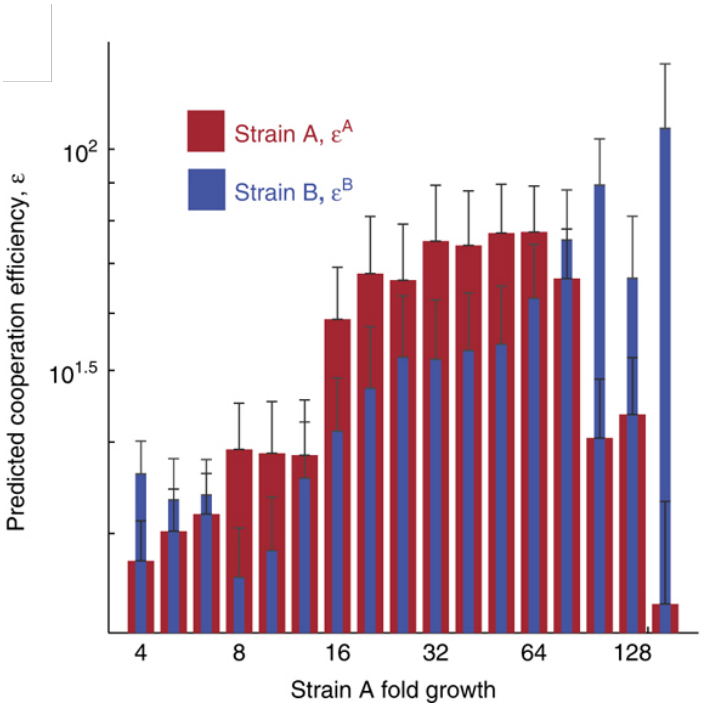

Figure 6.1: Prediction of co-growth properties based on flux-balance analysis (FBA) modeling. Cooperation efficiencies, ϵ were derived from a FBA model and are a measure of how well the strain can cheaply produce metabolites that promote high-growth of the partner auxotroph. Reproduced from [34].

SnoCAP for screening for *E. coli* productivity in co-culture with *T. reesei* A future direction of SnoCAP relevant to the work presented in this dissertation could be its application toward screening *E. coli* libraries for improved production when co-cultured with *Trichoderma* on a cellulosic carbon source. The *Trichoderma* would liberate glucose which would be consumed by itself, as well as by both *E. coli* strains (Fig. 6.2A). The co-growth should be determined by the productivity of the secretor strain, allowing screening for strains that are productive under the stresses of co-culture conditions (including limited glucose, suboptimal pH).

To test the feasibility of this, we cultured RUT-C30 spores with the sensor strain and either the high- or low-producing strain from Chapter 2 in tricultures in 200 μL volumes in 96-well microplates in TMM medium with 20 g/L cellobiose. The plate was incubated at 30 °C, with shaking, in a plate reader, reading optical density and green fluorescence at intervals. We observed that the high-producing *E. coli* strain led to increased total optical

density and fluorescence, as predicted (Fig. 6.2B). A control with only the sensor *E. coli* (no secretor strain) showed an eventual increase in fluorescence, indicating that it can get some of its needed metabolites from the *Trichoderma*. The low-producing secretor with sensor produced similar profiles with the sensor only.

We next tested for co-growth in droplets by encapsulating the cells such that all droplets contain *Trichoderma* spores and sensor *E. coli* cells and ∼10% contain a secretor cell. The cells are again in TMM medium with cellobiose. In this case, the *Trichoderma* spores germinated rapidly (with 12 hrs), but the *E. coli* did not show noticeable growth after 50 h. By this time the *Trichoderma* have grown outside of the droplets (Fig. 6.2C). It may be that this strain has too low of a production level under these settings to promote growth.

6.2.2 Future directions for *T. reesei/E. coli* co-culture development

It is clear from natural communities that microbial consortia are capable of performing complex tasks efficiently and robustly. Accumulation of a value-added product requires engineering a synthetic consortium. Despite significant research efforts to develop design principles for synthetic consortia, there are still relatively few examples of industrially implemented consortia, which highlights the need for further development of tools for engineering stable, productive co-cultures. In this dissertation we have begun investigation of isobutanol pathway integration and optimization, which we predict can lead to a more stable co-culture that can be subjected to further optimization by adaptive laboratory evolution.

Future work could be done on this project to integrate the *ilvCD* genes, combine this library with the *alsS* library, and screen by SnoCAP to identify even higher 2-KIV/isobutanol producers. Assessing the expression level and genomic location of these integrants should shed light on optimal expression levels of these genes and may provide useful knowledge for future pathway integration efforts. Screening can also be performed under different conditions (e.g., low pH conditions which are more optimal for the *T. reesei*). Ultimately, the best 2-KIV producers can then be combined with the CIChE integration of *kivd/adhA*

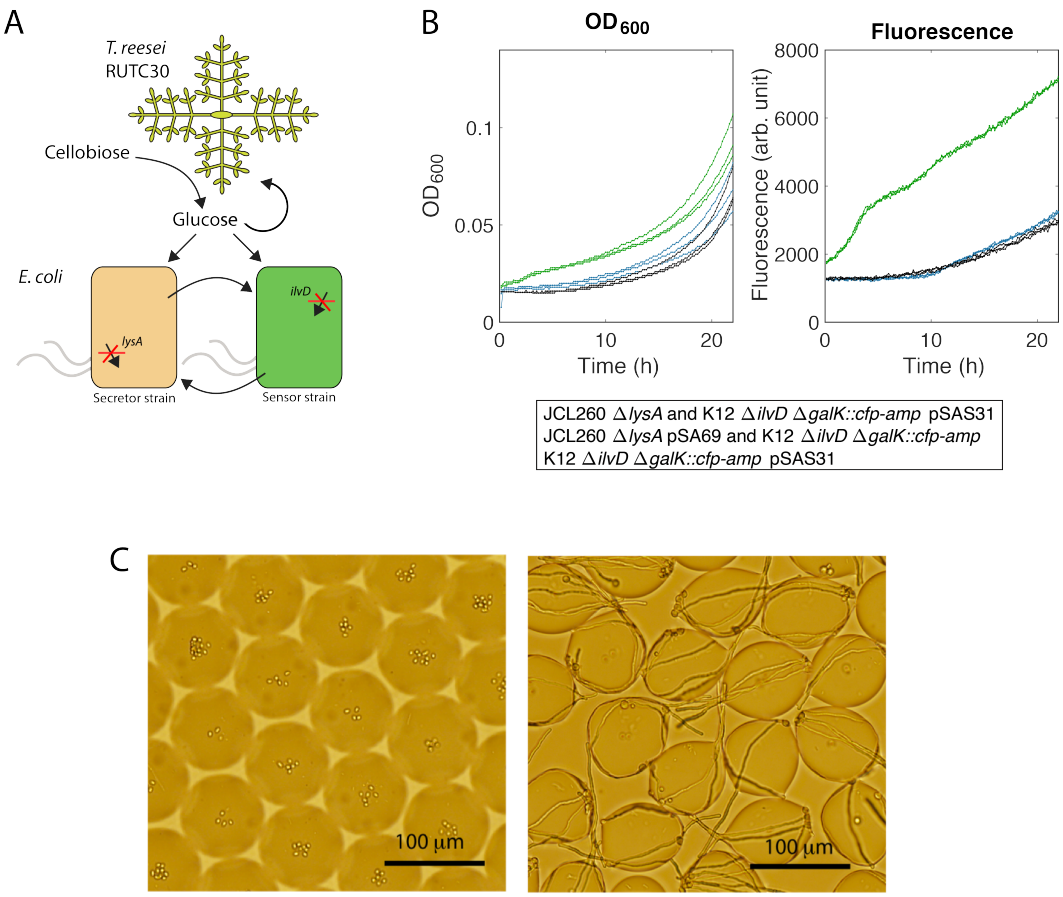

Figure 6.2: Preliminary investigation of SnoCAP for screening *E. coli* for productivity under co-culture with *T. reesei*. (A) Schematic of expected SnoCAP implementation with *Trichoderma*. (B) OD_{600} and fluorescence profiles of tricultures and a biculture with only the sensor *E. coli*. The replicates of each culture are shown, plotted in the same color. Due to setup error, the plate sat for ∼7 hr at room temperature before beginning the kinetic run in the plate reader. Therefore by the time indicated as 0 h there has likely already been some growth. Legend indicates the *E. coli* strains present in the culture; all cultures also contain RUT-C30. (C) Tricultures before and after 50 h incubation in droplets.

to produce fully integrated strains. We expect these strains will be good starting points for adaptive evolution efforts for isobutanol production when co-cultured with *Trichoderma reesei*.

References

[1] E.M.T. El-Mansi, C.F.A. Bryce, B. Dahhou, S. Sanchez, A.L. Demain, and A.R. Allman, editors. *Fermentation Microbiology and Biotechnology*. CRC Press, 2011.

[2] Madhavan Nampoothiri and Ashok Pandey. Genetic tuning of coryneform bacteria for the overproduction of amino acids. *Process Biochemistry*, 33(2):147–161

[3] Benjamin M. Woolston, Steven Edgar, and Gregory Stephanopoulos. Metabolic engineering: past and future. *Annual Review of Chemical and Biomolecular Engineering*, 4(1):259–288, 2013.

[4] Wenshan Liu and Rongrong Jiang. Combinatorial and high-throughput screening approaches for strain engineering. *Applied Microbiology and Biotechnology*, 99(5):2093–2104 0175–7598, 2015.

[5] Harris H. Wang, Farren J. Isaacs, Peter A. Carr, Zachary Z. Sun, George Xu, Craig R. Forest, and George M. Church. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*, 460(7257):894, 2009.

[6] Jing Du, Yongbo Yuan, Tong Si, Jiazhang Lian, and Huimin Zhao. Customized optimization of metabolic pathways by combinatorial transcriptional engineering. *Nucleic Acids Research*, 40(18):e142–e142

[7] Brian F. Pfleger, Douglas J. Pitera, Christina D. Smolke, and Jay D. Keasling. Combinatorial engineering of intergenic regions in operons tunes expression of multiple genes. *Nature Biotechnology*, 24(8):1027

[8] Chris J. Paddon and Jay D. Keasling. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology*, 12(5):355

[9] Douglas J. Pitera, Chris J. Paddon, Jack D. Newman, and Jay D. Keasling. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metabolic Engineering*, 9(2):193–207, 2007.

[10] Shota Atsumi, Taizo Hanai, and James C. Liao. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*, 451(7174):86, 2008.

[11] Kevin M. Smith and James C. Liao. An evolutionary strategy for isobutanol production strain development in *Escherichia coli*. *Metabolic Engineering*, 13(6):674–681, 2011.

[12] Antonino Baez, Kwang-Myung Cho, and James C Liao. High-flux isobutanol production using engineered escherichia coli: a bioreactor study with in situ product removal. *Applied microbiology and biotechnology*, 90(5):1681–1690 0175–7598, 2011.

[13] Sabine Bastian, Xiang Liu, Joseph T. Meyerowitz, Christopher D. Snow, Mike M.Y. Chen, and Frances H. Arnold. Engineered ketol-acid reductoisomerase and alcohol dehydrogenase enable anaerobic 2-methylpropan-1-ol production at theoretical yield in *Escherichia coli*. *Metabolic engineering*, 13(3):345–352

[14] James E. Bailey, Adriana Sburlati, Vassily Hatzimanikatis, Kelvin Lee, Wolfgang A. Renner, and Philip S. Tsai. Inverse metabolic engineering: a strategy for directed genetic engineering of useful phenotypes. *Biotechnology and Bioengineering*, 52(1):109–121, 1996.

[15] Jeffrey A. Dietrich, Adrienne E. McKee, and Jay D. Keasling. High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annual Review of Biochemistry*, 79:563–590, 2010.

[16] Stacy-Anne Morgan, Dana C. Nadler, Rayka Yokoo, and David F. Savage. Biofuel metabolic engineering with biosensors. *Current Opinion in Chemical Biology*, 35:150–158, 2016.

[17] Jyun-Liang Lin, James M. Wagner, and Hal S. Alper. Enabling tools for high-throughput detection of metabolites: metabolic engineering and directed evolution applications. *Biotechnology Advances*, 35(8):950–970, 2017.

[18] Christine Nicole S Santos and Gregory Stephanopoulos. Melanin-based high-throughput screen for l-tyrosine production in *Escherichia coli*. *Applied and Environmental Microbiology*, 74(4):1190, 02 2008.

[19] Annamaria Ruscito and Maria C. DeRosa. Small-molecule binding aptamers: selection strategies, characterization, and applications. *Frontiers in Chemistry*, 4:14

[20] Jeremy S. Paige, Thinh Nguyen-Duc, Wenjiao Song, and Samie R. Jaffrey. Fluorescence imaging of cellular metabolites with RNA. *Science*, 335(6073):1194–1194

[21] J.W. Payne, G. Bell, and C.F. Higgins. The use of an *Escherichia coli* Lys-auxotroph to assay nutritionally available lysine in biological materials. *Journal of Applied Microbiology*, 42(2):165–177, 1977.

[22] V.I. Chalova, W.K. Kim, C.L. Woodward, and S.C. Ricke. Quantification of total and bioavailable lysine in feed protein sources by a whole-cell green fluorescent protein growth-based *Escherichia coli* biosensor. *Applied Microbiology and Biotechnology*, 76(1):91–99, 2007.

[23] Moon Il Kim, Tae Jung Park, Nam Su Heo, Min-Ah Woo, Daeyeon Cho, Sang Yup Lee, and Hyun Gyu Park. Cell-based method utilizing fluorescent *Escherichia coli* auxotrophs for quantification of multiple amino acids. *Analytical Chemistry*, 86(5):2489–2496, 2014.

[24] Vesela I. Chalova, Sujata A. Sirsat, Corliss A. O'Bryan, Philip G. Crandall, and Steven C. Ricke. *Escherichia coli*, an intestinal microorganism, as a biosensor for quantification of amino acid bioavailability. *Sensors*, 9(9):7038–7057, 2009.

[25] Felix Bertels, Holger Merker, and Christian Kost. Design and characterization of auxotrophy-based amino acid biosensors. *PLoS One*, 7(7):e41349, 2012.

[26] P.R. Burkholder. Determination of vitamin B12 with a mutant strain of *Escherichia coli*. *Science*, 114:459–460, 1951.

[27] Thomas Hahn, Kristina Tag, Klaus Riedel, Steffen Uhlig, Keith Baronian, Gerd Gellissen, and Gotthard Kunze. A novel estrogen sensor based on recombinant *Arxula adeninivorans* cells. *Biosensors and Bioelectronics*, 21(11):2078–2085, 2006.

[28] Brian F. Pfleger, Douglas J. Pitera, Jack D. Newman, Vincent J.J. Martin, and Jay D. Keasling. Microbial sensors for small molecules: development of a mevalonate biosensor. *Metabolic Engineering*, 9(1):30–38, 2007.

[29] Brandon E.L. Morris, Ruth Henneberger, Harald Huber, and Christine Moissl-Eichinger. Microbial syntrophy: interaction for the common good. *FEMS Microbiology Reviews*, 37(3):384–406, 2013.

[30] Glen D'Souza, Shraddha Shitut, Daniel Preussger, Ghada Yousif, Silvio Waschina, and Christian Kost. Ecology and evolution of metabolic cross-feeding interactions in bacteria. *Natural Product Reports*, 35(5):455–488, 2018.

[31] K.C. Winkler, A.W. van Doorn, and A.F. Royers. Symbiosis of tryptophan-deficient mutants of *E. coli* B. *Recueil des Travaux Chimiques des Pays-Bas*, 71(1):5–14, 1952.

[32] Veikko Nurmikko. Biochemical factors affecting symbiosis among bacteria. *Experientia*, 12(7):245–284, 1956.

[33] P. Fildes. Production of tryptophan by *Salmonella typhi* and *Escherichia coli*. *Microbiology*, 15(3):636–642, 1956.

[34] Edwin H. Wintermute and Pamela A. Silver. Emergent cooperation in microbial metabolism. *Molecular Systems Biology*, 6:407–407, 2010.

[35] Michael T. Mee, James J. Collins, George M. Church, and Harris H. Wang. Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences*, 111(20):E2149–E2156, 2014.

[36] Vinuselvi Parisutham, Tae Hyun Kim, and Sung Kuk Lee. Feasibilities of consolidated bioprocessing microbes: from pretreatment to biofuel production. *Bioresource Technology*, 161:431–440, 2014.

[37] Vishnu Menon and Mala Rao. Trends in bioconversion of lignocellulose: biofuels, platform chemicals & biorefinery concept. *Progress in Energy and Combustion Science*, 38(4):522–550, 2012.

[38] Lee R. Lynd, Willem H. Van Zyl, John E McBride, and Mark Laser. Consolidated bioprocessing of cellulosic biomass: an update. *Current Opinion in Biotechnology*, 16(5):577–583, 2005.

[39] Daniel G. Olson, John E. McBride, A. Joe Shaw, and Lee R. Lynd. Recent progress in consolidated bioprocessing. *Current Opinion in Biotechnology*, 23(3):396–405, 2012.

[40] Lee R. Lynd, Mark S. Laser, David Bransby, Bruce E. Dale, Brian Davison, Richard Hamilton, Michael Himmel, Martin Keller, James D. McMillan, John Sheehan, et al. How biotech can transform biofuels. *Nature Biotechnology*, 26(2):169–172, 2008.

[41] Trevor R. Zuroff and Wayne R. Curtis. Developing symbiotic consortia for lignocellulosic biofuel production. *Applied Microbiology and Biotechnology*, 93(4):1423–1435, 2012.

[42] Ran Du, Jianbin Yan, Shizhong Li, Lei Zhang, Sandra Zhang, Jihong Li, Gang Zhao, and Panlu Qi. Cellulosic ethanol production by natural bacterial consortia is enhanced by *Pseudoxanthomonas taiwanensis*. *Biotechnology for Biofuels*, 8(1):10, 2015.

[43] Trevor R. Zuroff, Salvador Barri Xiques, Wayne R. Curtis, et al. Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium phytofermentans*/yeast co-culture. *Biotechnology for Biofuels*, 6(1):59, 2013.

[44] Jasmine Shong, Manuel Rafael Jimenez Diaz, and Cynthia H. Collins. Towards synthetic microbial consortia for bioprocessing. *Current Opinion in Biotechnology*, 23(5):798–802, 2012.

[45] Hao Song, Ming-Zhu Ding, Xiao-Qiang Jia, Qian Ma, and Ying-Jin Yuan. Synthetic microbial consortia: from systematic analysis to construction and applications. *Chemical Society Reviews*, 43(20):6954–6981, 2014.

[46] Jihyang Park, Alissa Kerner, Mark A. Burns, and Xiaoxia Nina Lin. Microdroplet-enabled highly parallel co-cultivation of microbial communities. *PLoS One*, 6(2):e17019, 2011.

[47] Woon Sun Choi, Dokyeong Ha, Seongyong Park, and Taesung Kim. Synthetic multicellular cell-to-cell communication in inkjet printed bacterial cell systems. *Biomaterials*, 32(10):2500–2507, 2011.

[48] Woon Sun Choi, Minseok Kim, Seongyong Park, Sung Kuk Lee, and Taesung Kim. Patterning and transferring hydrogel-encapsulated bacterial cells for quantitative analysis of synthetically engineered genetic circuits. *Biomaterials*, 33(2):624–633, 2012.

[49] Jing Zhang, Jingwen Zhou, Jie Liu, Kejie Chen, Liming Liu, and Jian Chen. Development of chemically defined media supporting high cell density growth of *Ketogulonicigenium vulgare* and *Bacillus megaterium*. *Bioresource Technology*, 102(7):4807–4814, 2011.

[50] Niels Klitgord and Daniel Segrè. Environments that induce synthetic microbial ecosystems. *PLoS Computational Biology*, 6(11):e1001002, 2010.

[51] Shen-Long Tsai, Garima Goyal, and Wilfred Chen. Surface display of a functional minicellulosome by intracellular complementation using a synthetic yeast consortium and its application to cellulose hydrolysis and ethanol production. *Applied and Environmental Microbiology*, 76(22):7514–7520, 2010.

[52] Simone Brethauer and Michael Hanspeter Studer. Consolidated bioprocessing of lignocellulose by a microbial consortium. *Energy & Environmental Science*, 7(4):1446–1453, 2014.

[53] Jeremy J. Minty, Marc E. Singer, Scott A. Scholz, Chang-Hoon Bae, Jung-Ho Ahn, Clifton E. Foster, James C. Liao, and Xiaoxia Nina Lin. Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. *Proceedings of the National Academy of Sciences*, 110(36):14592–14597, 2013.

[54] Mathew A. Rude and Andreas Schirmer. New microbial fuels: a biotech perspective. *Current Opinion in Microbiology*, 12(3):274–281, 2009.

[55] Robyn Peterson and Helena Nevalainen. *Trichoderma reesei* RUT-C30–thirty years of strain improvement. *Microbiology*, 158(1):58–68, 2012.

[56] Bingming Ou, Carolina Garcia, Yejun Wang, Weiping Zhang, and Guoqiang Zhu. Techniques for chromosomal integration and expression optimization in *Escherichia coli*. *Biotechnology and Bioengineering*, 115(10):2467–2478, 2018.

[57] Chung-Jen Chiang, Po Ting Chen, and Yun-Peng Chao. Replicon-free and markerless methods for genomic insertion of DNAs in phage attachment sites and controlled expression of chromosomal genes in *Escherichia coli*. *Biotechnology and Bioengineering*, 101(5):985–995, 2008.

[58] Juergen Mairhofer, Theresa Scharl, Karoline Marisch, Monika Cserjan-Puschmann, and Gerald Striedner. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. *Applied and Environmental Microbiology*, 79(12):3802–3812, 06 2013.

[59] Keith E.J. Tyo, Parayil Kumaran Ajikumar, and Gregory Stephanopoulos. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nature Biotechnology*, 27(8):760–765, 2009.

[60] Pengfei Gu, Fan Yang, Tianyuan Su, Qian Wang, Quanfeng Liang, and Qingsheng Qi. A rapid and reliable strategy for chromosomal integration of gene(s) with multiple copies. *Scientific Reports*, 5:9684, 2015.

[61] Jin Yin, Huan Wang, Xiao-Zhi Fu, Xue Gao, Qiong Wu, and Guo-Qiang Chen. Effects of chromosomal gene copy number and locations on polyhydroxyalkanoate synthesis by *Escherichia coli* and *Halomonas* sp. *Applied Microbiology and Biotechnology*, 99(13):5523–5534 0175–7598, 2015.

[62] Jacob A. Englaender, J. Andrew Jones, Brady F. Cress, Thomas E. Kuhlman, Robert J. Linhardt, and Mattheos A.G. Koffas. Effect of genomic integration location on heterologous protein expression and metabolic engineering in *E. coli*. *ACS Synthetic Biology*, 6(4):710–720

[63] Hironaga Akita, Nobutaka Nakashima, and Tamotsu Hoshino. Bacterial production of isobutanol without expensive reagents. *Applied Microbiology and Biotechnology*, 99(2):991–999, 2015.

[64] Marcelo C. Bassalo, Andrew D. Garst, Andrea L. Halweg-Edwards, William C. Grau, Dylan W. Domaille, Vivek K. Mutalik, Adam P. Arkin, and Ryan T. Gill. Rapid and

efficient one-step metabolic pathway integration in *E. coli*. *ACS Synthetic Biology*, 5(7):561–568, 07 2016.

[65] Adrian D. Haimovich, Paul Muir, and Farren J. Isaacs. Genomes by design. *Nature Reviews Genetics*, 16(9):501–516, 2015.

[66] Miso Park, Shen-Long Tsai, and Wilfred Chen. Microbial biosensors: engineered microorganisms as the sensing machinery. *Sensors*, 13(5):5777–5795, 2013.

[67] Naama Tepper and Tomer Shlomi. Computational design of auxotrophy-dependent microbial biosensors for combinatorial metabolic engineering experiments. *PloS One*, 6(1):e16274, 2011.

[68] Xize Niu, Fabrice Gielen, Joshua B. Edel, and Andrew J. deMello. A microdroplet dilutor for high-throughput screening. *Nature Chemistry*, 3:437, 2011.

[69] Alissa Kerner, Jihyang Park, Audra Williams, and Xiaoxia Nina Lin. A programmable *Escherichia coli* consortium via tunable symbiosis. *PLoS One*, 7(3):e34032, 2012.

[70] Xiaolin Zhang and Jennifer L. Reed. Adaptive evolution of synthetic cooperating communities improves growth performance. *PloS One*, 9(10):e108297, 2014.

[71] Ji-Hu Zhang, Thomas D.Y. Chung, and Kevin R. Oldenburg. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of Biomolecular Screening*, 4(2):67–73, 1999.

[72] Samay Pande, Shraddha Shitut, Lisa Freund, Martin Westermann, Felix Bertels, Claudia Colesie, Ilka B. Bischofs, and Christian Kost. Metabolic cross-feeding via intercellular nanotubes among bacteria. *Nature Communications*, 6:6238, 2015.

[73] Zhi-Jun Zhao, Chun Zou, Yi-Xing Zhu, Jun Dai, Sheng Chen, Dan Wu, Jing Wu, and Jian Chen. Development of L-tryptophan production strains by defined genetic

modification in *Escherichia coli*. *Journal of Industrial Microbiology & Biotechnology*, 38(12):1921–1929, 2011.

[74] Peter A. Carr, Harris H. Wang, Bram Sterling, Farren J. Isaacs, Marc J. Lajoie, George Xu, George M. Church, and Joseph M. Jacobson. Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Research*, 40(17):e132–e132, 2012.

[75] Olle Karlström. Methods for the production of mutants suitable as amino acid fermentation organisms. *Biotechnology and Bioengineering*, 7(2):245–268, 1965.

[76] F. Valle and A. Berry. *Metabolic Engineering of Escherichia coli for the Production of Aromatic Compounds. In: Metabolic Engineering, Lee, S. Y., and Papoutsakis, E. T. Eds. Marcel Dekker, Inc., New York, 79 (1999).* Metabolic Engineering. Taylor & Francis, 1999.

[77] João G.R. Cardoso, Ahmad A. Zeidan, Kristian Jensen, Nikolaus Sonnenschein, Ana Rute Neves, and Markus J. Herrgård. MARSI: Metabolite analogues for rational strain improvement. *Bioinformatics*, 34(13):2319–2321, 2018.

[78] Mingyue Fang, Tianmin Wang, Chong Zhang, Jili Bai, Xiang Zheng, Xuejin Zhao, Chunbo Lou, and Xin-Hui Xing. Intermediate-sensor assisted push–pull strategy and its application in heterologous deoxyviolacein production in escherichia coli. *Metabolic Engineering*, 33:41–51, 2016.

[79] Wenying Shou, Sri Ram, and Jose M.G. Vilar. Synthetic cooperation in engineered yeast populations. *Proceedings of the National Academy of Sciences*, 104(6):1877–1882, 2007.

[80] William R. Harcombe, Jeremy M. Chacón, Elizabeth M. Adamowicz, Lon M. Chubiz, and Christopher J. Marx. Evolution of bidirectional costly mutualism from byproduct

consumption. *Proceedings of the National Academy of Sciences*, 115(47):12000–12004, 2018.

[81] Shota Atsumi, Tung-Yun Wu, Eva-Maria Eckl, Sarah D. Hawkins, Thomas Buelter, and James C. Liao. Engineering the isobutanol biosynthetic pathway in *Escherichia coli* by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied Microbiology and Biotechnology*, 85(3):651–657, 2010.

[82] J.H. Miller. *A short course in bacterial genetics: A laboratory manual and handbook for Escherichia coli and related bacteria*. Cold Spring Harbor Laboratory Press, Plainview, NY, 1992.

[83] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A. Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2(1), 2006.

[84] Simanti Datta, Nina Costantino, and Donald L. Court. A set of recombineering plasmids for gram-negative bacteria. *Gene*, 379:109–115, 2006.

[85] Joshua Lederberg. A simple method for isolating individual microbes. *Journal of Bacteriology*, 68(2):258, 1954.

[86] Tomasz S. Kaminski, Ott Scheler, and Piotr Garstecki. Droplet microfluidics for microbiology: techniques, applications and challenges. *Lab on a Chip*, 16(12):2168–2187, 2016.

[87] Jean-Christophe Baret, Oliver J. Miller, Valerie Taly, Michaël Ryckelynck, Abdeslam El-Harrak, Lucas Frenz, Christian Rick, Michael L. Samuels, J. Brian Hutchison, and Jeremy J. Agresti. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab on a Chip*, 9(13):1850–1858, 2009.

[88] Benjamin L. Wang, Adel Ghaderi, Hang Zhou, Jeremy Agresti, David A. Weitz, Gerald R. Fink, and Gregory Stephanopoulos. Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. *Nature Biotechnology*, 32(5):473, 2014.

[89] Solvej Siedler, Narendar K. Khatri, Andrea Zsohár, Inge Kjærbølling, Michael Vogt, Petter Hammar, Christian F. Nielsen, Jan Marienhagen, Morten O.A. Sommer, and Haakan N. Joensson. Development of a bacterial biosensor for rapid screening of yeast p-coumaric acid production. *ACS Synthetic Biology*, 6(10):1860–1869, 2017.

[90] Joseph Abatemarco, Maen F. Sarhan, James M. Wagner, Jyun-Liang Lin, Leqian Liu, Wafa Hassouneh, Shuo-Fu Yuan, Hal S. Alper, and Adam R. Abate. RNA-aptamers-in-droplets (RAPID) high-throughput screening for secretory phenotypes. *Nature Communications*, 8(1):332, 2017.

[91] James M. Wagner, Leqian Liu, Shuo-Fu Yuan, Maya V. Venkataraman, Adam R. Abate, and Hal S. Alper. A comparative analysis of single cell and droplet-based FACS for improving production phenotypes: riboflavin overproduction in *Yarrowia lipolytica*. *Metabolic Engineering*, 47:346–356, 2018.

[92] Ming Li, Mark van Zee, Carson T. Riche, Bobby Tofig, Sean D. Gallaher, Sabeeha S. Merchant, Robert Damoiseaux, Keisuke Goda, and Dino Di Carlo. A gelatin microdroplet platform for high-throughput sorting of hyperproducing single-cell-derived microalgal clones. *Small*, 14(44):e1803315, 2018.

[93] Irit Levin-Reisman, Orit Gefen, Ofer Fridman, Irine Ronin, David Shwa, Hila Sheftel, and Nathalie Q. Balaban. Automated imaging with ScanLag reveals previously undetectable bacterial growth phenotypesscanlag reveals previously undetectable bacterial growth phenotypes. *Nature Methods*, 7(9):737–739, 2010.

[94] I. Levin-Reisman, O. Fridman, and N.Q. Balaban. ScanLag: High-throughput quantification of colony growth and lag time. *Journal of Visual Experiments*, 89(e51456), 2014.

[95] Marc Harper and Christopher J. Lee. Genome-wide analysis of mutagenesis bias and context sensitivity of N-methyl-N'-nitro-N-nitrosoguanidine (NTG). *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 731(1):64–67, 2012.

[96] Jimin Zheng and Zongchao Jia. Structure of the bifunctional isocitrate dehydrogenase kinase/phosphatase. *Nature*, 465(7300):961, 2010.

[97] E.M. el-Mansi, C. MacKintosh, K. Duncan, W.H. Holms, and H.G. Nimmo. Molecular cloning and over-expression of the glyoxylate bypass operon from *Escherichia coli* ML308. *The Biochemical Journal*, 242(3):661–665, 03 1987.

[98] Paula Arense, Vicente Bernal, Daniël Charlier, JoséLuis Iborra, Maria Remedios Foulquié-Moreno, and Manuel Cánovas. Metabolic engineering for high yielding L(-)-carnitine production in *Escherichia coli*. *Microbial Cell Factories*, 12(1):56, 2013.

[99] Frédéric Grenier, Dominick Matteau, Vincent Baby, and Sébastien Rodrigue. Complete genome sequence of *Escherichia coli* BW25113. *Genome Announcements*, 2(5):e01038–14, 2014.

[100] Sida Wang. *Characterizing Microbial Communities Using Droplet Microfluidic Technology*. PhD thesis, University of Michigan, 2017.

[101] Tobias W. Hofmann, Siegfried Hänselmann, Jan-Wilhelm Janiesch, Anne Rademacher, and Christian H.J. Böhm. Applying microdroplets as sensors for label-free detection of chemical reactions. *Lab on a Chip*, 12(5):916–922, 2012.

[102] Jeremy J. Agresti, Eugene Antipov, Adam R. Abate, Keunho Ahn, Amy C. Rowat, Jean-Christophe Baret, Manuel Marquez, Alexander M. Klibanov, Andrew D. Griffiths,

and David A. Weitz. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proceedings of the National Academy of Sciences*, 107(9):4004–4009, 2010.

[103] Todd P Lagus and Jon F Edd. High-throughput co-encapsulation of self-ordered cell trains: cell pair interactions in microdroplets. *Rsc Advances*, 3(43):20512–20522, 2013.

[104] Meng Ting Chung, Daniel Núñez, Dawen Cai, and Katsuo Kurabayashi. Deterministic droplet-based co-encapsulation and pairing of microparticles via active sorting and downstream merging. *Lab on a Chip*, 17(21):3664–3671, 2017.

[105] Lisa Mahler, Miguel Tovar, Thomas Weber, Susanne Brandes, Martin Michael Rudolph, Josef Ehgartner, Torsten Mayr, Marc Thilo Figge, Martin Roth, and Emerson Zang. Enhanced and homogeneous oxygen availability during incubation of microfluidic droplets. *RSC Advances*, 5(123):101871–101878, 2015.

[106] Fabrice Gielen, Raphaelle Hours, Stephane Emond, Martin Fischlechner, Ursula Schell, and Florian Hollfelder. Ultrahigh-throughput–directed enzyme evolution by absorbance-activated droplet sorting (AADS). *Proceedings of the National Academy of Sciences*, 113(47):E7383–E7389, 2016.

[107] Scott A. Scholz, Rucheng Diao, Michael B. Wolfe, Elayne M. Fivenson, Xiaoxia Nina Lin, and Peter L. Freddolino. High-resolution mapping of a standardized transcriptional reporter reveals dedicated high and low transcription domains in the *Escherichia coli* chromosome. *Under review*, 2018.

[108] Jihyang Park, Alissa Kerner, Mark A. Burns, and Xiaoxia Nina Lin. Microdroplet-enabled highly parallel co-cultivation of microbial communities. *PloS One*, 6(2):e17019, 2011.

[109] Yan-Yan Cui, Chen Ling, Yuan-Yuan Zhang, Jian Huang, and Jian-Zhong Liu. Production of shikimic acid from *Escherichia coli* through chemically inducible chromosomal evolution and cofactor metabolic engineering. *Microbial Cell Factories*, 13(1):21

[110] Dena H S Block, Razika Hussein, Lusha W Liang, and Han N Lim. Regulatory consequences of gene translocation in bacteria. *Nucleic Acids Res*, 40(18):8979–8992, Oct 2012.

[111] Jack A. Bryant, Laura E. Sellars, Stephen J.W. Busby, and David J. Lee. Chromosome position effects on gene expression in *Escherichia coli* K-12. *Nucleic Acids Research*, 42(18):11383–11392, 10 2014.

[112] Annette Sørensen, Mette Lübeck, Peter Lübeck, and Birgitte Ahring. Fungal beta-glucosidases: a bottleneck in industrial use of lignocellulosic materials. *Biomolecules*, 3(3):612–631, 2013.

[113] Shou Takashima, Akira Nakamura, Makoto Hidaka, Haruhiko Masaki, and Takeshi Uozumi. Molecular cloning and expression of the novel fungal β-glucosidase genes from *Humicola grisea* and *Trichoderma reesei*. *The Journal of Biochemistry*, 125(4):728–736 1999.

[114] Sang Yup Lee, Hyun Uk Kim, Tong Un Chae, Jae Sung Cho, Je Woong Kim, Jae Ho Shin, Dong In Kim, Yoo-Sung Ko, Woo Dae Jang, and Yu-Sin Jang. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis*, 2(1):18–33, 2019.