UNIVERSITY OF CALIFORNIA SAN DIEGO

The Characterization and Improvement of Gene Circuits for Cancer Diagnostics and Therapies

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Bioengineering

by

Aayush Somani

Committee in charge:

      Professor Jeff Hasty, Chair
      Professor Nan Hao
      Professor Rob Knight

2024

The Thesis of Aayush Somani is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AcCoA ....Acetyl Coenzyme A

AHL ....Acyl-homoserine Lactone

ALE.. Adaptive Laboratory Evolution

ATP.. Adenosine Triphosphate

AZT...3-azido-2′,3′-dideoxythymidine

CDS.. Coding Sequence

CGSC .... Coli Genetic Stock Center

CRAΦ ....Competence-Reducing Acinetobacter Phage

CRC..Colo-Rectal Cancer

DLG .... Diffusion Limited Growth

ECM...Extra-Cellular Matrix

EcN ....*E. coli* Nissle 1917

GFP ....Green Fluorescent Protein

HGT .... Horizontal Gene Transfer

HlyE ....Hemolysin E

IPTG...Isopropyl ß-D-1-thiogalactopyranoside

IS ....Insertion Sequence

KanR ....Kanamycin Resistance

LB ....Luria-Bertani

N1...Neutral 1

N2...Neutral 2

ON.. Overnight Culture

Ori .... Origin of Replication

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Jeff Hasty for his support as my PI and the chair of my committee. His feedback and mentorship really allowed me to flourish as a scientist in his lab, for which I am incredibly grateful. I would also like to thank the rest of my committee, Professor Nan Hao and Professor Rob Knight, for their feedback and help with my thesis projects.

I would also like to acknowledge all of my fellow Hasty lab members. I am grateful for their support and knowledge, especially when the experiments would not go my way. I would especially like to thank my mentors Joanna Zhang and Dr. Rob Cooper, as well as Austin Doughty for his insights into my patterning project. Additionally, many of the conclusions reached in the SLC patterning project would not have been possible without the expertise and knowledge of Dr. Lev Tsimring. I am also incredibly grateful to him for his help in creating the growth model, making a write up for it, and taking the time to explain the MATLAB code to me.

Lastly, I would like to thank my friends and family for helping me through the past year and providing the support necessary to make it through. I truly could not have done it without them.

ABSTRACT OF THE THESIS

The Characterization and Improvement of Gene Circuits for Cancer Diagnostics and Therapies

by

Aayush Somani

Master of Science in Bioengineering

University of California San Diego, 2024

Professor Jeff Hasty, Chair

A recent application for synthetic biological gene circuits has been in the realm of cancer diagnostics and therapeutics. Improvement and characterization of these circuits was conducted using three different methods. Firstly, using the synchronized lysis circuit (SLC), delivers cancer therapeutics in waves following oscillating lysis event, the spatiotemporal growth studied in motile bacteria[1,2]. The resulting patterns showed that the SLC strain was rendered immotile due

to the stress of the circuit while the wild type showed chemotactic behavior, avoiding the SLC colony. It was discovered that bacteria enter a state of stress caused by overflow metabolism due to high glucose concentrations, resulting in lowered cAMP levels and loss of flagellar development and forcing fimbriae-based movement. This combination of wild type fimbriae motility with high glucose levels and non-motility of SLC strains was used to further observe dual-strain motility which creates flower shaped patterns. Additionally, improvements were made to a biosensor, which utilized horizontal gene transfer (HGT) to sense genetic material indicating presence of colorectal cancer (CRC) tumors[3]. Replacement of the tetracycline repression mechanism with a *lacI* based repression system was attempted. Further stability improvements were made in the sensor by deleting a prophage sequence and several insertion sequence copies, which increased sensitivity to tumor DNA. Furthermore, bacteria evolved using adaptive laboratory evolution (ALE) were characterized for growth and circuit behavior non-traditional media. Lastly, an inducible GFP lysis circuit was developed to better compare the lysis response between two strains of *E. coli*, MG1655 and Nissle 1917.

1.1 The History of Synthetic Biology

The concept of synthetic biology dates back to 1973, where Stanley Cohen first created a plasmid and then ported it into *E. coli*[4]. The multidisciplinary field involves engineering genetic circuits with biological tools, akin to the approach of electrical circuit engineering. Researchers constantly investigate and map the fundamental building blocks of biology, ranging from DNA promoters, ribosome binding sites (RBS), coding sequences (CDS), and terminators. These sequences are then put together to create inherent gene networks within living organisms that are often highly complex. Consequently, synthetic biology adopts an approach that involves first deconstructing these intricate networks into simpler subunits, enabling the analysis of their dynamic interactions.
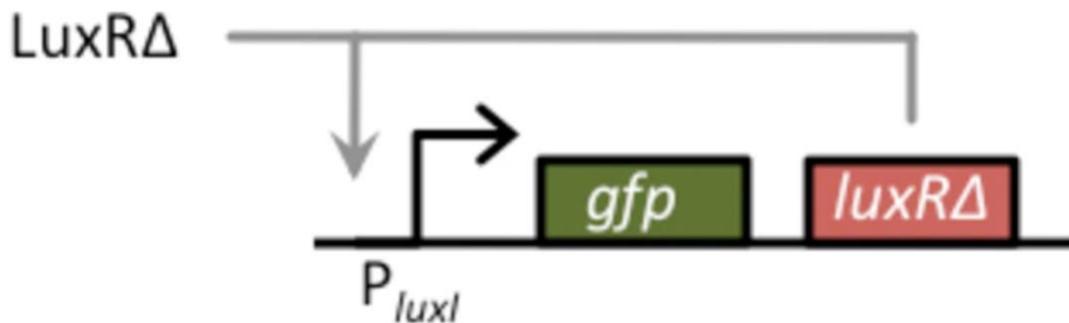
Figure 1.1: A basic gene circuit diagram, depicting a positive feedback loop[5]**.** The promoter *pLuxI* drives GFP and *luxR* production. *LuxR* then proceeds to act as a transcription factor (TF) to activate and further drive *pLuxI* transcription through a positive feedback loop. Figure adapted from Nistala et al.

The development of genetic circuits integrates mathematical modeling to first simulate the behavior, and then later building the model in vitro. Inversely, some methodologies involve first recognizing the phenomena in vitro, and then developing the model to prove understanding

of the process. Recent advancements in molecular biology and sequencing techniques facilitate the construction of genetic networks, allowing researchers to glean into multicellular system interactions by breaking it down into the simplest mechanisms. Furthermore, the systematic design of synthetic networks permits the extraction and concentration on specific subsystems within natural organisms. The analysis of these simpler systems, or modules, provides valuable insights into characterizing their behavior and contributions to underlying cellular functions. The capacity to isolate subsystems, such as toggle switches, feedback loops, and oscillators, offers the potential to extend these circuits into stackable modules to increase the complexity of the system.

The objective of synthetic biology as a whole is to build progressively complex networks that mimic natural systems while retaining the capability to experimentally isolate and test each component of the system. This approach enables the addition of complexity to basic circuits, allowing for experimental exploration, characterization, and optimization of these circuits. Synthetic gene networks, embedded in a cell's genome or ported into the chassis as a plasmid, have a range of diverse applications, including cellular biosensing, biomaterial synthesis, targeted apoptosis of tumor cells, and potential integration with bioelectronic applications. We look at two such gene networks in the projects detailed below, namely a negative feedback loop that can deliver cancer therapeutics in cycles, and a DNA biosensor which can detect cancerous genes.

1.2 The Synchronized Lysis Circuit (SLC)

The SLC was first developed by employing quorum sensing mechanisms found in *Vibrio fischeri*: a gram-negative, bioluminescent bacterium that uses the gene *luxI* to drive its luminescence, producing the small molecule acyl-homoserine lactone (AHL), which can easily diffuse across the cell membrane and enable intercellular coupling[1]. Originally, this led to the

development of a genetic clock circuit to couple GFP expression between cells whenever the

quorum threshold, set by the AHL level and cell population, was reached. The gene circuit for

the genetic clock is shown below in Fig. 1.2. A positive feedback loop utilizing quorum sensing

and a negative feedback loop that degrades AHL were combined to result in a quorum sensing

genetic clock that fully coupled the cell interaction and allowed for GFP fluorescence to be

brought to a singular time scaling, allowing for clear oscillatory behavior.
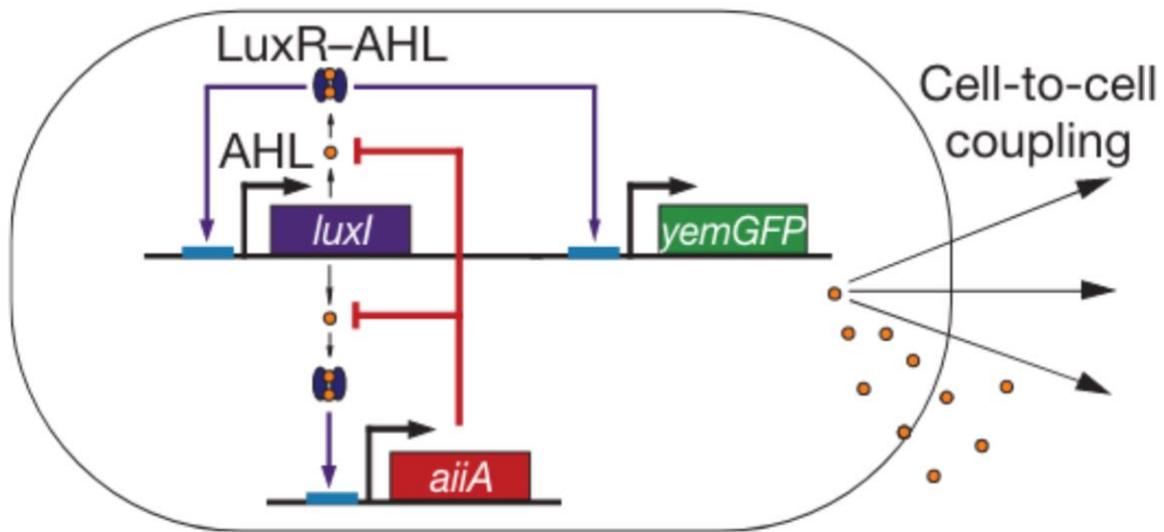
Figure 1.2: Gene circuit for a quorum sensing genetic clock[1]. *pLux* drives *luxI* production, which in turn makes AHL. *LuxR*, which is constitutively expressed, combines with AHL to become an activating TF to further drive *LuxI* and GFP production. Additionally, *LuxR*-AHL drives *aiiA* production, which degrades AHL. Figure adapted from Din et al.

With this system established, the researchers began to build up, replacing the negative

feedback loop with a lysing system, rather than degrading the AHL. Thus, the SLC was created,

with the colony reaching a quorum threshold, resulting in synchronous lysis of 90% of the

bacteria[2]. At the time of lysis, all proteins in the cell were released after the cells underwent lysis

and the 10% of cells left behind could begin to grow again, resulting in oscillatory behavior. The

protein release was a key feature that allowed the use of the SLC for cancer therapy.
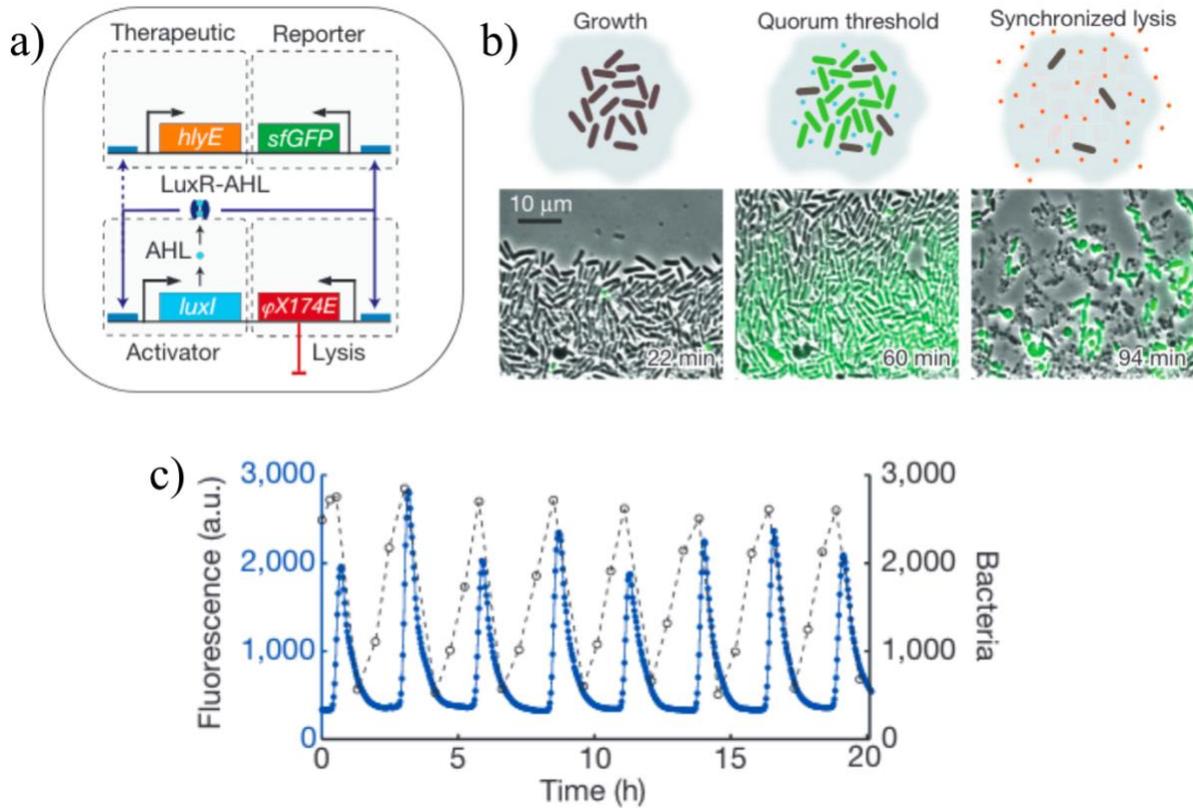
Figure 1.3: SLC diagram and behavior[2]. a) Circuit diagram for the SLC, describing the quorum sensing AHL based system, which drives production of a cancer therapeutic *hlyE* and GFP reporter, as well as the E lysis protein derived from a bacteriophage, ΦX174. b) Behavior of the SLC in a microfluidic chamber. At quorum, the entire trap fluoresces and immediately lyses, leaving behind 10% of the remaining population, which regrows and lyses again. c) Time course study of the fluorescence and optical density (OD) over time, depicting the oscillatory nature of the SLC. Figure adapted from Din et al.

As seen in Figure 1.3c, the SLC results in oscillatory behavior in the cells, with the peak

triggering the positive feedback quorum sensing loop and causing 90% lysis with the production

of X174E. The therapeutic agent, hemolysin E (*hlyE*), is a pore-forming toxin that attacks cell

wells to prevent synthesis[6].

1.2.1 SLC Application in Cancer Therapy

With this SLC created and tested, our lab began to branch into the realm of cancer therapies. The properties of *hlyE* allow its use as a cancer therapeutic. When the SLC is cocultured with tumor cells, the *hlyE* is released into the surrounding environment during lysis events. This *hlyE* toxin then targets the cell membrane of the cancerous cells and results in tumor size reduction. This proves that the SLC has applications as a cancer therapeutic since coculture with the tumor can successfully deliver toxins and eliminate growth. The added benefit of using the SLC to deliver *hlyE* is the repeated cycles of lysis to release the toxin into the surrounding environment. *HlyE* by itself is too large to be exported easily out of the cell, so the population lysis event is necessary to achieve delivery of the toxin, at a high concentration at once, to the cancer cells. Additionally, the SLC repeatedly lyses to achieve a cycle of *hlyE* buildup and release, simulating repeated dosage of the therapeutic. Previous research used a microfluidic testing model to coculture HeLa cells with bacteria harboring the SLC plasmid and constitutive *hlyE* production, and there was a clear reduction in cell viability, as seen in Figure 1.4.
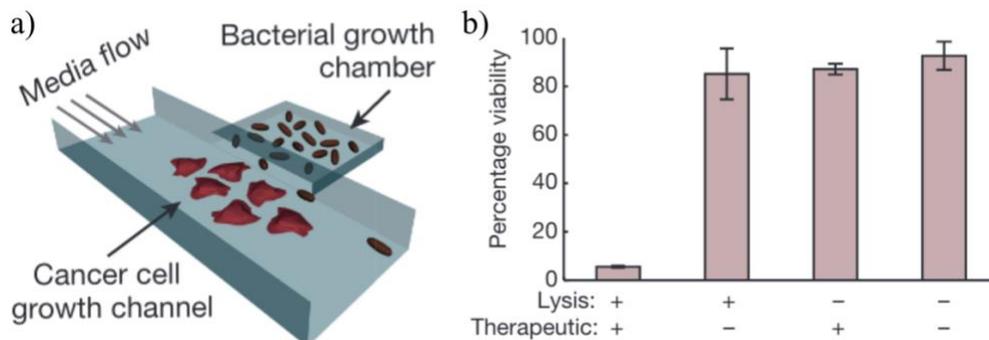
Figure 1.4: Microfluidic testing setup and HeLa cell viability results[2]. a) Microfluidic schematic of setup involving coculturing of HeLa cells and SLC in *S. Typhimurium*. b) Percentage viability of HeLa cells co-cultured with supernatant from *S. Typhimurium* culture with the SLC + *hlyE*, the SLC only, constitutive *hlyE* only, or no plasmid. Figure adapted from Din et al.

Additionally, oral delivery of the SLC bacteria combined with cytotoxic chemotherapy agent 5-FU, an antimetabolite, resulted in prolonged life for tumor model mice[2]. Since the SLC is aimed to be used in future studies as a potential cancer therapy, characterization of the strain is necessary. Understanding the pattern formation of the strain can lead to a clearer understanding of the mechanics and circuit behavior. Tumors are typically larger than the microfluidic traps circuits are studied in, so exploring the behavior of these circuits on a larger spatial scale like agar plates is important for characterization.

1.3 SLC Characterization via Patterning

Patterning experiments are quite popular in the realm of synthetic biology. They allow for a proof of concept for a circuit and allow for a 2-dimensional visualization of the circuit as the colony spreads on an agar plate. Sometimes, fractal patterns emerge allowing for the development of robust models that further characterize the circuit and allow for a better understanding of the functionality.

One such example of a patterning project is the work of a previous member of the lab, Liyang Xiong, who combined the non-motile *E. coli* and relatively more motile *A. baylyi* in a single spot on a plate to create flower shaped fractal patterns[7]. From this biological observation *in vitro*, he was then able to develop a robust model that accounted for the various diffusive properties of each bacterial strain and recreated the patterns mathematically. I aimed to follow a similar route with any patterns observed by plating the SLC on a medium that allowed for motility, by creating a mathematical model that could emulate the patterns formed *in vitro*.

As mentioned, SLC patterning is vital to understanding the mechanics of the chassis within the tumor environment. The results would be useful for the characterization of the circuit and ensure that mathematical models correlate to experimental models, proving an understanding of every aspect. Knowing how the bacteria carrying the SLC plasmid was moving within the

tumor environment, where delivery and treatment is a necessity, could prove essential in knowing how to fine-tune the parameters to ensure optimal delivery of the therapeutic. It would be incredibly helpful to be able to convert this into a mathematical model, so that the parameters are easily visualized and a more robust system can then be developed to ensure optimal delivery based on the surrounding conditions.

1.4 Materials and Methods

Initial experiments with the SLC patterning involved obtaining a motile form of MG1655. Previously, a group at the Laboratory of Genetics in Wisconsin sequenced a strain of *E. coli* which was known to have flagellar operons related to flagella production, which was then later submitted to the Yale Coli Genetic Stock Center (CGSC) under the identification number 7740[8]. This strain was obtained and transformed via heat shock with the SLC plasmids. Additionally, a positive feedback quorum sensing circuit with no lysis, only quorum sensing, was transformed for comparative study. A constitutive red fluorescent protein (RFP) with chloramphenicol resistance was also transformed into this strain so the plating could be conducted side by side on the same selection plate. This strain was labeled the quorum sensing control (QSC). Plate reader experiments on a Tecan plate reader were conducted to ensure the circuit behavior was as expected, which can be found in the appendix.
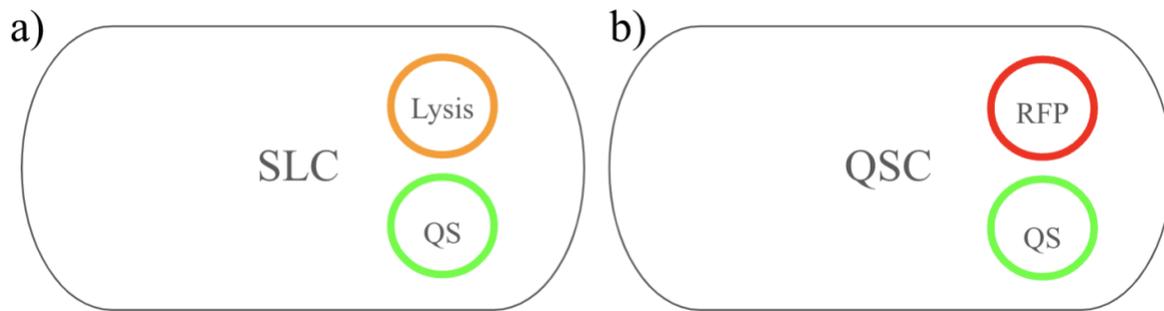
Figure 1.5: Diagram depicting plasmids in the SLC and QSC strains. a) The SLC contains a plasmid with the lysis gene driven by *pLuxI* and constitutive *hlyE* production. Additionally, there is the quorum sensing (QS) plasmid which has GFP, *luxR*, and *luxI* all driven by *pLuxI*. These plasmids are named, respectively, pZA35 and pTD103. See appendix A1 for more details. b) The QSC strain contains only the QS plasmid. To be able to see the strain on plate images and be able to use the same antibiotics for growth as the SLC, a plasmid with constitutive RFP and chloramphenicol resistance was also transformed.

Previous motility experiments with bacteria recommended an agar concentration of 0.5% to test swarming function of bacteria[9]. The normal accepted media for growing the strain was Luria-Bertani (LB) broth, which was made using the recommended concentration of 25 g/L. The appropriate amount of agar was added to reach 0.5% concentration. Additionally, a 0.35% LB agar was made to see if the agar density affected how fast the bacteria spread, and if any patterns seen were more striking.

For initial patterning experiments, a selection plate of kanamycin and chloramphenicol was made using the 0.5% LB agar. The SLC and QSC were grown overnight and then passaged at a 100X dilution the following morning. Once they grew to an OD of about 0.4, indicating early exponential phase, the two strains were plated side by side in 2 μL spots on the 0.5% LB agar. Two versions of the plate were made, one with 1% glucose to serve as a negative control for the SLC, and one without glucose. This is because literature suggests that glucose binds the CAP-cAMP activating complex found between the *luxI* and *luxR* promoters, which decreases or

slows *luxR* expression due to the inhibitory binding[2]. The intent for the glucose plate was to serve as a negative control where the SLC would not experience lysis or at least be at a slower timescale. Once spotted with the strains, the plate was left in an incubator at 37°C overnight and imaged the next day for patterns using a plate imager that can detect fluorescence.

1.5 Preliminary Goals and Results

Originally, the goal was to detect a pattern of interest with the SLC growing on a non-glucose plate. The constant death and regrowth cycle was expected to push dead bacteria to the outer edges or perhaps avoid areas where cells burst to see potential fractal patterns. However, this hypothetical pattern did not take place, with something else of interest happening instead.

The preliminary results were a bit surprising, in that they produced an unexpected pattern in the negative control, and a disinteresting pattern in the experimental plate. In Figure 1.6 below, the growth patterns are seen on 0.5% agar. The plate in Figure 1.6a was intended as the experimental setup to see the SLC working as intended, with no glucose inhibition preventing or slowing lysis. However, instead of the expected fractal patterns that result from dead bacteria being pushed away or avoided, the colony remained highly dense and relatively non-motile. Additionally, the expression of GFP indicates that the SLC colony has reached quorum, and the cells must be actively lysing. The QSC colony behaved similarly, expressing both RFP and GFP, also indicating it is at quorum and the constitutive RFP is working as intended. Meanwhile, in Figure 1.6b we see the plate that was intended as the control. Here we also see the SLC colony behave the same way, with a dense, non-motile, fluorescent spot. The QSC, however, spread in a concentric ring pattern, which breaks on the edge. There is only RFP expression, with no GFP expression, indicating the colony is not sensing it is at quorum. Additionally, the plate smelled distinctly of acetic acid which was observed when uncapping the plate for imaging. This

combination suggested that the QSC colony engaged in nutrient diffusion limited growth and overflow metabolism, which will be expanded upon in the coming chapters.
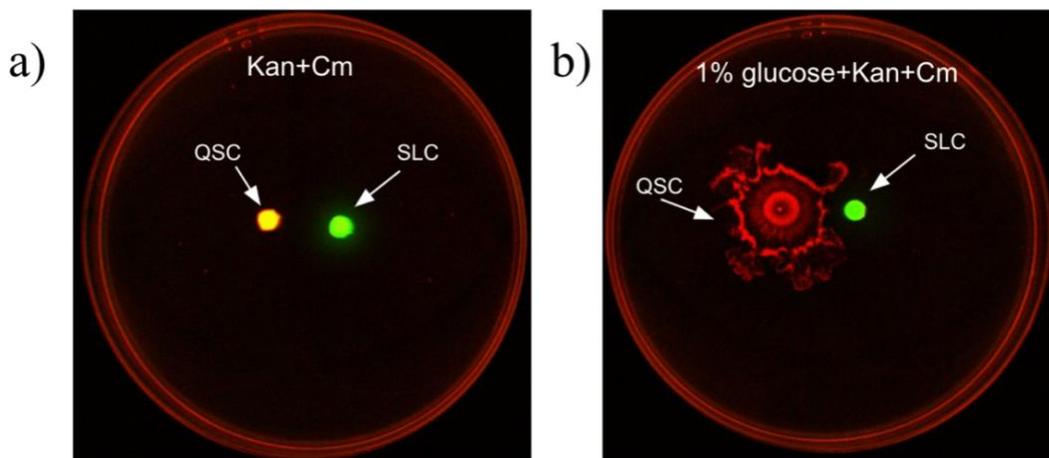
Figure 1.6: Initial patterning experiment with SLC and QSC on 0.5% LB agar plates. a) Plate without glucose, intended for a fully functioning SLC strain to be active. Result ended with little motility and highly dense colonies. b) Plate with glucose, intended as a negative control to slow SLC activity. The SLC strain stayed non-motile and dense, but the QSC colony made a strange concentric pattern that resembles nutrient diffusion limited growth. Red ring around the plate is an artifact of illumination.

Once this was observed in 0.5% LB agar, the experiment was performed again with 0.35% LB agar to see if the pattern could be amplified with less resistance for the bacteria's motility. Figure 1.7 shows the results of the experiment with 0.35% LB agar and glucose. Here, the same diffusion limited growth is seen in the QSC, with a switch after a certain point which could be due to reasons like a change in agar density from drying out or an internal clock. Additionally, amplification of this pattern allowed visualization of the QSC behavior around the SLC, as it seems to avoid the colony in a clear zone. This was especially interesting because it could indicate that under the stress caused by overflow metabolism and diffusion limited growth, the QSC colony becomes hypersensitive to the nearby SLC colony and avoids growth in that region. It could be due to the cell lysate, *hlyE*, or AHL, but these causes needed to be

investigated further. Additional experiments were conducted with pure wild type (WT) motile

bacteria and QSC colonies, to explore if the behavior was an effect of the circuit or the strain

itself.
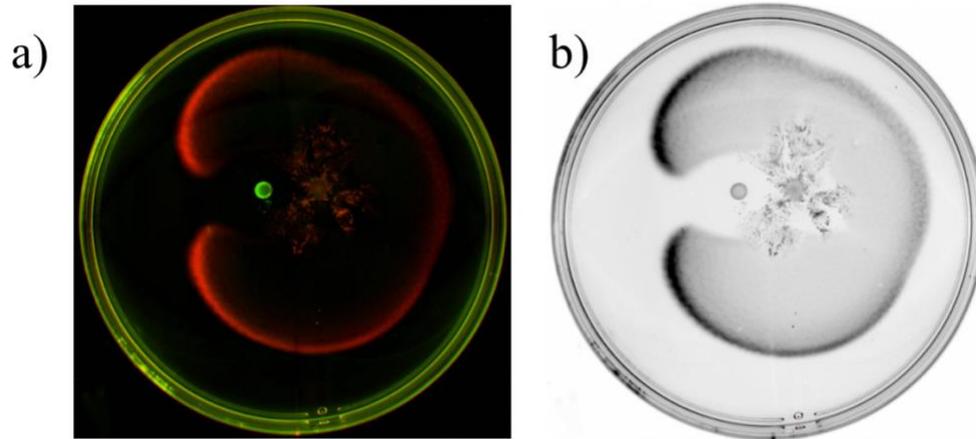
Figure 1.7: SLC and QSC patterned on 0.35% LB agar with 1% glucose. a) GFP and RFP
expression colorized, showing the same dense colony and non-motile behavior of the SLC. b)
The same plate shown with brightfield imaging, depicting the avoidance of the SLC colony
better, as well as the switch from the nutrient diffusion limited growth pattern to a smoother
pattern.

To test if the patterning were functions solely limited to the SLC and QSC dynamics, a

few combinations were plated. Firstly, two QSC colonies were plated side by side to see if the

avoidance only occurred when near an SLC colony. Additionally, motile and non-motile WT

MG1655 was plated to see if similar patterns occurred on glucose. Figure 1.8 and 1.9 detail the

experiments conducted. As seen below, Figure 1.8 shows that the QSC colonies grow into each

other without avoidance when no glucose is present. Additionally, there does seem to be a slight

repelling behavior between the two colonies when glucose is present, though the same pattern

does not occur with the motile colony surrounding the denser one as in Figure 1.7. These results

do seem to indicate that glucose causes a switch in colony growth dynamics which is further

investigated below. Figure 1.9 confirms that MG1655 7740, the motile WT strain, also forms

similar patterns with the presence of glucose, though the fluorescence is weaker due to no GFP.
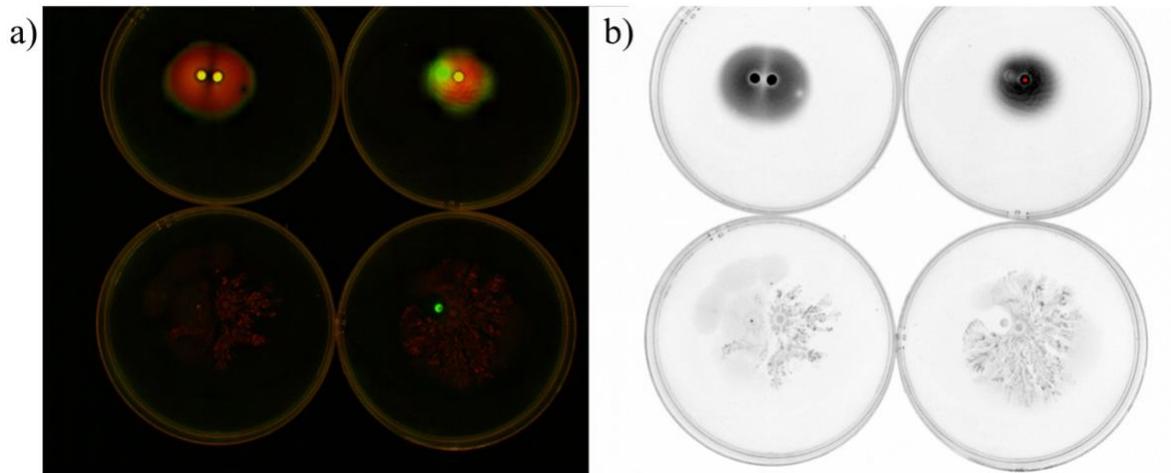
Figure 1.8: QSC and SLC colonies plated together with different combinations, with and without glucose. a) GFP and RFP visualized. b) Brightfield imaging. The top left plate shows two QSC colonies plated together with no glucose. Top right shows QSC and SLC plated together with no glucose. Bottom left shows two QSC colonies plated on a glucose plate, showing both the switch in patterning and a repelling behavior. Bottom right shows the previously observed pattern with a QSC and SLC colony on glucose.
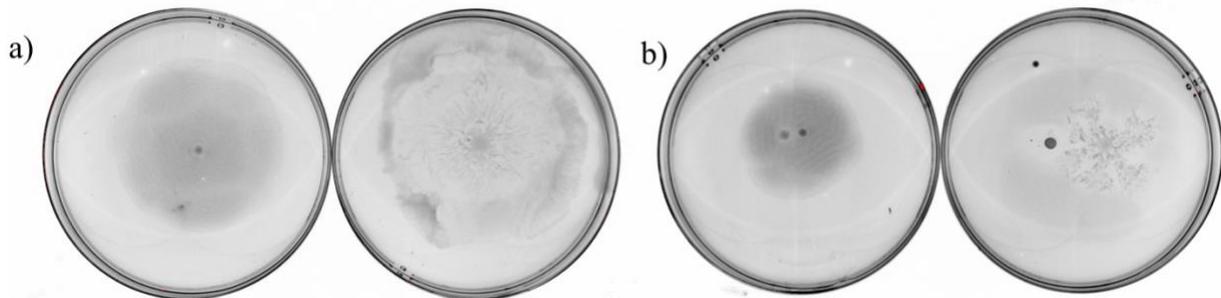
Figure 1.9: WT MG1655 7740 tests confirm similar behavior as QSC colony. a) Left plate shows WT on no glucose and right plate shows that on glucose, the WT strain forms a similar pattern as the QSC. b) When plated next to the SLC colony, the non-glucose (left) and glucose (right) plates show very similar behavior as in Figure 1.8.

This confirms that the behavior is a function of WT bacteria dynamics and glucose

concentration, rather than the specific QSC strain dynamics which was made. With the

patterning not being caused by the QSC itself, all that was left to test was if there was something

in the SLC colony specifically that was causing the avoidance of the other colony.

1.6 Nutrient Diffusion Limited Growth

The specific pattern formed by the WT and QSC strains on glucose needed to be further

characterized. Literature research shows that the pattern seen in the above figures resembles

nutrient diffusion limited growth (DLG). This effect is typically seen in *Bacillus subtilis* in

literature, but the patterns found in the figure below closely resembles the pattern seen by the

motile MG1655. DLG or diffusion-limited aggregation, as termed in the paper, is described as a

fractal growth pattern characterized by openly outward radiating structures that are randomly

branched[10].
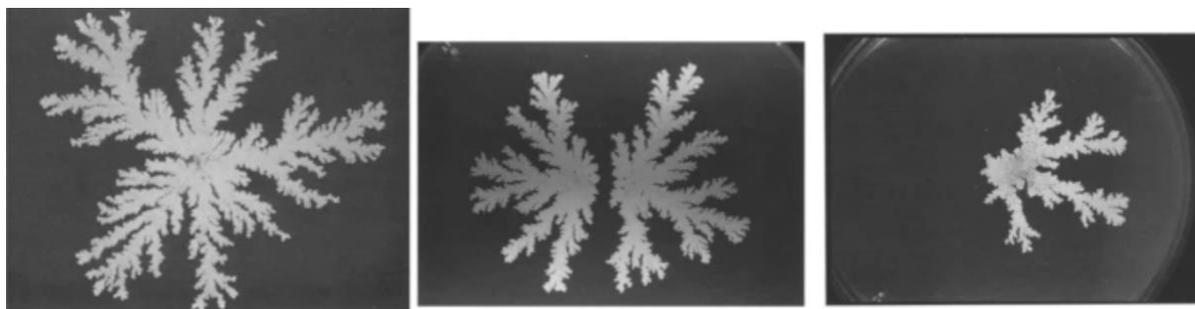
Figure 1.10: Nutrient diffusion limited growth patterns seen in *B. subtilis*[10]. The growth is
characterized by random branched structures radiating out from the center colony. From left to
right, we see three different patterns. The left image shows a single colony branching outward
randomly. The center image shows two colonies repelling each other. The right image shows
directed growth as the colony only grows in one direction. Figure adapted from Matsushita et al.

These patterns seen in *B. subtilis* closely resemble those seen formed in Figure 1.7,

before the switch from branching to smoother colony growth occurs. Additionally, the center

image in Figure 1.10 resembles the experiment in Figure 1.8 where two QSC colonies were

plated next to each other on a glucose plate. This suggests that the QSC may also be undergoing

DLG when placed on excess glucose. Nutrient DLG was found to have been characterized by the

simultaneous random branching of a colony in search of nutrients as well as the repulsion of two neighboring colonies. We can see the random branching in Figure 1.7, as well as the repulsion in Figure 1.8 with the 2 QSC colonies placed next to each other on glucose. It could also be that the dense SLC colony is a repelling point in the nutrient DLG behavior of the colonies.

Though the patterns resemble DLG growth, the conditions did not make sense for that growth to take place. DLG typically takes place where there is a limited supply of nutrient, but in this case, glucose was present in excess. Though it is possible that another nutrient was limited in the agar, there must have been some other mechanism forcing this type of growth in the colonies. The irregular growth pattern resembled a form of chaos, suggesting the colony had entered a state of stress causing the DLG pattern of growth, where the colony began to prioritize survival and made pathways following the highest nutrient source concentration. It would make sense for why the repelling behavior also took place, since survival of the fittest would prioritize spots in the agar not already occupied by an existing colony. This phenomenon needed to be further studied as DLG alone could not explain it given the existing conditions.

1.7 Overflow Metabolism in Cells

Another key observation was a strong smell of acetic acid when the plates were uncapped for imaging. Typically, the presence of acetic acid is a sign of anaerobic respiration, which is concerning since the plate is intended for aerobic respiration. Acetate is also an indicator of a stress response from bacteria, indicating the excess glucose was causing a state of stress for the strain. Literature shows that bacteria can engage in overflow metabolism, resulting in a metabolic imbalance and excess acetate production, when exposed to unlimited concentrations of glucose[11]. The biochemistry behind this phenomenon is explained via the loads imposed in the bacterial Krebs cycle. Under aerobic conditions, with unlimited glucose supply, *E. coli* displays a high substrate consumption rate, and consequently, a high growth rate[11]. This quickly results in a

14

metabolic imbalance, where excess acetyl coenzyme A (AcCoA) is generated and surpasses the threshold capacity for the tricarboxylic acid (TCA) cycle which generally is responsible for consuming the AcCoA and producing high energy adenosine triphosphate (ATP). Instead, the excess AcCoA is diverted to the phosphotransacetylase-acetate kinase pathway[12]. Here, acetate is a main byproduct, leading to eventual conversion to acetic acid, and the associated vinegar-like smell. In addition to the loss of productivity from conversion of the AcCoA, the acetate byproduct is known to be a toxic organic acid in the bacterial ecosystem. Hence, further loss in productivity occurs and the bacteria is also sent into a highly stressed state due to the production of this toxic molecule.

This accidental byproduct could be one explanation for why DLG occurs in these colonies with unlimited glucose source, as the acetate most likely sends the colony into a hyper-stressed state where it begins to treat glucose as a scarce commodity. Since acetate induces a stress response in bacteria, it makes sense that the effects of overflow metabolism result in chaotic growth, typically seen in stressed colonies that resort to chaotic patterns. This pattern is also seen in a study that grew *Bacillus thuringiensis* on 1% glucose plates. Seen below, the same fractal patterns emerge, with outward facing branches that seem to choose a random direction of growth.
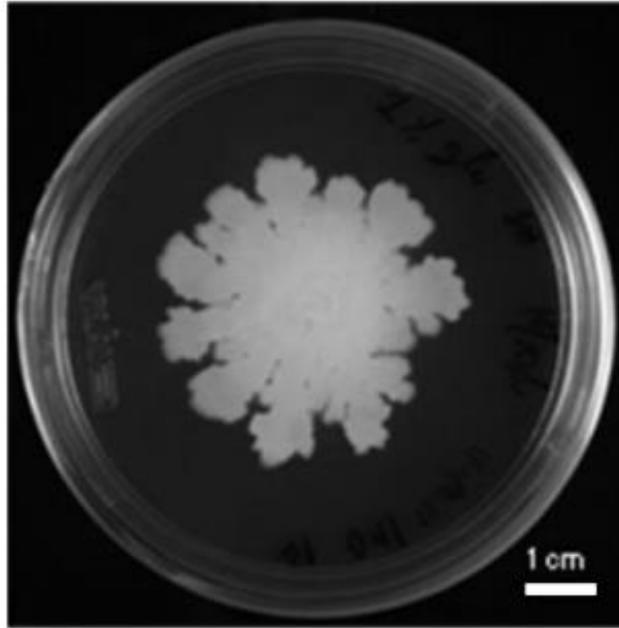
Figure 1.11: *Bacillus thuringiensis* **grown on a 1% glucose agar plate**[13]**.** Similar DLG patterns as the QSC and motile WT colonies form, with outward branching structures radiating in a seemingly random pattern. Figure adapted from Banerjee et al.

As seen in Figure 1.11, the bacteria form similar branching patterns that prove this type of growth most likely originates from an excess supply of glucose. The overflow metabolism results in excess acetate production, which is toxic to bacteria and induces stress, that most likely contributes to the pattern of growth observed. In order to confirm if acetate does in fact induce stressful growth in bacteria, I grew the QSC and SLC strains on 3% acetate plates. This amount was calculated to be around equimolar to the amount of carbon that would be supplied by 1% glucose, but also hopefully induce stress with the presence of acetate. I wanted to confirm that the presence of acetate induced random growth patterns that resembled DLG in the colonies.

Seen below in Figure 1.12 are the QSC and SLC colonies grown on acetate. It appears that acetate completely hindered motility in the QSC and none of the DLG patterning is seen at all. The SLC colony grows densely and has a slight region of growth diffusing outwards, so it appears that the growth was more than it would typically be on a glucose plate. There could be

an element of stochasticity here for the SLC being able to become more motile than the QSC colony, but the presence of acetate alone does not lead to DLG patterning and it actually results in a non-motile switch for the QSC colony.
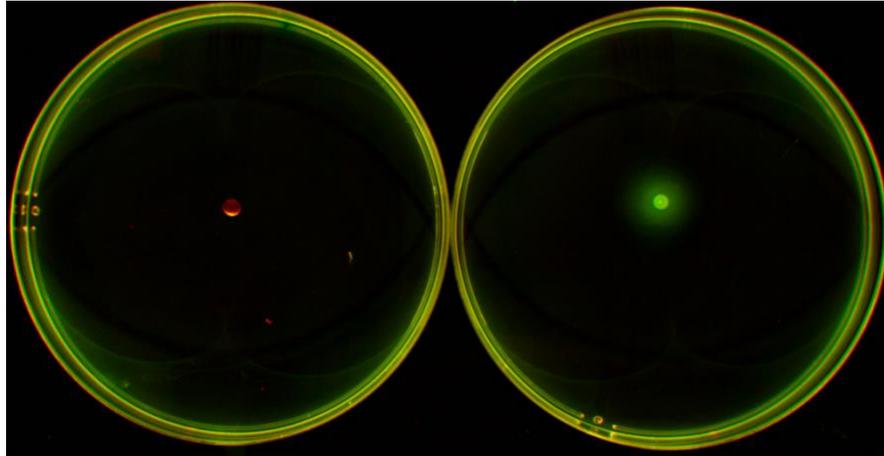
Figure 1.12: QSC (left) and SLC (right) colonies plated on LB agar with 3% acetate. The QSC colony is a very sparse, non-motile colony, indicating that acetate inhibits growth. The SLC colony is slightly denser and grows more than it usually would on a glucose plate.

Unfortunately, the acetate growth experiment only proved that it inhibits growth for the QSC colony, leaving the DLG patterning during overflow metabolism still a mystery. Though more experiments certainly need to be performed to investigate this mechanism thoroughly, it can be concluded that both the excess of glucose and acetate work together to cause DLG type patterning in motile WT and QSC colonies.

1.8 Flagellar development of *E. coli*

Another research paper suggested that the flagellar development of commensal, non-pathogenic, *E. coli* is dependent on cAMP levels[14]. We also know that glucose and cAMP levels are correlated, with elevated glucose corresponding to a drop on cAMP[15]. The data from this paper suggests that the branching pattern is seen as a result of incomplete flagellar development.

Below, Figure 1.13 displays the exact same pattern seen in our motile MG1655 and QSC strains due to a deletion in the gene *fliC*, which is related to flagellar component creation. The patterns match what is seen in Figure 1.7, with a normal radial spread without fluctuations for the WT and the random branched pattern with the deletion of *fliC*.
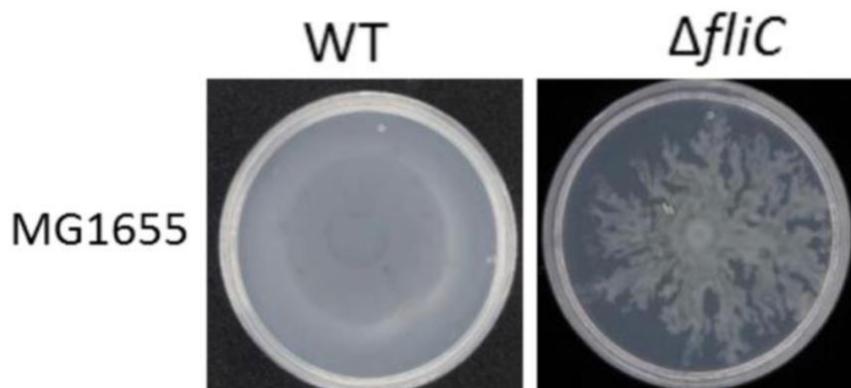
Figure 1.13: MG1655 (both WT and with a *fliC* deletion) plated on agar[14]**.** The *fliC* deletion leads to very similar patterns as the glucose supplied strains, suggesting that flagellar development may be a contributing factor. Figure adapted from Ambagaspitiye et al.

Additionally, research suggests that flagellar development of commensal *E. coli* changes with glucose addition. MG1655 is categorized as commensal rather than uropathogenic, meaning that it is nonpathogenic and does not cause disease[16]. Multiple sources cite that flagella development is inhibited by glucose addition, due to decreased cAMP levels repressing *fliC* expression[17,18]. Figure 1.14 below shows flagellar development while swimming and for surface motility for commensal and uropathogenic bacteria below. Clearly, the development is impacted, and this stunted growth must be the reason for DLG type patterning to arise. Since the chemotaxis system is also inhibited this may be a key factor to study for other behavior shown by the various colonies on glucose plates. From these studies, we can conclude that perhaps the elevated glucose levels impact flagellar development in a similar fashion to the deletion of the *fliC* gene, which leads to changes in the patterns formed by the strain. This conclusion is

supported by research that found *fliC* expression to decrease in glucose and sucrose

concentrations above 80 mM and 40 mM, respectively[19]. This is only observed in WT and QSC

strains while the SLC strain remains completely non-motile, indicating that another mechanism

is taking place in that circuit to prevent motility.
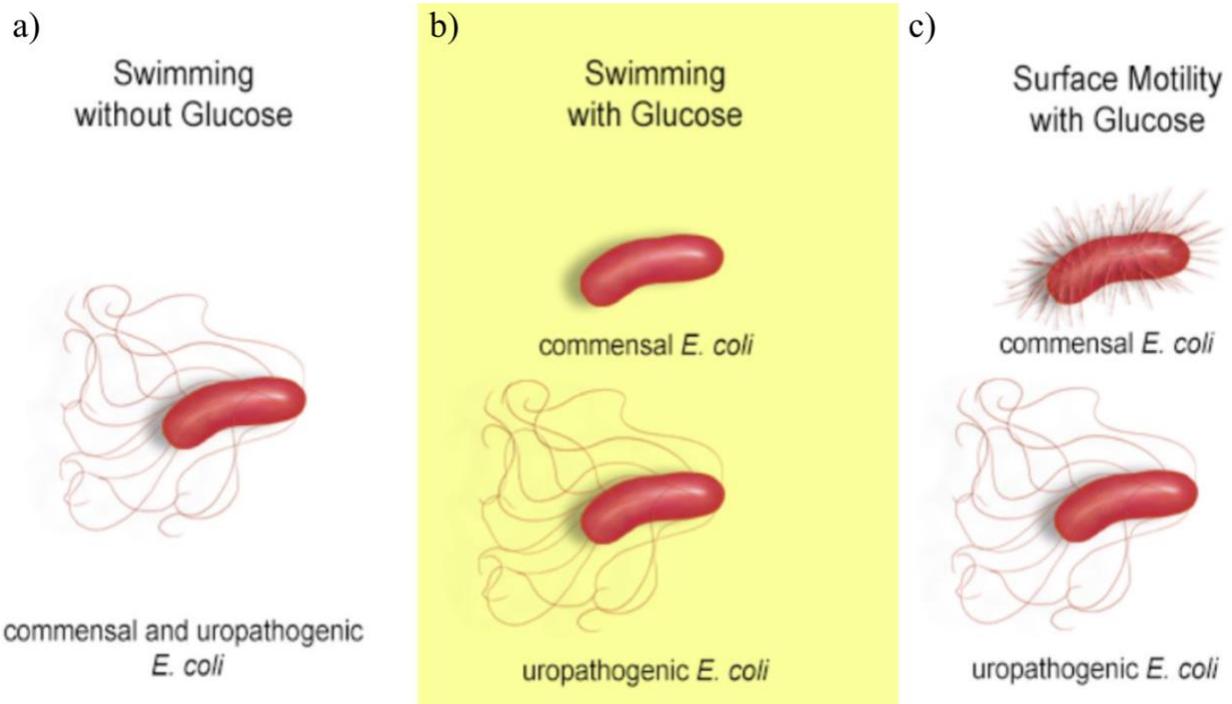
Figure 1.14: Development of appendages of commensal and uropathogenic *E. coli* with and without glucose[14]. Commensal *E. coli* lose flagella while swimming and have shorter fragments during surface motility. Figure adapted from Ambagaspitiye et al.

We can conclude from the research that the MG1655, both WT and QSC, must resemble

the commensal bacteria in Figure 1.14c, where rather than flagellar development, they only have

shorter structures called fimbriae. These short pili-like structures must lead the bacteria to form

DLG-like patterns, perhaps in combination with the secreted acetate inducing a negative

chemotaxis gradient and general stress. For the SLC strain, which does not make similar DLG

patterns, the lysis circuit must prevent motility all together and keep the colony dense and

perhaps lyse immediately. Additionally, another research paper by the same group that found this phenomenon claimed the patterns arose from type I pilus-mediated motility[20]. As seen in Figure 1.15, MG1655 with the *fliC* gene deleted resembles the same pattern as our QSC and WT strains in glucose. The paper claims that the motility is fimbriae-driven, as the flagella do not form at all. Looking more in depth at the 7740 strain used in the previous experiments, it was noted that there was an IS1 element in the regulatory region of the *flhDC* operon, which drives transcription of class II flagellar genes, that causes hypermotility of the strain[21]. The 7740 strain does not contain any pre-existing *fliC* deletion, but it is likely that the overexpression of the *flhDC* operon, which has downstream effects on *fliC*, results in repression of the gene with elevated glucose levels. With this repression, the 7740 strain behaves like it has a *fliC* deletion.
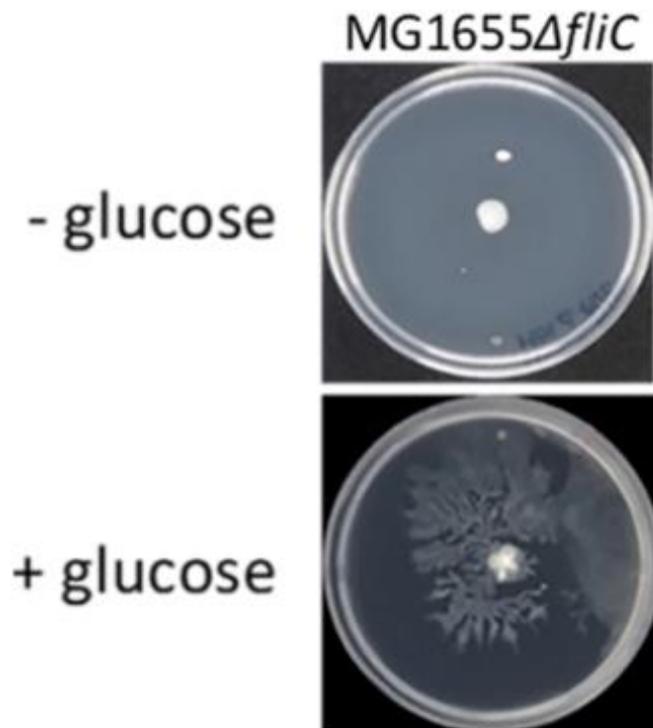
Figure 1.15: Motility assay for MG1655 with *fliC* deletion, on non-glucose and glucose plates[20]**.** With glucose, MG1655 with a *fliC* deletion undergoes type I pilus-mediated surface motility. Figure adapted from Sudarshan et al.

Type I fimbriae-driven surface motility is different from flagellar motility. While flagella typically rotate clockwise or counter-clockwise to drive bacteria along a chemical gradient, fimbriae engage in a type of motility called twitching[14,22]. It essentially works like a grappling hook, adhering to the surface and pulling the bacteria along, seemingly randomly. The patterns resemble DLG aggregates, suggesting that perhaps the pili-driven motility is highly dependent on a chemotaxis gradient.

1.9 Investigating Hypersensitivity to Cell Lysate Under Stress

With the previous processes studied from literature review and experiments, it was concluded that the patterns seen in the WT and QSC colonies resulted from a combination of the stress induced from overflow metabolism, incomplete flagellar development from elevated glucose levels, and the DLG growth pattern that follows. The primary cause is likely the loss of the flagella due to a high glucose concentration which regulates cAMP levels and therefore decreases *fliC* expression, as discovered in previous literature[17–19]. This results in primarily fimbriae-driven surface motility, or twitching, which closely resembles patterns seen in DLG. It still needed to be determined what exactly about the SLC colony was causing the other colony to avoid it in high glucose concentration conditions. Since the SLC I transformed into the chassis contained both hemolysin E (*hlyE*) as the therapeutic and the E lysis protein that caused cell lysis, I theorized that the motile cell colony was sensing these "dangerous" proteins as a stress response and then avoiding it. The hypothesis was that the hlyE and E lysis protein were secreted into the surrounding agar after the SLC colony lysed, and the QSC colony, in its stressed state, became hypersensitive to the cell lysate.

To test if hypersensitivity to cell lysate was a possibility, I had to artificially create bacterial cell lysate from both WT bacteria and SLC bacteria. In order to preserve the proteins of interest from the lysate, a cold shock method of lysing was used. The SLC strain was grown in 5

mL of LB overnight and then spun down into a pellet. Then, the tube was placed for 2 minutes in a dry ice and ethanol bath, and then transferred to an ice bath for 8 minutes. This was done 3 times, as prescribed in the methods used in previous literature to obtain cell lysate without accidentally denaturing the proteins of interest[23]. Once the cells were thoroughly lysed, the pellet was resuspended in phosphate buffered saline (PBS), so the pH was stabilized and did not affect protein conformation. Half of the lysate was filtered since theoretically, the protein size was smaller than 0.22 microns, and any cells still alive would be removed. The other half was kept as unfiltered lysate just in case there was something causing the colony avoidance that was larger than the filter pore size.

For this experiment, the QSC strain was spotted in the center of an LB agar plate with 1% glucose. The lysate was spotted in three dots surrounding the QSC colony. The filtered and unfiltered SLC lysate were both tested, along with 10 nM of AHL in PBS buffer just to see if the quorum sensing played a role.
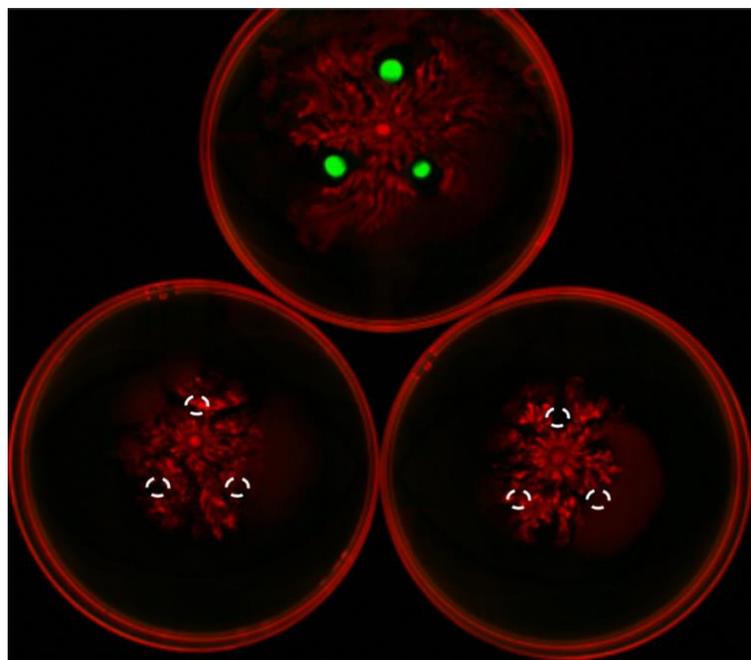
Figure 1.16: Testing if QSC colony avoids lysate and AHL. Top plate shows unfiltered SLC lysate spotted around the central QSC colony, which resulted in growth from some remaining live cells. Bottom left plate is filtered SLC lysate and bottom right plate is 10 nM AHL diluted in PBS, both of which did not appear to be sensed by the QSC colony (spots outlined in white).

Figure 1.16 shows the results above. Surprisingly, the filtered lysate and AHL had no impact on the spatial avoidance of the QSC colony, and the strain grew right over the spots. The unfiltered lysate most likely had some living cells remaining as growth was seen, as shown with the GFP fluorescence, and spatial avoidance occurred with that colony. From these results we can conclude that the cell lysate and AHL was most likely not causing the avoidance and it had to do with the living colony, or something larger than the pore size of the filter: 0.22 microns.

With these results, *hlyE*, E lysis protein, and AHL were ruled out as reasons causing avoidance of the SLC colony. Next, the chemotaxis system of *E. coli* was investigated to see it that was a factor.

1.10 Chemotaxis of Bacteria Undergoing Type I Fimbriae Driven Surface Motility

$E.$ $coli$ is able to sense concentration gradients of various attractants and repellants, and is able to swim up or down them repectively[22]. It typically does so with rotation of its flagella, counterclockwise to propel it in the direction of attractants, and clockwise to cause tumbling and reverse the direction of movement. It also uses a response regulator, $cheY$, that favors tumbling and clockwise rotation when it binds to the flagella after phosphorylation.

However, we already established that on glucose plates, flagella development is stunted, and the primary movement results from twitching from fimbriae. Pilus driven movement is extensively studied in $Pseudomonas$ $aeruginosa$, and it typically involves the use of type IV fimbriae for biofilm formation, which can sense nutrient rich areas[24–26]. It is established from previous research that only type I pili are involved in motility for MG1655, which is a commensal type of $E.$ $coli$[14]. Type I fimbriae are typically associated with adhesion to glycoproteins or glycolipids that have mannose[27]. Agar is composed of galactose in various forms, so fimbriae should not bind that tightly to the surface of the agar. It has been proposed that $fimH$, which is the lectin located at the end of the type I fimbriae, can weakly bind to media not containing mannose[28,29]. It is likely that the basis of the motility is dependent on biomechanical properties of the fimbriae components, such as flexibility[30].

The conclusion from this must be that fimbriae-mediated surface motility requires conformational changes in the fimbriae. Additionally, intracellular factors like overflow metabolism, must also affect the conformation of the fimbriae, potentially driving chemotaxis in a different way. One element to investigate further was whether the dense SLC colony was acting as a nutrient sink to discourage the other colony from entering that nutrient sparse area. To test this, a non-motile WT MG1655 was plated next to the motile MG1655 7740 from the CGSC to see the behavior.
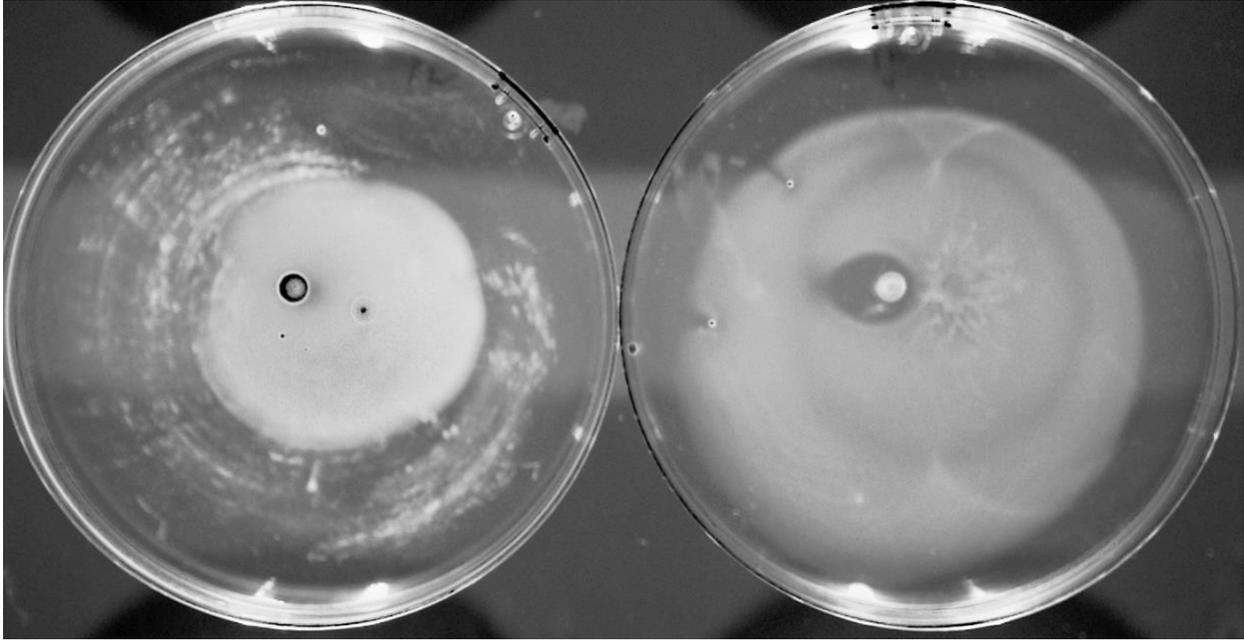
Figure 1.17: Non-motile WT and motile WT MG1655 plated next to each other on regular 0.35% LB agar (left) and with 1% glucose (right). Patterns resemble Figure 1.8 where without glucose the colonies grow into each other whereas with glucose the DLG pattern arises in the motile WT and colony avoidance occurs.

As seen in Figure 1.17, the non-motile colony also causes avoidance from the motile WT colony on glucose. This shows that the avoidance was not limited to SLC behavior alone and it is due to an inherent mechanism in the chassis. Likely, what is happening, is that the motile WT colony is undergoing biofilm formation via type I fimbriae and avoids areas with a lower source of nutrients, which the dense non-motile colony has started to use up. Similar to how DLG takes place, it is likely that the twitching mechanism of the pili combined with the fimbriae elements like *fimH* rely on surrounding nutrient concentration to determine spatial growth patterns. Additionally, since conformational changes are required, the overflow metabolism producing acetate probably allows for such changes.

This duality of a more motile strain and a less motile strain reminded me of a paper produced in the lab years ago where *E. coli* and *A. baylyi* were plated together to result in beautiful flower shaped colonies[7]. We had a similar situation with the QSC and SLC strains, with

the added component of enabled crosstalk between the colonies due to AHL diffusion. I decided to recreate the experiments in that paper with my strains to see if interesting patterns could arise on glucose plates.

1.11 Recreating Flower Shaped Patterns

The paper I referenced in the previous section involved a spatial patterning project that found an interesting phenomenon in bacterial growth and created a mathematical model. Here, *A. baylyi* is known to be an incredibly motile species while *E. coli* is less motile. When the 2 strains are grown to exponential phase and plated in a 10:1 ratio of the motile strain being at a higher proportion, flower like patterns emerge as a result, seen in Figure 1.18.

Figure 1.18: Flower like patterns emerge from a 10:1 ratio of *A. baylyi* and *E. coli* grown for three days on a 0.5% LB agar plate[7]**.** Figure adapted from Xiong et al.

The above pattern arises from the highly motile strain pushing out the less motile one, causing cusps and branches rather than a smooth radial outward pattern. It arises from the *E. coli* destabilizing the colony front by hindering *A. baylyi* expansion. I decided to recreate the experiment with my strains, instead growing them on 0.35% LB agar plates with 1% glucose to force the fimbriae mediated motility. Figure 1.19 details the results below.
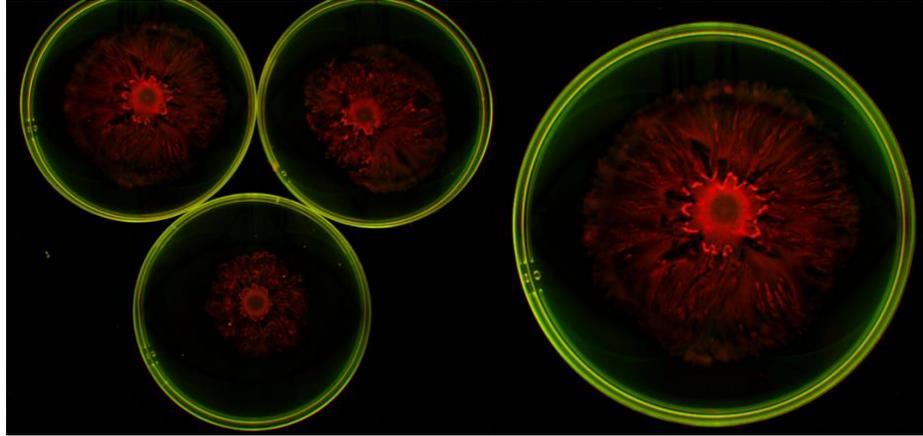
Figure 1.19: Flower pattern emerges from 10:1 Ratio of QSC and SLC strain on 1% glucose and 0.35% LB agar plate. Left shows the triplicate run all results in similar patterns while the right shows a zoom in of one of the replicates.

Though the flower pattern is not as brilliant or striking, the pattern seen above sees some similarities to the previous paper. Around the edge of the colony before the strain appears to break out, there are cusps or curved dense lines resembling a colony front that broke. Additionally, there are areas of less opaque colony presence that could signify some lysing from the SLC occurring. Comparing this to the individual behavior of the two colonies from Figure 1.7 and 1.8, it is clear that the emerging colony fronts are a result of the combination.

1.12 Modeling the Phenomena

With the help of Dr. Lev Tsimring, I developed a model of the fast-moving motile colony avoiding the non-motile colony. The behavior is very similar to a Keller-Segel model, which is the first model of chemotaxis developed around 1970, and this version is essentially a 2 dimensional version[31].

We introduce two cell density fields, $n_1(x, y, t)$ and $n_2(x, y, t)$, as well as the nutrient concentration field $c(x, y, t)$. The first density field represents motile bacteria that are also chemotactic to the nutrient concentration. The second density field represents the non-motile strain of bacteria. Both cell fields are growing logistically with rate λ, which is assumed to be

constant, until the total cell density reaches the media carrying capacity density n_{max}, which is described by the growth function $f(n_{1,2}) = \lambda n_{1,2}(1 - (n_1 + n_2)/n_{max})$. Both cells are consuming nutrients with rate k, the there is a sink term for nutrients $s = -k(n_1 + n_2)$. The motile cells diffuse with diffusion constant D_n, while nutrient diffuses with the diffusion constant D_c. The chemotaxis of the motile strain is described by the term $\nabla \cdot (n\nabla c/(c + c_0))$. The full model reads as follows:

$$\frac{\partial n_1}{\partial t} = D_n \nabla^2 n_1 - \chi \nabla \cdot \left[\frac{n_1}{c+c_0} \nabla c\right] + \lambda n_1 (1 - \frac{n_{max}}{n_1+n_2}) \tag{1}$$

$$\frac{\partial n_2}{\partial t} = \lambda n_2 (1 - \frac{n_{max}}{n_1+n_2}) \tag{2}$$

$$\frac{\partial c}{\partial t} = D_c \nabla^2 c - k(n_1 + n_2) \tag{3}$$

In (x, y) coordinates explicitly,

$$\frac{\partial n_1}{\partial t} = D_n \left(\frac{\partial^2 n_1}{\partial x^2} + \frac{\partial^2 n_1}{\partial y^2}\right) - \chi \frac{\partial}{\partial x}\left[\frac{n_1}{c+c_0}\frac{\partial c}{\partial x}\right] - \chi \frac{\partial}{\partial y}\left[\frac{n_1}{c+c_0}\frac{\partial c}{\partial y}\right] + \lambda n_1(1 - \frac{n_{max}}{n_1+n_2}) \tag{4}$$

$$\frac{\partial n_2}{\partial t} = \lambda n_2 (1 - \frac{n_{max}}{n_1+n_2}) \tag{5}$$

$$\frac{\partial c}{\partial t} = D_c \left(\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right) - k(n_1 + n_2) \tag{6}$$

For initial conditions, we assume that $c(x, y, 0) = C_0 = const$ and $n_{1,2}$ are non-zero only in small, separated, localized regions. We use split-step method with the 1st order explicit integrator for the reaction terms and the Fourier method for diffusion and chemotactic differential operators. The model was built in MATLAB and produced the results seen below.
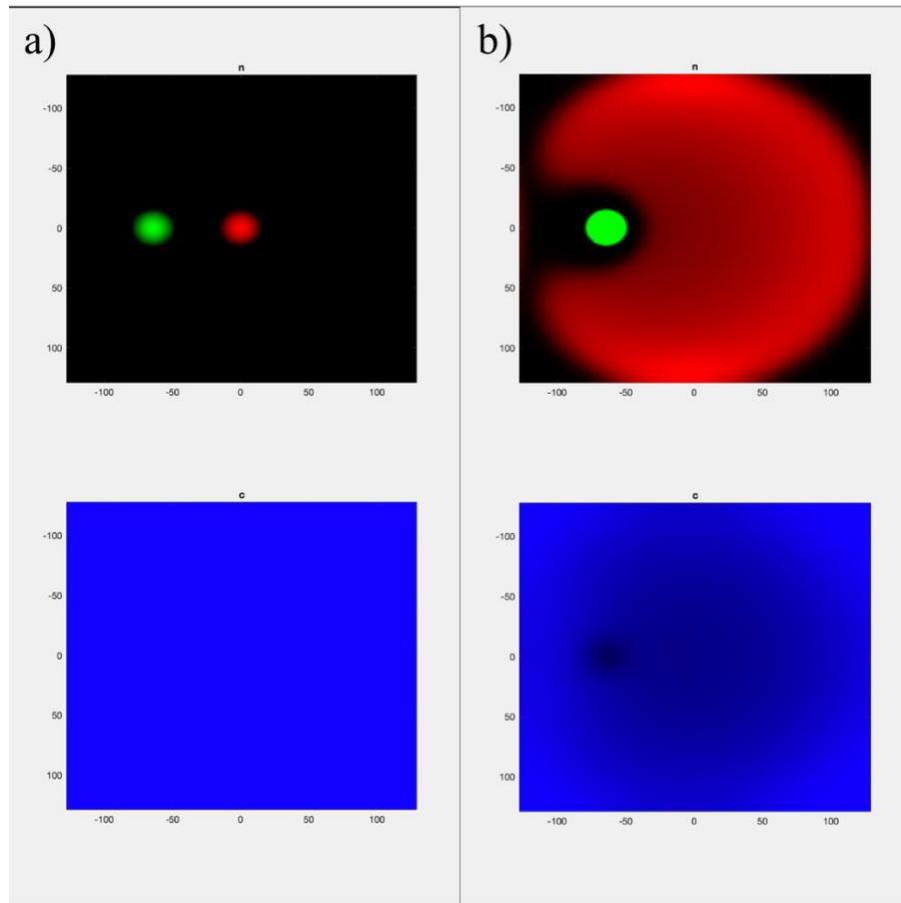
Figure 1.20: Keller-Segel model results for chemotactic motile strain. a) At t=0 the colonies are both small. b) After some time, the red colony, which is faster and more motile, spreads quickly and shows avoidance of the smaller non-motile green colony.

The model results in Figure 1.20 depicts the relationship seen in our *in vitro* studies. The colony avoidance evidently occurs, though the twitching motility is a little more difficult to model. Most research suggests twitching is entirely random and results in a random-walk growth pattern, which I do not believe is the case. My studies have shown that there is an element of nutrient and surrounding factors that influence the motility of the bacteria. As such, a model that combines the supposed random walk of twitching motility as well as chemotaxis needs to be developed. Unfortunately, I ran out of time for this project, but a model showing the dual-strain

growth pattern is also necessary and can perhaps be based on the previous multi-strain growth model developed in the lab.

1.13 Discussion and Future Directions

From the experiments conducted so far, a few conclusions were made. Firstly, no interesting patterns of note arise from the SLC and QSC colonies on normal agar plates. With glucose, the QSC colony starts growing in random, chaotic patterns resembling DLG. The SLC colony completely loses its motility, likely due to the excess stress from overflow metabolism and the heavy burden of the plasmid circuit. The QSC colony, and motile WT MG1655, both lose flagellar development on glucose plates and the movement becomes driven by pilus structures via twitching. Additionally, the more motile colony prioritizes regions with higher nutrients, so the less motile colony acts as a resource sink which makes the motile colony move around it and avoid growing into the area. This is likely a biofilm survival mechanism that avoids survival of the fittest situations where resources will have to be competed for.

Though the original purpose of this project was SLC characterization, it extended to a study of multi-species colony dynamics. On non-glucose environments, the SLC behaves predictably, growing as a normal dense colony. Instead, the surprising discovery to come out of this was the fact that glucose, which originally was only used in the lab to slow lysis, also affects the flagellar development pathway with the same cAMP molecule that inhibits quorum sensing. This ends up forcing a switch to twitching motility in WT bacteria. However, in the SLC strain, even twitching motility is not observed and it appears that the stress from the acetate and lysis circuit prevents motility altogether. It is also likely that the density of the colony causes such a high AHL concentration that lysis occurs at the border and the strain cannot break out.

From here, I decided to branch into multi-species growth dynamics by recreating the flower shapes using a non-motile and motile strain. The added element of AHL cross-

communication could introduce a varying element that affects the growth in some way. Figure 1.18 shows that interesting patterns arise from the combination of two strains. The GFP expression overlays entirely with the RFP expression suggesting a complete mixture of the two strains. In the future, I would suggest creating a different reporter molecule for the SLC and removing the GFP expression in the QSC strain or switching it out for another color. This would enable easier distinguishing of the QSC from the SLC strain and we can visualize where the SLC strain ends up in the spanning colony.

If I had a little more time on this project, I would have liked to make a time-lapse of the multi-strain growth. In addition to visualizing the growth in real-time, which undoubtedly would help with modeling and seeing potential switches take place, it would also provide a visual that helps emphasize the pattern. The lab had a device we termed the Milliscope, which involved a box containing a fan and humidifier that could control temperature and humidity via an Arduino. There was also a camera with optic filters capable of capturing GFP and RFP channels, which I could control via Micromanager. Unfortunately, I was not able to get the system to work properly, but in the future, I hope that the system can be used to create a time-lapse. Visualizing the colony fronts forming in real-time as well as the chemotactic nature of the WT next to the SLC would be incredibly useful in characterizing the behavior with models, and ensuring the fit is accurate to the time-lapse.

Additionally, models that describe this new growth pattern can most likely be made using the base of the existing model that shows colony front destabilization, from the original flower pattern paper. An additional term for AHL levels and a kill coefficient for the SLC strain will need to be introduced, and then studied to see if the model runs correspond to the images seen above.

Though the initial results of the project did not seem promising for crucial SLC characterization in a spatiotemporal growth sense, the findings from the dual strain growth experiment appear promising as an add-on to the flower pattern study. Introducing the element of lysis and crosstalk via quorum sensing signals proves that the pattern appears to change slightly. Additionally, there is the switch from flagella-mediated movement to fimbriae-mediated movement on glucose plates for *E. coli*. *A. baylyi* already engages in twitching motility which allows it to spread further and faster[7]. These phenomena combined lead to the patterns seen above, which need to be modeled and validated. Doing so will confirm that we understand the mechanisms taking place and allow for further characterization of the SLC colony when placed with other WT strains that engage in twitching motility. It could lead to a less localized therapeutic and allow for faster movement along the tumor where treatment is intended, covering a higher surface area for a widespread zone of treatment. Though further studies are definitely required in this area, I believe these experiments provide some insight and a definitive starting point.

Chapter 2 - IMPROVEMENT AND CHARACTERIZATION OF THE CANCER SENSOR

2.1 HGT in *A. baylyi*

Acinetobacter baylyi is a highly competent bacteria that is well studied in the field of

biology, and is relatively non-pathogenic in humans[32–34]. One interesting characteristic of the

bacteria is that it can engage in a mechanism referred to as horizontal gene transfer (HGT)[3].

HGT describes the exchange of genetic material between organisms outside the typical "vertical"

transfer from parent to offspring[35]. HGT is known to be an adaptation method for many bacteria

and archaea, which enables the evolution of the species by uptake of surrounding genetic

material which could potentially allow for better survival. Through this process, the bacteria can

easily uptake extraneous DNA and incorporate it into their genome. This mechanism, combined

with the fact that bacteria can easily find their way into the gut or gastrointestinal tract, allows

for the possibility of a bacterial "sensor" of DNA of interest, for which the output could easily be

measured by examining stool or urine samples[36,37]. It is well known that bacterial diagnostic

tools exist already, with some examples being microbes that react to gut inflammation or

intestinal bleeding[36,38]. So, it can easily follow that a bacterial diagnostic tool can be designed to

sense surrounding DNA, employing the native HGT mechanism. Specifically, mutated

oncogenes like KRAS, which is known to be an important factor in CRC development, can be

detected by placing the sensor next to affected cells[39]. Using a CRISPR mechanism to cut up WT

KRAS DNA and then allowing HGT to pick up the mutated KRAS signaling oncogenic activity,

a successful tumor-sensing diagnostic can be designed.

The key takeaway used in the construction of an *A. baylyi* "sensor" was the original

intake of the DNA from either the extra-cellular matrix (ECM), or even from the released

contents after *A. baylyi* kills the neighboring cells. From here, the sensor needs to have homology

arms embedded in the genome, in order to recognize and bind to parts of the extraneous DNA that needs to be detected, and then use homologous recombination to integrate the foreign DNA strand. A cascade event will then be triggered from the resulting "knockout" of the genome and lead to a positive signal in the bacteria indicating successful integration of the extraneous DNA. This design is exactly what Dr. Rob Cooper used in his construction of the CRC sensor, described in the next section.

2.2 *A. baylyi* as a CRC Sensor

As alluded to above, the HGT mechanism of *A. baylyi* was tapped into to create a sensor for oncogenes specific to CRC. The specific technology is referred to as CATCH, a CRISPR-discriminated horizontal gene transfer strategy, which involves the genomic integration of target DNA, and the subsequent output signal that notifies successful integration. The workflow is detailed in Figure 2.1 below, adapted from Dr. Cooper's published paper on the technology[3].
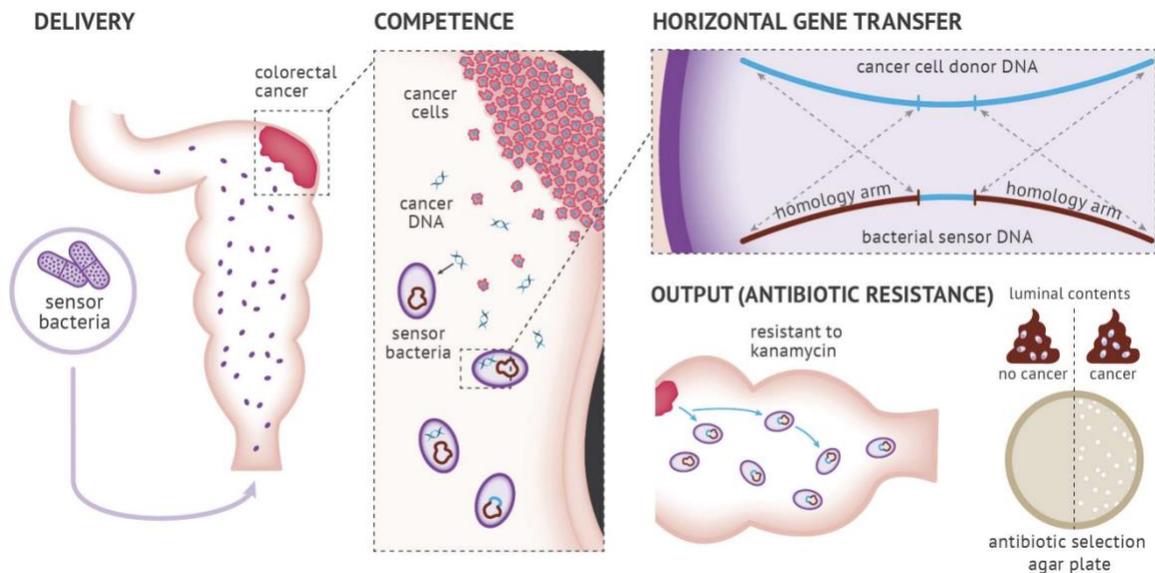
Figure 2.1: Workflow describing the CATCH system[3]. Sensor bacteria containing HGT capabilities, homology arms flanking a recognition sequence, and CRISPR gRNA that cuts up sequences that do not have the specific oncogenic mutation are delivered to the gut to sense CRC. They uptake the DNA and integrate it into their own genome, producing an output of kanamycin resistance. Luminal contents containing the bacteria are streaked onto antibiotic selection plates and if colonies appear, the DNA was sensed. Figure adapted from Cooper et al.

As described in Figure 2.1, engineered *A. baylyi* are to be administered rectally, utilizing their natural competence to uptake tumor DNA shed into the colorectal lumen from the cancerous cells. This experiment was conducted in mice hosting engineered tumor cells. In this simple experiment, the tumor donor DNA was strategically engineered with a kanamycin resistance (*kanR*) cassette flanked by KRAS homology arms. Concurrently, sensor bacteria are engineered with complementary KRAS homology arms to facilitate homologous recombination. Through horizontal gene transfer from tumor DNA, sensor bacteria acquire kanamycin resistance, allowing for their quantification from luminal contents via serial dilution on antibiotic selection plates. Tests were conducted with a 2 kb *kanR* and GFP insert that swapped out a spectinomycin resistance (*specR*) gene that was already present in the sensor DNA genome. Another construct involved an 8 base pair insert to repair a defective *kanR* gene.
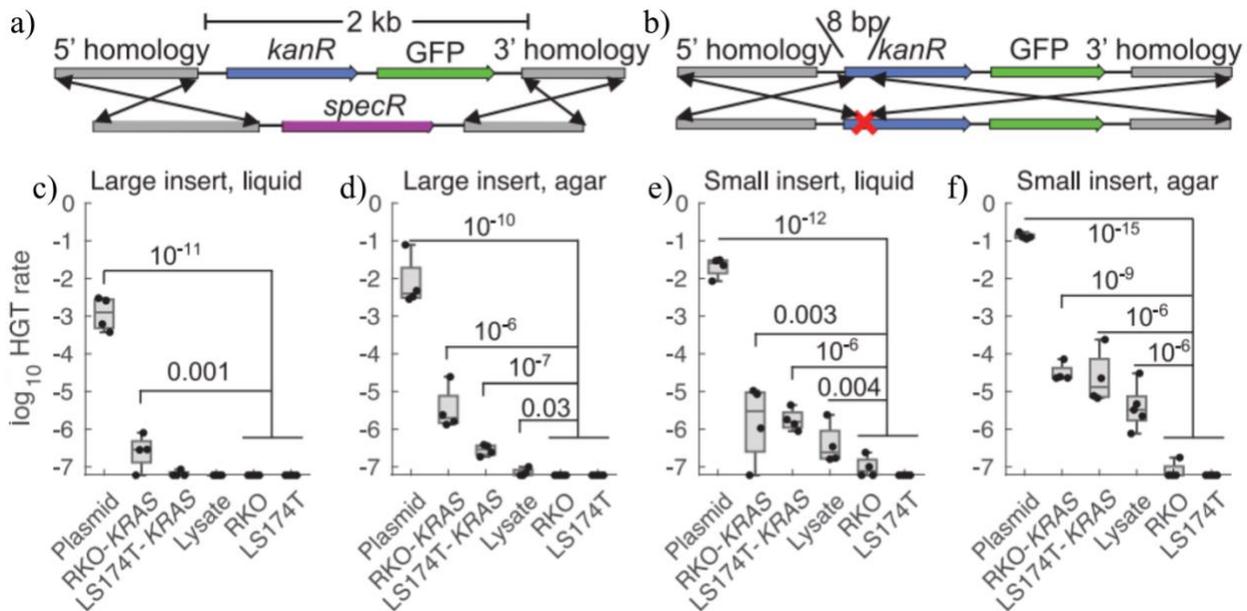
Figure 2.2: Constructs sensing KRASG12D DNA *in vitro*[3]. Donor DNA was derived from plasmid PCRs, purified cancer cell genomic DNA, or raw lysate that were incubated with the biosensor. HGT included either a large, 2-kb insert (a) or a small, 8-bp insert to repair two stop codons (b), in both cases providing kanamycin resistance. c-f) Biosensors were incubated with plasmid DNA, purified RKO-KRAS or LS174T-KRAS genomic DNA, or raw RKO-KRAS lysate, all containing the donor cassette, or purified RKO or LS174T genomic DNA as controls. Figure adapted from Cooper et al.

From Figure 2.2, it is apparent that the constructs are able to sense DNA coming from the plasmid source, regardless of the insert size or incubation medium. RKO and LS174T lysate containing KRAS DNA are also able to be sensed, though at a lower HGT rate. Raw lysate from RKO is able to be sensed on agar mediums, but there is a false positive background on the agar as well, seen in Figure 2.2e-f where a small level of RKO control still yielded colonies with *kanR*. Though there is an issue of false positives, for the most part, this simple construct can sense engineered tumor DNA, in plasmid DNA, genomic extract, and raw lysate.

Dr. Cooper further engineered the sensor bacteria to detect natural, nonengineered tumor DNA sequences. It was able to discriminate oncogenic mutations at the single-base pair level, specifically, the KRAS G12D mutation that is associated with CRC presence. The simplified construct is seen in Figure 2.3. KRAS is a gene present in all wild type somatic cells. When it obtains a mutation, such as the G12D mutation, the cell experiences oncogenic activity[39]. In clinical settings, the issue lies in ensuring the sensor DNA does not accidentally yield false positive reads when it intakes WT KRAS DNA. *A. baylyi* contains an endogenous type I-F CRISPR-CAS system called Cascade, which comes into use for the posed problem[40]. The protospacer adjacent motif (PAM) of *A. baylyi* was used to insert a CRISPR spacer that targeted the WT KRAS DNA, which would then degrade any WT KRAS that entered the cell. The spacer was embedded into another neutral genome location. This enabled only KRAS G12D to be sensed by the bacteria, and act as a true cancer sensor. This construct utilized the *P_LtetO-1* promoter driving the output gene, and suppressed by the *tetR* gene[41]. The *tetR* gene was also flanked by KRAS homology arms intended to be knocked out by the tumor DNA of interest. This would subsequently stop repression of the output gene and provide a positive signal.
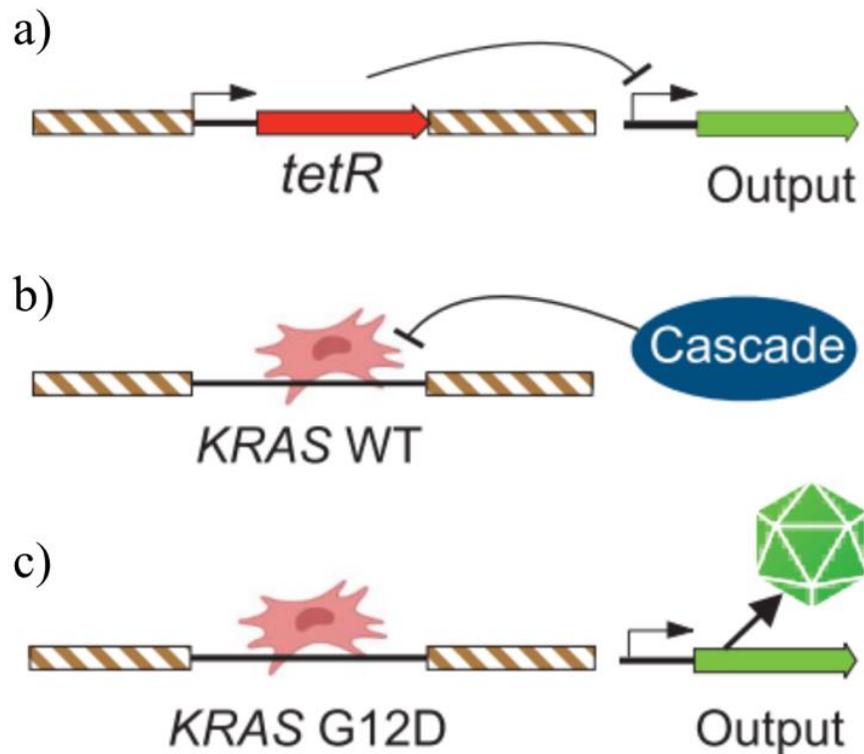
Figure 2.3: Construct detection of non-engineering tumor DNA[3]**.** a) *tetR* located between the homology arms on the *A. baylyi* genome represses expression of the output gene. b) Target DNA with wildtype KRAS sequence is recognized and degraded by the type I-F CRISPR-Cas effector complex, Cascade. c) Target DNA with the KRASG12D mutation avoids degradation, replaces *tetR* in the biosensor genome, and relieves repression of the output gene. Figure adapted from Cooper et al.

2.2.1 Issues with the Current Design

Currently, the cancer sensing system in *A. baylyi* uses the tetracycline repression (*tetR*) system to suppress either Kanamycin resistance or GFP expression. As described, the mechanism involves the sensing of mutated KRAS G12D DNA. HGT occurs and the homology arms flanking the tetR sequence in the genome match up with the KRAS sequence in the foreign DNA and kick out the tetR gene using homologous recombination. One major issue with the use of a *tet* repression system in ADP1 is toxicity from *tet* itself. It was somewhat difficult to transform the tetR-producing gene into the genome of ADP1 as any successful transformation ended up dying very quickly, suggesting the *tet* may have been repressing vital functions necessary for

37

survival within the cell. The only successful transformation ended up being a mutated version of the *tet* system that was temperature sensitive, allowing repression to occur only at 30 °C.

In addition to the inherent toxicity of the system that interfered with testing and general strain survival, there were issues of a high mutation rate in the genome and linear DNA once integrated into the bacteria, which caused false positives to occur. The *tetR* system or the KRAS gene sometimes mutated in the negative controls and yielded positive results when none should have occurred. We believe that the toxicity of the *tetR* system is related to this issue, as the higher stressed state the ADP1 is placed in results in a higher mutation rate in an attempt to survive the stress.

Lastly, there was an issue regarding competence in the *A. baylyi* due to a filamentous phage that reduces it. In ADP1, there was found to be a filamentous phage existing in the genome named the competence-reducing Acinetobacter phage (CRAΦ)[42]. Reactivation and evolution of CRAΦ can occur in stressful situations for the ADP1 bacteria, which reduces the potential for HGT to occur.

2.3 Proposed Solutions

With the current issues with the design posed above, improvement of the biosensor needed to occur from three avenues. We had to deal with 1) the inherent toxicity of the tetR system, 2) the high mutation rate that caused false positive reads, which may also be related to the *tetR* toxicity, and 3) the existing CRAΦ prophage that emerges when the bacteria are highly stressed, reducing competency.

To solve the first issue, this project will attempt to swap out the *tetR* system for another repression system, *lacI*. *LacI* is a lactose repressor protein derived from the *lac* operon from *E. coli*[43]. The *lacI* output will be suppressing the output gene driven by the *pTrc* promoter, which is a combination of the trp and lac promoter, able to be suppressed by *lacI*[44]. Additionally, *lacI*

suppression has been successfully transformed in ADP1 before, so this swap should pose a simple synthetic biology problem[45].

To solve the issue of high mutation rates, we hope that the *lacI* system will produce less stress on the bacteria and lead to a higher stability of the system. Additionally, there have been recorded issues of transposable insertion sequence (IS) elements in ADP1 which can inactivate genes[46]. The group that identified these IS elements managed to delete all 6 copies of IS1236 from the genome of ADP1, seeing a 7- to 21-fold reduction in the mutation rate of a reporter gene. This strain with the deletion is referred to as ISx.

Lastly, the competency of the *A. baylyi* was improved with the deletion of the CRAΦ prophage[40]. Deleting the phage resulted in an increase in triple transform efficiency, which should hopefully also correlate to a higher competency with respect to the non-engineered KRAS sequence. This strain is referred to as ΔCRA. Much of the biosensor's activity is limited by whether the DNA is actually taken in, so deleting the phage that prevents uptake should certainly improve the sensing capabilities.

With the ISx and ΔCRA strains established, all that remained was the subsequent characterization of the sensor with these improvements, as well as further characterization after switching the *tetR* system for *lacI*. Detailed in the next few chapters are the results of this project.

2.4 Materials and Methods

Firstly, the DNA constructs to be switched out needed to be developed. This was not a typical plasmid transformation into a strain of bacteria, as we needed to ensure only one copy remained in each individual bacterium. As such, the circuit was embedded into the genome itself using homologous recombination. The existing circuit involved two regions on the ADP1 genome, both of which were confirmed previously to be archaic remnants full of nonsense DNA.

These two regions were dubbed neutral 1 (N1) and neutral 2 (N2). The N1 region contained the repressor gene flanked by the KRAS homology arms. This part is what engaged in HGT and knocked out the repressor. The N2 region hosted the output, with an antibiotic resistance gene driven by the promoter that was respectively suppressed by the repressor.

For the previous design, the N1 region had *tetR* flanked by KRAS while the N2 region had *pTet* driving KanR. So, to replace the system with *lac*-based repression, the *tetR* needed to be swapped for *lacI* and the *pTet* promoter needed to be swapped for *pTrc*, which was repressed by *lacI*. Swapping out these individual DNA parts would prove to be difficult since there is no way to confirm which colonies hosted the corrected system visually. Instead, I decided to work from the ground up, starting with a simple ADP1 construct with only GFP in the N1 region. This strain was labeled ADP1 N1 GFP. From here, the idea would be to design two constructs. The first was the *kanR* gene driven by *pTrc*, flanked by N2 homology arms to be directly inserted into the N2 region (N2-pTrc-KanR). The second construct would be N1 homology arms flanking KRAS homology arms that flanked the *lacI* gene, driven by a constitutive promoter (N1-KRAS-LacI). The *lacI* and *pTrc* sequences were PCR'd from the pBWB294 plasmid (#140636, Addgene)[47]. The N1-KRAS and N2 homology arms as well as the *kanR* sequence were obtained from PCR off of existing DNA constructs used to build the first iteration of the biosensor, labeled Ntrl1_KRAS_space and Ntrl2_PLtet_UTR1. PCR primers are detailed in Table 2.1 below. For the first construct intended to sit in the N1 region, three separate parts were built. The three parts were the 5' end of the N1 and KRAS homology arm, the *lacI* gene, and the 3' end of the KRAS and N1 homology arm. The second construct had four parts: the 5' end of the N2 homology arm, the *pTrc* promoter, the *kanR* gene, and the 3' end of the N2 homology arm. The primers were designed with overlaps intended for Golden Gate assembly. Using the restriction

enzyme *BsaI*, T4 DNA Ligase, and *DpnI* to chop up the backbone, the protocol from NEB was

followed to build the constructs using PCR pieces. Sequencing was conducted to confirm the

product had the desired sequence.

Table 2.1: PCR primers used to build DNA constructs.

Primer Name	Template Strand	Primer Sequence
oAS_Ntrl1_5'_F	Ntrl1_KRAS_space	CTGACAACCGACTACCGTCC
oAS_KRAS_5'_R_2	Ntrl1_KRAS_space	CCCGGTCTCATGCTGACAGCTATCTCCACT
oAS_LacI_F	pBWB294	CCCGGTCTCAAGCAGCGGCCGCTTCTAGAACTAT
oAS_LacI_R_2	pBWB294	CCCGGTCTCAAGGAACAGCCATACCACAGCTTCC
oAS_KRAS_3'_F_2	Ntrl1_KRAS_space	CCCGGTCTCATCCTTGGGAGTATGTCAGGGT
oAS_Ntrl1_3'_R	Ntrl1_KRAS_space	ATTGCTCGGCACTTCGTTTA
oAS_Ntrl2_5'_F	Ntrl2_PLtet_UTR1	CGTTACCAGCTCCACCAGTT
oAS_T2_R	Ntrl2_PLtet_UTR1	CCCGGTCTCATTCCTCGAAAAAAAAAGCCCGCAC
oAS_pTrc_F	pBWB294	CCCGGTCTCAGGAAGCTGTGGTATGGCTGT
oAS_pTrc_R	pBWB294	CCCGGTCTCATAAGTGAACTTGGGCCCTGT
oAS_KanR_F	Ntrl2_PLtet_UTR1	CCCGGTCTCACTTAAATTTTGTTTAACTTTAAGAAGGAGA
oAS_KanR_R	Ntrl2_PLtet_UTR1	CCCGGTCTCACGTAGCGTTCACCGACAAACAACA
oAS_Ntrl2_3'_F	Ntrl2_PLtet_UTR1	CCCGGTCTCATACGAAGCAAAGCACATGCC
oAS_Ntrl2_3'_R	Ntrl2_PLtet_UTR1	TGAGCTAAACAAGCTGCAAAAAC

Once these constructs were designed, the idea was to first allow homologous

recombination of the N2 region into ADP1 N1 GFP, to transform *kanR* into the genome.

Selection for this step was quite easy as colonies that appeared on a kanamycin plate should have

integrated the construct successfully. Once a colony was selected for growth overnight, the

genomic DNA was extracted and used for PCR to amplify the N2 region. The product was

sequenced to confirm the correct sequence before continuing to build the sensor. The next step

was transformation of our N1-KRAS-LacI construct into the N1 region, knocking out the

preexisting GFP gene. Colonies were screened for loss of GFP using colored filters over a blue

light. Again, colonies were picked for genomic DNA extraction, with PCR to amplify the N1

region this time. The product was sent in for sequencing once again. PCR primers used for entire

N1 and N2 region amplification can be seen in Table 2.2 below.

Table 2.2: PCR primers used to amplify N1 and N2 regions from genomic DNA.

Primer Name	Primer Sequence
oAS_Ntrl1_5'_F_mod	TTTTAGATGATAAGGGTTCTGC
oAS_Ntrl1_3'_R_mod	TAAACGTTTACCTAAATCACGC
oAS_Ntrl2_5'_F	CGTTACCAGCTCCACCAGTT
oAS_Ntrl2_3'_R	TGAGCTAAACAAGCTGCAAAAAC

Initially, this final step was conducted without antibiotic plates, since the *lacI*, if

successfully transformed, would already begin repressing *kanR* production. I quickly found that

contamination was a huge issue as most white colonies that I picked somehow yielded a random

specR gene in the N1 region instead of *lacI*. I started adding IPTG to the cultures and plates to

inactivate *lacI* repression, so I could grow the colonies on kanamycin and remove any unwanted

contaminants. Sadly, the *specR* gene kept making a reappearance. I believed I was successful

with this transformation when finally, I picked a white colony and ran a colony PCR to amplify

the N1 region. Sequencing successfully yielded the desired LacI sequence flanked by KRAS

homology arms. Unfortunately, when continuing to the next step for testing required kanamycin

42

dilution levels, I found that the supposedly correct strain grew in liquid cultures containing 2x and even 3x the amount of standard kanamycin, without IPTG. This indicated that repression was not occurring, which was puzzling. PCR off the genomic DNA extract kept yielding no bands in the gel electrophoresis, suggesting that the N1 region somehow got deleted in the bacteria. The fact that the colony PCR worked, however, could have meant that the linear DNA strand was stuck in the ECM or outside of the genome, somehow. Unfortunately, time ran out to complete this project and I was unable to successfully develop the *lacI*-based cancer biosensor.

I had to proceed with characterization of the ISx and CRAΦ deletion strains containing the *tetR* system. For this experiment, I used primers to amplify KRAS DNA from RKO and LS174T genomic DNA that was provided to me by Dr. Cooper, via PCR. I took seven different strains, listed in Table 2.3 below, and cultured them overnight. The following day, they were centrifuged, and the pellet was washed and diluted 10-fold. The washed cells were then incubated for two hours in a shaking incubator with 100 ng of the DNA obtained from PCR. A no DNA control was run as well. From here, I serially diluted each culture to six dilution levels of 10x, ranging from a dilution to the order of 1E-1 to 1E-6, and spotted them onto triplicate runs of LB only and Kan-containing 2% LB agar plates. The colonies were counted after a day of growth to generate results for the stability and successful sensing rate.

Results of the stability and HGT characterization can be seen in the following section. The KRAS sequence obtained from the RKO and LS174T lysate was around 1.5 kb in size. The RKO segment was a WT KRAS sequence while the LS174T had the mutated G12D version. Expected results should hopefully see the strains with WT KRAS targeting sequence only sense DNA from LS174T. Additionally, we should also see higher stability in the ISx strains as well as higher competency or HGT uptake rate.

Table 2.3: Strains tested for characterization.

Strain Name	Description
ADP1F	Biosensor without any PAM sequence
ADP1F CRRand	Biosensor with a random CRISPR spacer, which would not degrade any DNA
ADP1F CRKRAS	Biosensor with the WT KRAS targeting spacer
ISx	Biosensor without any PAM sequence, with IS and CRAΦ deletion
ISx CRRand	Biosensor with a random CRISPR spacer, which would not degrade any DNA, with IS and CRAΦ deletion
ISx CRKRAS	Biosensor with the WT KRAS targeting spacer, with IS and CRAΦ deletion
ISx Spaced CRKRAS	Biosensor with the WT KRAS targeting spacer, and spaced homology arms to match the size of the donor cassette, with IS and CRAΦ deletion

2.5 Characterization of Biosensor After Improvements

Characterization of the improved biosensor involved testing the HGT capabilities compared to previous designs. As stated in the methods, the experiment was done in liquid culture, adding 100 ng of KRAS DNA, from either the RKO or LS174T line, to the washed cells. The total HGT rate was measured over a given span of two hours. Two measurement replicates were plated on each plate with three total technical replicates. The standard error was determined from the individual HGT rate of the three technical replicates. Additionally, the log was taken to scale the results better, and the lower limit of detection was noted as one colony counted at dilution level one divided by the average count of all colonies from the LB-only plates. This allowed us to see if the data was real and above the lower limit of detection when comparing to the no DNA control. The lower bound was set as the lower limit of detection and any data below that was artificially raised to be at the limit of detection for visualization. Figure 2.4 details the findings.
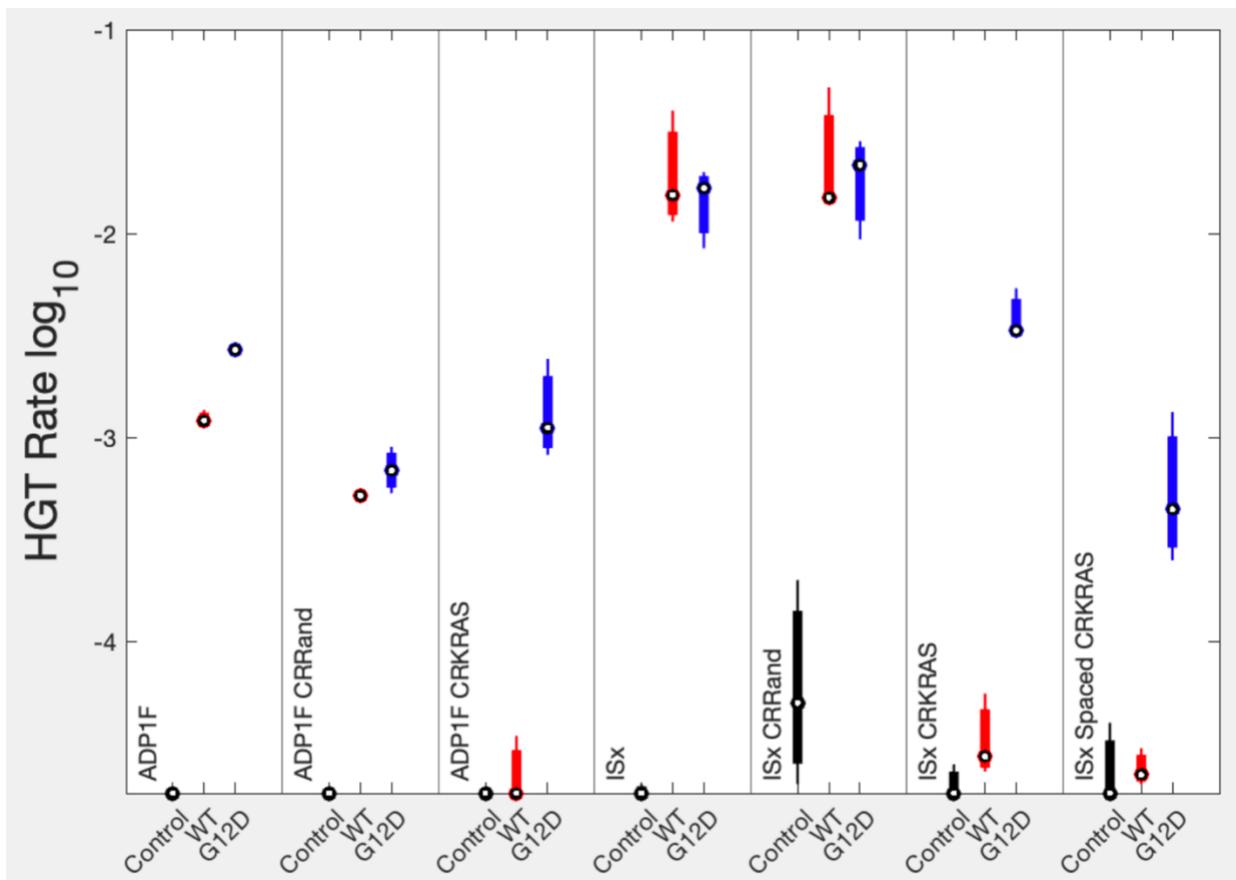
Figure 2.4: HGT rate in $\log_{10}$ for various versions of the biosensor, tested on KRAS WT, KRAS G12D, and control with no DNA. Log of limit of lower detection was calculated to be approximately -4.74, and replicates with 0 colonies were artificially raised to that level, so they would appear on the chart. This is seen for the control strains of AD1F, ADP1F CRRand, ADP1F CRKRAS, and ISx. WT KRAS came from RKO cells while G12D mutated KRAS came from LS174T cells. n = 3.

From Figure 2.4, a few things can be noted. First, we can look at the mutation rate of each strain, as seen from calculating the HGT rate with no DNA. If Kan resistance was developed without any DNA to sense, we can assume the strain developed a mutation or there is some leakage of the repressed gene. So, this is measured as mutation rate or just background noise. All strains without the ISx and CRA deletion that measured no colonies with serial dilution had to be artificially raised to the limit of lower detection. However, ISx CRRand, ISx CRKRAS, and ISx spaced CRKRAS measured slightly above the background noise. This is

surprising since we expected the ISx strains to acquire fewer mutations and generally be more stable.

On a brighter note, it appears that ISx strains generally had a higher HGT rate when cultured with DNA. Looking at the ISx versus ADP1F strain, both RKO and LS174T DNA were taken up at a higher rate. The same is true when looking at the ISx CRRand versus the ADP1F CRRand, as well as the ISx CRKRAS versus the ADP1F CRKRAS. For each strain tested, the ISx version appeared to have increased in competency. However, we should note that ISx CRRand and ISx CRKRAS also have a higher background from the mutations with no DNA present.

Additionally, it appears that the constructs can successfully distinguish between WT KRAS and KRASG12D when provided with PAM spacers targeting WT KRAS. The strains containing random or no PAM spacers also behaved as predicted. We see that ADP1F, ADP1F CRRand, ISx, and ISx CRRand all detect KRAS DNA without discrimination. This makes sense since there is no degradation of the WT KRAS occurring in these strains, so HGT still occurs. In both ISx CRKRAS and ADP1F CRKRAS, the mutated KRAS is sensed at a significantly higher rate than the WT KRAS. Looking at the ISx spaced CRKRAS sensor, it seems that it, unfortunately, detected KRASG12D at a lower rate than the version without the spacer, but we still see a decreased HGT rate for the WT KRAS, indicating the PAM spacer works as intended.

From the data, we can draw the conclusion that the deletion of IS and CRA causes an increased sensitivity to DNA in liquid culture. However, background noise also seems to increase which was not expected. Additionally, the PAM spacer targeting WT KRAS successfully degrades the WT DNA but leaves the mutated version alone, allowing for distinguishing between non-cancerous and cancerous cells.

2.6 Discussion

  As mentioned previously, I was unsuccessful in building the *lacI*-based repression system in the biosensor. The genomic integration of the *kanR* into the N2 region was successful, likely due to a selection pressure being added with kanamycin resistance. Unfortunately, the second step was too difficult with the method used. It involved knocking out the GFP gene but there was no additional selection pressure allowing for easy transformation and subsequent selection. The resulting dilution and plating resulted in at most two to three white colonies every time the transformation was performed. And every time, sequencing of the white colony yielded a mutated colony that was not what I was looking for.

  I believed that I had been successful after performing a colony PCR on a white colony that appeared one time. The product was the correct size and sequencing yielded the desired sequence containing the *lacI* insert. Unfortunately, PCR on the genomic DNA was unsuccessful in amplifying the N1 region where the insert was supposed to be, indicating that perhaps the N1 region was not in the genome anymore and somehow had been deleted. Additionally, the colony grew in twice and thrice the usual amount of kanamycin, without IPTG, which deforms *lacI* to prevent repression, indicating that either the repression was not working or was extremely weak. In both cases, it was clear that whatever strain I had right now could not be a better sensor than the existing ADP1 construct containing *tetR*. Time ran out to try other transformation methods for this project, as well as some shipping issues, but I detail a method that could work in the following section.

  Since the *lacI*-based biosensor construction failed, what remained was the characterization of the other improvements made on the sensor. The deletion of the IS and CRA phage needed to be studied to see if stability had been improved a little with the *tetR* based system. The results indicate that competency improved with the ISx strains. All versions of the

47

ISx strains had a much higher HGT rate compared to the ADP1F counterparts that did not have the improvements. Additionally, the PAM spacer targeting the WT KRAS was proven to successfully degrade only WT KRAS and allow for increased sensitivity to only KRASG12D.

Unfortunately, it must also be noted that the background noise or mutation rate was higher for the ISx strains. This is surprising because the IS deletion should have reduced the mutation rate according to the group that performed it. I believe that the stress imposed by the *tetR* system could be activating other repair mechanisms in the cell, such as error-prone polymerases, and deletion of the IS causes more stress. Even though the background noise is higher, the increased competency is certainly a benefit, indicating some minor success in improving the sensitivity of the biosensor. Additionally, it seems that the spacer added in the homology arms to decrease conformational discrepancies did not work as intended. The ISx spaced CRKRAS strain actually had a lower level of detection compared to the strain without the spacer, indicating that the size of the spacer needs to be modified, as it currently is hindering HGT rates. Further stability improvements regarding background noise reduction can certainly be made for the strain, which are referenced below.

2.7 Future Directions

2.7.1 Dual Cassette Selection

Since the current method of transformation of non-selective DNA elements into ADP1 is quite difficult and relies a lot on chance by hoping that the diluted volume has the transformed colony present, another method is sorely needed. Research has suggested a double-kill gene cassette is useful in transforming non-selective DNA segments into *A. baylyi*[48]. This method uses an inducible kill gene already present in the bacteria, which when not transformed out via crossover, can kill the bacteria with the provided selection conditions. Two such inducible kill genes were presented in the paper, referred to as *sacB* and *hok*. *SacB* results in death with the

presence of sucrose, and IPTG activates hok. I believe a method to transform the ADP1 could be to create a new strain, with a blank N2 region. In the N1 region, where *lacI* needs to be inserted, the *sacB* and *hok* genes can be inserted first, with an antibiotic resistance so selection of colonies bearing *sacB* and *hok* is made easy. Then, for the genome transformation, the KRAS *lacI* insert can first be added to the N1 region, and plated on sucrose and IPTG-containing plates to activate the kill gene unless the *sacB* and *hok* are removed. Once selection of *lacI* bearing colonies is successful, the *kanR* gene can be inserted into the N2 region and plated on kanamycin plates with IPTG to stop *lacI* suppression. This should ideally force a selection pressure for the otherwise non-selection-bearing *lacI* insert.

Another method involves the use of 3-azido-2′,3′-dideoxythymidine (AZT). A group of researchers were able to do a transformation of non-selective segments using, first, a DNA cassette with a dual positive and negative selection functionality introduced with *tdk* and *kanR*[49]. Successful transformants were first selected on the positive selection agent, kanamycin, meaning it also had the *tdk* gene. In the second transformation event, a rescue cassette was introduced to knock out the *tdk-kanR* segment and replace it with the gene of interest. This batch was plated on LB containing AZT, which would kill the bacteria if it still contained *tdk*, which is a thymidine kinase. So, this method can also be used to create our *lacI*-containing strain. The *tdk-kanR* cassette can first be introduced into the N1 region and selected for on a kanamycin plate. Then, the LacI rescue cassette can then be introduced and selected for using AZT. Finally, the *pTrc-kanR* segment can be inserted into the N2 region and selected for using IPTG and kanamycin again. This method should ideally provide an easier selection process compared to the knockout of GFP.

2.7.2 Further Stability Improvements

From the results above we have concluded that the ISx strain sees an increase in background noise, which is the opposite of what was expected. Sensitivity was increased to KRAS G12D however, so the results show that we are on a promising track for the improvement of the biosensor. There are a few more steps that can be taken regarding reduction of mutagenesis rates in ADP1. In *E. coli*, researchers were able to see reduced mutagenesis rates with the deletion of genes associated with stress-inducible error-prone DNA polymerases[50]. Targeted genes included *polB*, *dinB*, and *umuDC*, seeing close to a 50% decrease in mutation rates, especially when placed in stressful environments. Perhaps a similar study can be performed on ADP1, as *dinB* and a homolog of *umuDC* exist in the genome[51]. Additionally, other error-prone polymerases are also activated with the SOS response to DNA damage. *UmuDAb* is another important SOS gene that activates *ddrR* upon recognizing DNA damage. The mechanism also involves *lexA*, which unbinds from the SOS genes when DNA damage is recognized. All of these genes are of interest as they relate to repair of DNA damage and are more prone to errors when performing the repair. If these genes can be deactivated, forcing less error-prone polymerase to take over when DNA damage occurs, perhaps the stability of the biosensor can be increased, reducing the background mutation rate. Additionally, it might be worth looking into the SOS response and trying to deactivate it altogether or attach a kill gene to the response, so any hint of a mutation will kill the bacteria and prevent a false positive read from the sensor.

These are all potential avenues to explore in the future and could lead to a higher stability of the sensor. Along with the replacement of the *tetR* gene with *lacI*, increasing the stability in these various ways will serve to prepare the transition of this cancer biosensor from a lab setting to a plausible alternative diagnostic tool for CRC tumors.

Chapter 3 – IMPROVING CIRCUIT BEHAVIOR WITH ALE

3.1 Issues with Circuit Behavior in Non-Traditional Growth Media

As discussed throughout this thesis, synthetic biology enables the design of novel circuits that detect, treat, and quantify a multitude of phenomena. Two such strains are mentioned above, with the SLC cancer therapeutic and the HGT cancer biosensor[2,3]. Other examples involve strains that can detect inflammation by passing through the gut microbiome[37]. One thing these strains all have in common is the necessity for their presence in non-traditional growth media. In the lab, the typical method of culture for bacteria involves a growth medium like LB broth. This broth contains a carbon source via yeast extract as well as important amino acids, salts, and various other ingredients that all contribute to the proliferation of the bacteria[52]. However, these strains are all developed with the intention of detecting, treating, and quantifying phenomena in the human body. Areas like the gut microbiome, tumor microenvironment, and general tracts traveled by bacteria are not anywhere near the conditions simulated with plain LB. Complex growth environments lead to increasing unpredictability of circuit behavior, which is a well-established fact in synthetic biology[53]. Unfortunately, real-world conditions, where these complex growth environments are typically found, are what synthetic circuits need to perform their intended function in for any hope of translation from the lab to a viable therapeutic or diagnostic.

The human body is incredibly complex, and the conditions bacteria are subjected to in lab cannot quite encapsulate every detail of the environment. Nevertheless, there are methods to simulate specific niches of the localized treatment area, such as the microenvironment in the tumor or the gut microbiome. Simulating these conditions, along with the use of adaptive laboratory evolution (ALE) to force genomic changes in the chassis harboring the circuit, can

lead to better overall performance of the circuit in non-traditional growth media. Previously, we had already seen improvement of the SLC in different media like M9 lactate, which is a minimal media containing the carbon source lactate, simulating a media more representative of the tumor environment, with elevated lactate[54].

As previously discussed, synthetic biological circuits have many applications *in vivo*. However, the growth conditions in lab using LB are not representative of the conditions faced by the bacteria in their intended locations of treatment or diagnostic. For example, the tumor microenvironment (TME), where the SLC is intended to be applied, is a complex ecosystem surrounding cancer cells, composed of immune cells, stromal cells, blood vessels, ECM components, and various small molecules[55]. It plays a critical role in tumor progression, metastasis, and response to therapies. It has even been recently discovered that commensal bacteria are present throughout the tumor microbiome, which just goes to show how complex the system is[56]. With all these complex growth conditions, testing circuit behavior in simple LB conditions is not representative of the environment intended for the therapeutic or diagnostic circuit. As such, circuit behavior in conditions representative of the TME is required and suggests that the main sources of stress in the TME need to first be identified.

One significant aspect of the TME is the metabolic reprogramming observed in cancer cells, which affects the ECM. Tumor cells often exhibit increased glycolysis, even in the presence of oxygen, leading to elevated lactate production and hypoxic environments in a phenomenon referred to as the Warburg effect[57]. Lactate acts as a signaling molecule, influencing various aspects of tumor biology, including immune cell function, angiogenesis, and metastasis, while simultaneously providing an alternative carbon source to glucose.

Alternative carbon sources will undoubtedly affect the performance of our synthetic circuits intended for use within the TME. As such, there is a need for the rapid improvement of the chassis housing these diagnostics and therapeutics. With the use of ALE, we can quickly adapt bacteria to grow better in conditions such as elevated lactate.

3.2 Adaptive Laboratory Evolution

ALE is a vital scientific approach that enables the study of evolutionary processes in a controlled laboratory environment. There has been a significant resurgence of ALE experiments in the last 25 years, with *E. coli* and *S. cerevisiae* being the primary organisms of interest[58]. In microbial ALE, microorganisms are cultured under precisely defined conditions for extended periods, ranging from weeks to years, rapidly evolving the phenotypes to thrive in the established conditions. Microbial cells offer several advantages for ALE studies: they have simple nutrient requirements, are easily cultivated in the laboratory, and have fast doubling times, enabling hundreds of generations to be cultivated within weeks or months. Combined with bioinformatic tools like high-throughput single cell sequencing, ALE allows for clear associations between phenotypic and genotypic changes, along with the specific growth environments that drive the trait selection.
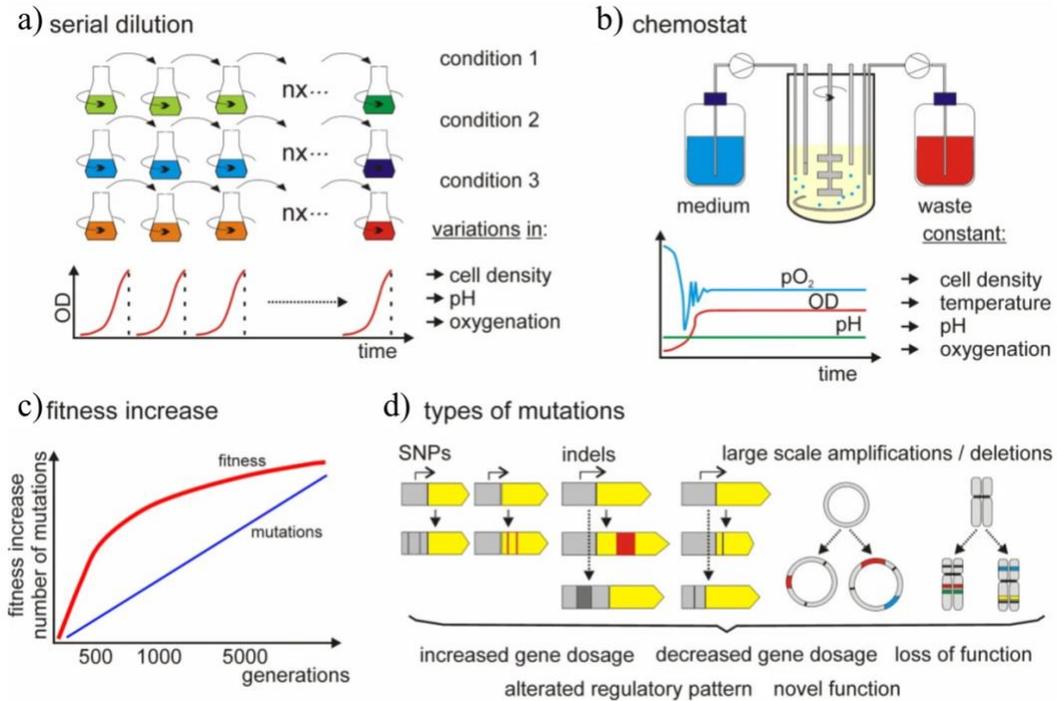
Figure 3.1: Adaptive laboratory evolution process[58]**.** a) One method of performing ALE is via serial passages in flasks, ensuring the desired growth conditions for directed mutagenesis. b) A second method can be via chemostat, where several parameters can be held constant. c) The fitness increase during ALE is rapid in the beginning stages but generally falls off as generations increase, while the number of mutations is steadily increasing. Network complexity leads to a decreasing beneficial effect of additional mutations. d) Typical mutations identified from ALE studies. Single nucleotide polymorphisms (SNPs) and insertions and deletions (indels) contribute to regulatory and fitness changes during the selection for improved survivability. Figure adapted from Dragosits et al.

Recent advancements in technologies such as transcriptional profiling and next-generation DNA sequencing have also facilitated the correlation of phenotypes with genotypes through whole-genome re-sequencing, further advancing the utilization of ALE experiments to characterize genes associated with phenotypic changes for survival[59]. These insights include understanding the genetic basis of increased fitness in non-traditional media and second-order effects during evolution. Overall, ALE experiments continue to offer valuable insights into the mechanisms driving microbial adaptation and evolution, contributing to our broader understanding of evolutionary processes. Furthermore, ALE can be used to direct mutagenesis to

ensure survivability in non-traditional media with specific stresses. For example, we can use ALE to drive adaptation to factors present in the TME, like elevated lactate.

Here, ALE was used to promote the survival and circuit behavior of MG1655 in a minimal media, M9, with an alternative carbon source, lactate[54]. We characterized the lysis recovery and growth rates in the non-traditional media. The 89th passage of the MG1655 strain, labeled BOP27, were selected for their performance and then tested. Since the SLC strain uses AHL as the quorum sensing signal molecule, it was necessary to characterize the response to AHL in the evolved media. The AHL inducible section of the SLC was taken to observe only AHL induced GFP expression, which was then observed in the evolved MG1655 in M9 lactate.

Additionally, though the SLC was originally built in *E. coli* MG1655, *E. coli* Nissle 1917 (EcN) is a better suited strain for clinical application due to a favorable safety profile and relevance in the treatment and detection of cancers[60]. To be able to compare circuit behavior between MG1655 and EcN, especially aspects like lysis response, I modified a lysis circuit with constitutive GFP to make both the E lysis protein and GFP inducible with AHL. This would then directly indicate the amount of lysis protein being generated, rather than indicate total protein generation as indicated by the constitutive GFP production. Previous experiments suggest that EcN produces less protein overall than MG1655, so having a direct comparison of the lysis protein production specifically would allow for a better characterization of the lysis response.

3.3 Materials and Methods

Previously, the first strain, wildtype *E. coli* K-12 MG1655, was evolved in M9, a minimal media, with lactate as the sole carbon source[54]. In batch culture growth and passage, it was observed that the growth rate in M9 lactate increased with subsequent evolutionary passages in all replicates. An individual clone from each replicate evolution was sequenced for mutations at the end of the evolution. Mutations in *ppsA*, phosphoenolpyruvate synthase in the

gluconeogenesis pathway, and *ilvH*, acetolactate synthase subunit, were discovered. Additionally, *rhsE*, a pseudogene, was also identified as a convergent mutation. The 89th passage was chosen to conduct my individual studies, as the accumulated mutations were found to allow for enhanced metabolic activity in M9 lactate.

The first part of my contribution to this work involved the study of an inducible GFP circuit, driven by the *pLuxI* promoter which is induced with AHL. This would allow for a characterization of the AHL induction which is important to the function of the SLC strain. This plasmid was named M4, and I transformed it into the evolved MG1655, as well as the WT version. The plasmid was simple, with *pLuxI* driving GFP, constitutive *luxR* and *specR* for AHL inducibility and positive selection, respectively. Once these circuits were transformed using heat shock, plate reader experiments were performed to visualize OD and GFP curves for each strain in the non-traditional media.

The second part of this project involved the creation of an inducible lysis circuit which also induced GFP expression. Since EcN was a strain of interest for the application of the SLC *in vivo*, an assay needed to be designed to look at the lysis response of each strain compared to how much overall protein and how much lysis protein was generated. Previously, the inducible lysis circuit being used was named pJZ50v2, which had the X174E lysis protein driven by *pLuxI*. GFP production was constitutive but was not able to accurately represent the production of lysis protein following 1 nM AHL induction, instead indicating more of a baseline protein production level in the strain. I modified the plasmid to delete the constitutive promoter on the GFP and added an RBS in between the X174E protein and the GFP. Now, hypothetically, the E lysis protein generation could be characterized after AHL induction, before lysis, to see the buildup of only that protein, indicated by GFP level. Then, the lysis magnitude could be calculated to better

understand the amount of lysis protein necessary to reach a certain lysis magnitude for each

strain. This plasmid was transformed into WT MG1655 and EcN via heat shock and

electroporation, respectively. Seen below are the gene circuit diagrams describing the

mechanism within the M4, pJZ50v2, and inducible lysis plasmids.
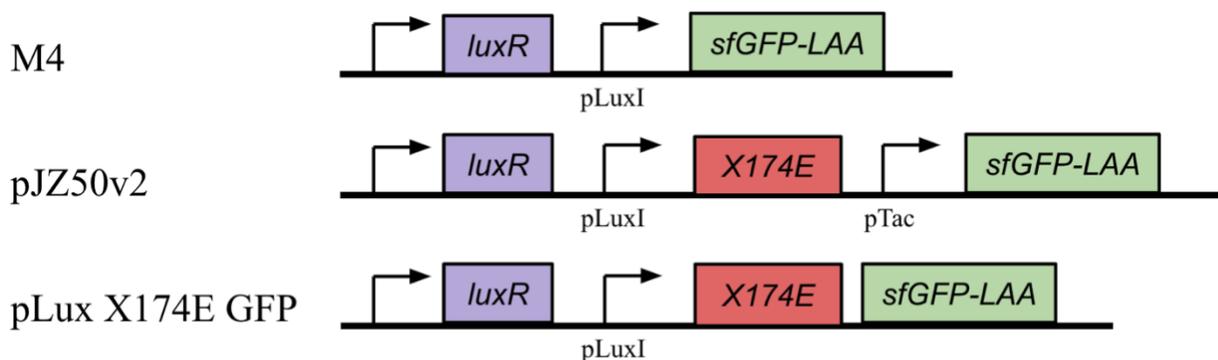
Figure 3.2: Gene circuit diagrams for M4, pJZ50v2, and pLux X174E GFP. The M4 circuit
is simple, with constitutive *luxR* production and *pLuxI* (induced by AHL) driving production of
GFP with an LAA degradation tag. In pJZ50v2, the GFP production is also made constitutive
with the *pTac* promoter. Additionally, *pLuxI* now drives X174E lysis protein production. In *pLux*
X174E GFP, which is the name of the inducible GFP lysis circuit, the *pTac* constitutive promoter
is removed and replaced with an RBS, so the GFP is now linked to the *pLuxI* promoter.

3.4 Results

To begin this experiment, MG1655 was transformed with the M4 plasmid, which allowed

for GFP expression induced by AHL addition. Overnights (ONs) were grown in M9 lactate, and

the WT and the evolved strain, P89, were both tested to characterize growth and GFP expression.

An AHL induction concentration of 1 nM was used to test sensitivity to AHL in the evolved

media. The results characterized are seen in Figure 3.3 and 3.4 below.
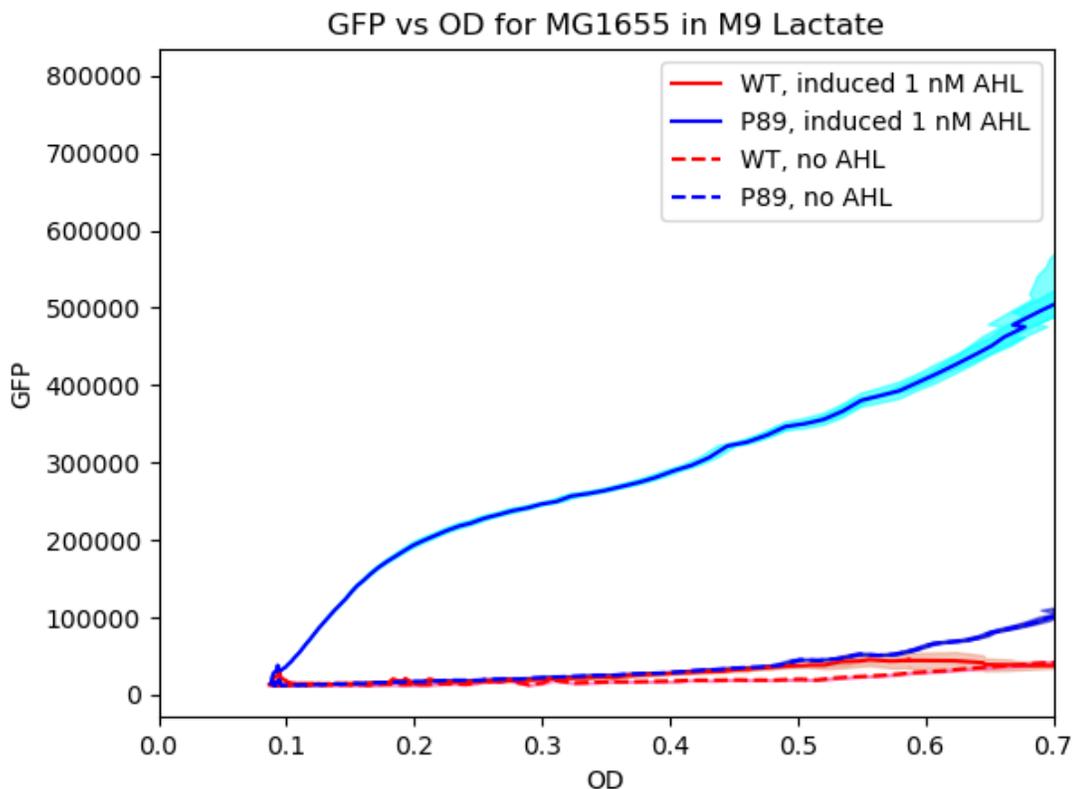
Figure 3.3: GFP vs OD curves for M4 transformed into MG1655, grown in M9 lactate. The evolved strain, when induced with 1 nM AHL, is the only strain that produces GFP. The WT strain is unable to produce any GFP when induced, suggesting that sensitivity to AHL in the evolved media is better for the evolved strain. All data points represent mean of GFP and OD values (solid or dashed line) ± SEM of GFP and OD values (shaded areas), n = 3.

As seen in Figure 3.3, the only strain that produces GFP is the evolved P89 version, when induced with 1 nM of AHL. The cultures that did not receive AHL produced no GFP as expected, though there is a slight production for the evolved strain as OD passes 0.6, suggesting some leakiness. When the WT strain received AHL, no GFP was produced at all. From these results, we can surmise that the evolved strain has an increased sensitivity to AHL in M9 lactate, suggesting the accumulated mutations help circuit performance. The individual growth curves tracked by the OD can be seen in Figure 3.4.
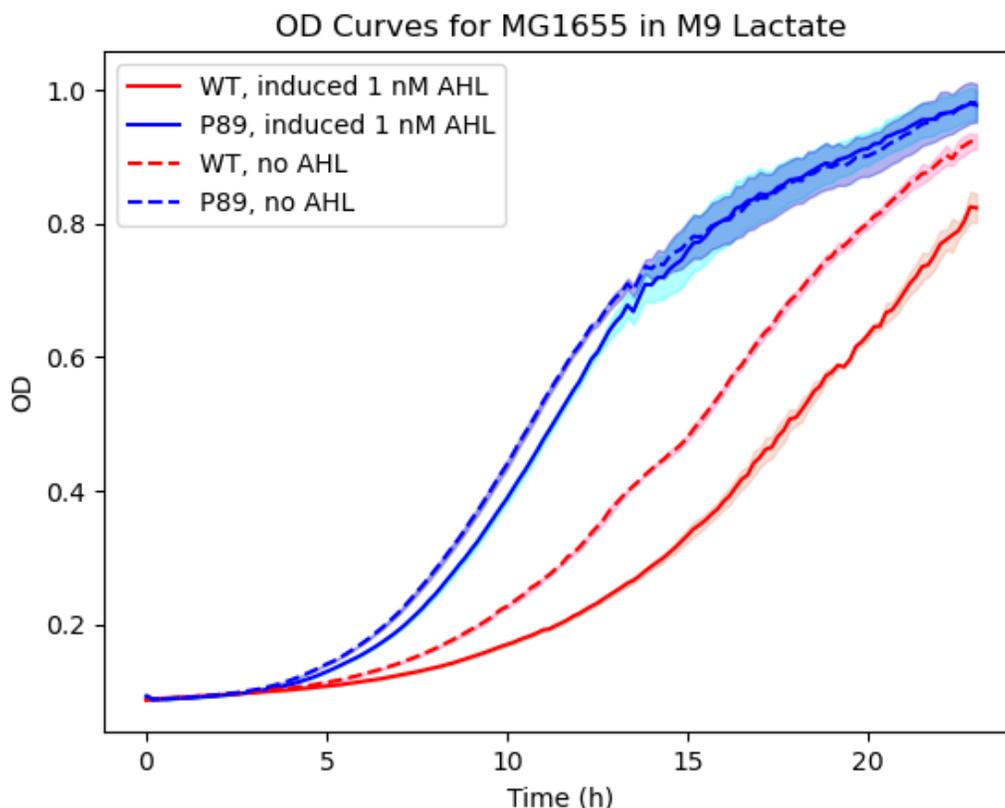
Figure 3.4: OD curves for M4 transformed into MG1655, grown in M9 lactate. Both the uninduced and induced evolved strain have a higher growth rate than the WT strains. Additionally, it appears that induction with AHL causes the growth rate of the WT strain to further decrease, while the evolved strain, when induced, sees a minimal drop in the growth rate. All data points represent mean (solid or dashed line) ± SEM (shaded areas), n = 3.

Above, we see that the growth rates of the evolved P89 strain, regardless of induction or not, are higher than the WT. Additionally, it appears in the WT that induction with 1 nM AHL results in a significant decrease in the growth rate, suggesting the burden is extremely high when attempting to produce GFP in the M9 lactate media. For the evolved strain, induction sees a considerably less significant drop in the growth rate, meaning that the accumulated mutations help relieve some of the burden the strain is subjected to when performing the circuit functions in M9 lactate.

Moving forward, using the transformed pJZ50v2 and inducible lysis strains in the WT MG1655 and EcN, lysis induction curves were generated. Once batches reached an OD of 0.3,

they were induced with 10 nM of AHL, and the GFP vs OD curve was plotted up until lysis occurred. The lysis response was measured by dividing the lysis magnitude by the GFP vs. OD slope.
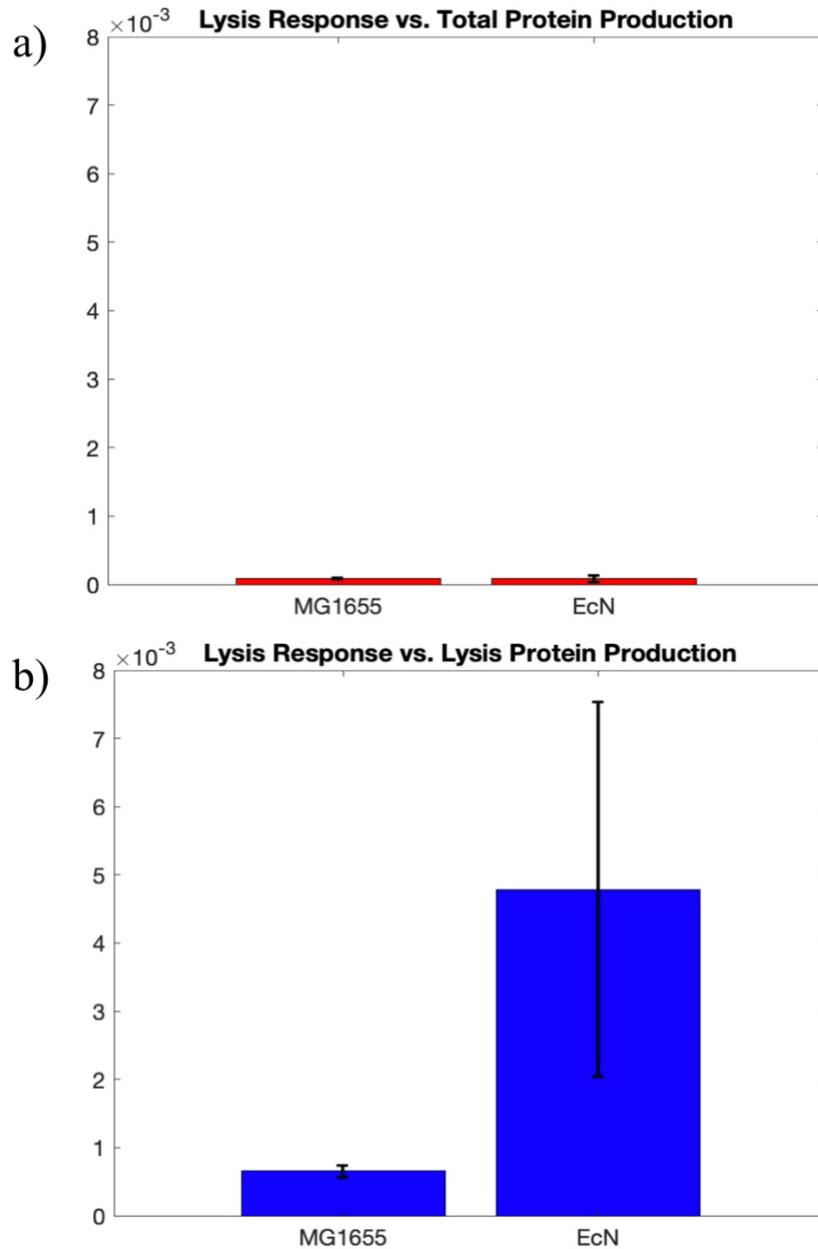
Figure 3.5: Lysis magnitude divided by GFP vs. OD slope for each strain. a) Data for pJZ50v2. This gives an indication of the lysis magnitude per total protein production. b) Data for *pLuxI* X174E GFP. This gives an indication of the lysis magnitude per lysis protein production. Error bars represent SEM, n = 3.

As seen in Figure 3.5a, comparison of the lysis response using constitutive GFP expression, which gives an indication of the average protein production, does not yield a significant difference between the two strains. However, from Figure 3.5b, EcN appears to have a higher lysis response to the amount of lysis protein compared to MG1655, suggesting a higher sensitivity to the E lysis protein.

3.5 Discussion

As seen in the above results, the strain evolved with ALE in M9 lactate shows an increased sensitivity to AHL in the evolved media, as well as a resistance to the burden caused by induction. P89 clearly produces a higher amount of GFP than the WT strain when induced with 1 nM of AHL. Additionally, the OD curves show that the drop in the growth rate for the evolved strain, when induced, is insignificant compared to the drop in the growth rate seen from inducing the WT strain. This suggests that the circuit function and sensitivity improved with ALE, and the accumulated mutations help decrease the burden in the evolved media. As a proof of concept, using the M4 plasmid as the basis of the study, we have proven that ALE is a viable option in improving chassis health and performance in nontraditional media. Further experiments can be run using this evolved strain in M9 lactate to test other inducible circuit components, such as IPTG or *tet* systems.

The second part of this project suggests that EcN has a higher sensitivity to the lysis protein compared to MG1655. Though the error bars are significant in Figure 3.5, the experiment works as a proof of concept to indicate the lysis efficiency, calculated from lysis magnitude per lysis protein made. Further experiments need to be run to better characterize the inducible lysis system in different strains. This will further aid in determining the lysis response of various strains before one can be chosen for the optimal circuit behavior as a therapeutic. Additionally, models can be made to predict the lysis response in these various strains, perhaps using a

parameter to indicate sensitivity to the lysis protein, which can further be confirmed using the inducible GFP lysis circuit.

Now that it has been proven that MG1655 has a higher sensitivity to AHL in M9 lactate following evolution via ALE, future directions would include testing these strains *in vivo*. Since high lactate levels are a characteristic of the TME, testing the circuit behavior *in vivo* for its intended purpose would further prove the value of ALE in improving circuits for therapies. The experiments performed in the original SLC publication should be rerun with the evolved strains containing the SLC to compare performance against organoids and mice tumor models. The implications of this work are a huge step in optimizing synthetic circuit behavior for use in cancer therapeutics, and many other conditions where non-traditional stresses may be imposed on the strain.

REFERENCES

1. Danino, T., Mondragón-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463,** 326–330 (2010).

2. Din, M. O., Danino, T., Prindle, A., Skalak, M., Selimkhanov, J., Allen, K., Julio, E., Atolia, E., Tsimring, L. S., Bhatia, S. N. & Hasty, J. Synchronized cycles of bacterial lysis for in vivo delivery. *Nature* **536,** 81–85 (2016).

3. Cooper, R. M., Wright, J. A., Ng, J. Q., Goyne, J. M., Suzuki, N., Lee, Y. K., Ichinose, M., Radford, G., Ryan, F. J., Kumar, S., Thomas, E. M., Vrbanac, L., Knight, R., Woods, S. L., Worthley, D. L. & Hasty, J. Engineered bacteria detect tumor DNA. (2023).

4. Cohen, S. N., Chang, A. C. Y., Boyer H. W. & Helling, R. B. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Nat Acad Sci USA* (1973).

5. Nistala, G. J., Wu, K., Rao, C. V. & Bhalerao, K. D. A modular positive feedback-based gene amplifier. *J. Biol. Eng.* **4,** 4 (2010).

6. Bernhardt, T. G., Struck, D. K. & Young, R. The Lysis Protein E of φX174 Is a Specific Inhibitor of the MraY-catalyzed Step in Peptidoglycan Synthesis*. *J. Biol. Chem.* **276,** 6093–6097 (2001).

7. Xiong, L., Cao, Y., Cooper, R., Rappel, W.-J., Hasty, J. & Tsimring, L. Flower-like patterns in multi-species bacterial colonies. *eLife* **9,** e48885 (2020).

8. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B. & Shao, Y. The Complete Genome Sequence of *Escherichia coli* K-12. *Science* **277,** 1453–1462 (1997).

9. Palma, V., Gutiérrez, M. S., Vargas, O., Parthasarathy, R. & Navarrete, P. Methods to Evaluate Bacterial Motility and Its Role in Bacterial–Host Interactions. *Microorganisms* **10,** 563 (2022).

10. Matsushita, M. & Fujikawa, H. Diffusion-limited growth in bacterial colony formation. *Phys. Stat. Mech. Its Appl.* **168,** 498–506 (1990).

11. Fuentes, L. G., Lara, A. R., Martínez, L. M., Ramírez, O. T., Martínez, A., Bolívar, F. & Gosset, G. Modification of glucose import capacity in Escherichia coli: physiologic consequences and utility for improving DNA vaccine production. *Microb. Cell Factories* **12,** 42 (2013).

12. Delgado, J. & Liao, J. C. Inverse Flux Analysis for Reduction of Acetate Excretion in Escherichia coli. *Biotechnol. Prog.* **13,** 361–367 (1997).

13. Banerjee, P., Soni, J., Purwar, H., Ghosh, N. & Sengupta, T. K. Probing the fractal pattern and organization of *Bacillus thuringiensis* bacteria colonies growing under different conditions using quantitative spectral light scattering polarimetry. *J. Biomed. Opt.* **18,** 035003 (2013).

14. Ambagaspitiye, S., Sudarshan, S., Hogins, J., McDill, P., De Nisco, N. J., Zimmern, P. E. & Reitzer, L. Fimbriae and flagella mediated surface motility and the effect of glucose on nonpathogenic and uropathogenic *Escherichia coli*. Preprint at https://doi.org/10.1101/840991 (2019)

15. Kimata, K., Takahashi, H., Inada, T., Postma, P. & Aiba, H. cAMP receptor protein–cAMP plays a crucial role in glucose–lactose diauxie by activating the major glucose transporter gene in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **94,** 12914–12919 (1997).

16. Zheng, J., Lou, L., Fan, J., Huang, C., Mei, Q., Wu, J., Guo, Y., Lu, Y., Wang, X. & Zeng, Y. Commensal *Escherichia coli* Aggravates Acute Necrotizing Pancreatitis through Targeting of Intestinal Epithelial Cells. *Appl. Environ. Microbiol.* **85,** e00059-19 (2019).

17. Park, S., Park, Y.-H., Lee, C.-R., Kim, Y.-R. & Seok, Y.-J. Glucose induces delocalization of a flagellar biosynthesis protein from the flagellated pole. *Mol. Microbiol.* **101,** 795–808 (2016).

18. Dobrogosz, J. & Hamilton, P. B. Biochemical and biophysical research communications. *Biochem. Biophys. Res. Commun.* **42,** (1971).

19. Avelino-Flores, F., Soria-Bustos, J., Saldaña-Ahuactzi, Z., Martínez-Laguna, Y., Yañez-Santos, J. A., Cedillo-Ramírez, M. L. & Girón, J. A. The Transcription of Flagella of Enteropathogenic Escherichia coli O127:H6 Is Activated in Response to Environmental and Nutritional Signals. *Microorganisms* **10,** 792 (2022).

20. Sudarshan, S., Hogins, J., Ambagaspitiye, S., Zimmern, P. & Reitzer, L. Nutrient and Energy Pathway Requirements for Surface Motility of Nonpathogenic and Uropathogenic Escherichia coli. *J. Bacteriol.* **203,** (2021).

21. Gauger, E. J., Leatham, M. P., Mercado-Lubo, R., Laux, D. C., Conway, T. & Cohen, P. S. Role of Motility and the *flhDC* Operon in *Escherichia coli* MG1655 Colonization of the Mouse Intestine. *Infect. Immun.* **75,** 3315–3324 (2007).

22. Hansen, C. H., Endres, R. G. & Wingreen, N. S. Chemotaxis in Escherichia coli: a molecular model for robust precise adaptation. *PLoS Comput. Biol.* **preprint,** e1 (2005).

23. Shehadul Islam, M., Aryasomayajula, A. & Selvaganapathy, P. A Review on Macroscale and Microscale Cell Lysis Methods. *Micromachines* **8,** 83 (2017).

24. Shi, W. & Sun, H. Type IV Pilus-Dependent Motility and Its Possible Role in Bacterial Pathogenesis. *Infect. Immun.* **70,** 1–4 (2002).

25. Hospenthal, M. K., Costa, T. R. D. & Waksman, G. A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nat. Rev. Microbiol.* **15,** 365–379 (2017).

26. O'Toole, G. A. & Kolter, R. Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Mol. Microbiol.* **30,** 295–304 (1998).

27. Werneburg, G. T. & Thanassi, D. G. Pili Assembled by the Chaperone/Usher Pathway in Escherichia coli and Salmonella. *EcoSal Plus* **8,** 10.1128/ecosalplus.ESP-0007–2017 (2018).

28. Mobley Harry L. T., Donnenberg Michael S., & Hagan Erin C. Uropathogenic Escherichia coli. *EcoSal Plus* **3,** 10.1128/ecosalplus.8.6.1.3 (2009).

29. Pratt, L. A. & Kolter, R. Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Mol. Microbiol.* **30,** 285–293 (1998).

30. Hospenthal, M. K. & Waksman, G. The Remarkable Biomechanical Properties of the Type 1 Chaperone-Usher Pilus: A Structural and Molecular Perspective. *Microbiol. Spectr.* **7,** 10.1128/microbiolspec.psib-0010–2018 (2019).

31. Arumugam, G. & Tyagi, J. Keller-Segel Chemotaxis Models: A Review. *Acta Appl. Math.* **171,** 6 (2021).

32. Chen, T.-L., Siu, L.-K., Lee, Y.-T., Chen, C.-P., Huang, L.-Y., Wu, R. C.-C., Cho, W.-L. & Fung, C.-P. *Acinetobacter baylyi* as a Pathogen for Opportunistic Infection. *J. Clin. Microbiol.* **46,** 2938–2944 (2008).

33. Palmen, R., Buijsman, P. & Hellingwerf, K. J. Physiological regulation of competence induction for natural transformation in *Acinetobacter calcoaceticus*.

34. Young, D. M., Parke, D. & Ornston, L. N. Opportunities for genetic investigation afforded by *Acinetobacter baylyi*, a nutritionally versatile bacterial species that is highly competent for natural transformation. *Annu. Rev. Microbiol.* **59,** 519–551 (2005).

35. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* **16,** 472–482 (2015).

36. Danino, T., Prindle, A., Kwong, G. A., Skalak, M., Li, H., Allen, K., Hasty, J. & Bhatia, S. N. Programmable probiotics for detection of cancer in urine. *Sci. Transl. Med.* **7,** (2015).

37. Riglar, D. T., Giessen, T. W., Baym, M., Kerns, S. J., Niederhuber, M. J., Bronson, R. T., Kotula, J. W., Gerber, G. K., Way, J. C. & Silver, P. A. Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35,** 653–658 (2017).

38. Mimee, M., Nadeau, P., Hayward, A., Carim, S., Flanagan, S., Jerger, L., Collins, J., McDonnell, S., Swartwout, R., Citorik, R. J., Bulović, V., Langer, R., Traverso, G., Chandrakasan, A. P. & Lu, T. K. An ingestible bacterial-electronic system to monitor gastrointestinal health. *Science* **360,** 915–918 (2018).

39. Vogelstein Bert, Fearon Eric R., Hamilton Stanley R., Kern Scott E., Preisinger Ann C., Leppert Mark, Smits Alida M.M., & Bos Johannes L. Genetic Alterations during Colorectal-Tumor Development. *N. Engl. J. Med.* **319,** 525–532 (1988).

40. Cooper, R. M. & Hasty, J. One-Day Construction of Multiplex Arrays to Harness Natural CRISPR-Cas Systems. *ACS Synth. Biol.* **9,** 1129–1137 (2020).

41. Lutz, R. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* **25,** 1203–1210 (1997).

42. Renda, B. A., Chan, C., Parent, K. N., & Barrick, J. E. Emergence of a Competence-Reducing Filamentous Phage from the Genome of *Acinetobacter baylyi* ADP1. *J. Bacteriol.* **198,** 3209–3219 (2016).

43. Matthews, K. S. & Nichols, J. C. in *Prog. Nucleic Acid Res. Mol. Biol.* (ed. Moldave, K.) **58,** 127–164 (Academic Press, 1997).

44. Till, P., Toepel, J., Bühler, B., Mach, R. L. & Mach-Aigner, A. R. Regulatory systems for gene expression control in cyanobacteria. *Appl. Microbiol. Biotechnol.* **104,** 1977–1991 (2020).

45. Arvay, E., Biggs, B. W., Guerrero, L., Jiang, V. & Tyo, K. Engineering Acinetobacter baylyi ADP1 for mevalonate production from lignin-derived aromatic compounds. *Metab. Eng. Commun.* **13,** e00173 (2021).

46. Suárez, G. A., Renda, B. A., Dasgupta, A. & Barrick, J. E. Reduced Mutation Rate and Increased Transformability of Transposon-Free *Acinetobacter baylyi* ADP1-ISx. *Appl. Environ. Microbiol.* **83,** e01025-17 (2017).

47. Biggs, B. W., Bedore, S. R., Arvay, E., Huang, S., Subramanian, H., McIntyre, E. A., Duscent-Maitland, C. V., Neidle, E. L. & Tyo, K. E. J. Development of a genetic toolset for the highly engineerable and metabolically versatile Acinetobacter baylyi ADP1. *Nucleic Acids Res.* **48,** 5169–5182 (2020).

48. Harms, K., De Vries, J. & Wackernagel, W. A double kill gene cassette for the positive selection of transforming non-selective DNA segments in Acinetobacter baylyi BD413. *J. Microbiol. Methods* **69,** 107–115 (2007).

49. Suárez, G. A., Dugan, K. R., Renda, B. A., Leonard, S. P., Gangavarapu, L. S. & Barrick, J. E. Rapid and assured genetic engineering methods applied to *Acinetobacter baylyi* ADP1 genome streamlining. *Nucleic Acids Res.* **48,** 4585–4600 (2020).

50. Csörgő, B., Fehér, T., Tímár, E., Blattner, F. R. & Pósfai, G. Low-mutation-rate, reduced-genome Escherichia coli: an improved host for faithful maintenance of engineered genetic constructs. *Microb. Cell Factories* **11,** 11 (2012).

51. Hare, J. M., Ferrell, J. C., Witkowski, T. A. & Grice, A. N. Prophage Induction and Differential RecA and UmuDAb Transcriptome Regulation in the DNA Damage Responses of *Acinetobacter baumannii* and *Acinetobacter baylyi*. *PLoS ONE* **9,** e93861 (2014).

52. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. *Escherichia coli* Physiology in Luria-Bertani Broth. *J. Bacteriol.* **189,** 8746–8749 (2007).

53. Lezia, A., Miano, A. & Hasty, J. Synthetic Gene Circuits: Design, Implement, and Apply. *Proc. IEEE* **110,** 613–630 (2022).

54. Zhang, J. T., Lezia, A., Emmanuele, P., Wu, M., Olson, C. A., Feist, A. M. & Hasty, J. Host evolution improves genetic circuit function in complex growth environments. Preprint at https://doi.org/10.1101/2024.03.13.583595 (2024)

55. Anderson, N. M. & Simon, M. C. The tumor microenvironment. *Curr. Biol.* **30,** R921–R925 (2020).

56. Li, L., Chandra, V. & McAllister, F. Tumor-resident microbes: the new kids on the microenvironment block. *Trends Cancer* **10,** 347–355 (2024).

57. Spencer, N. Y. & Stanton, R. C. The Warburg Effect, Lactate, and Nearly a Century of Trying to Cure Cancer. *New Perspect. Acid-Base Balance* **39,** 380–393 (2019).

58. Dragosits, M. & Mattanovich, D. Adaptive laboratory evolution – principles and applications for biotechnology. *Microb. Cell Factories* **12,** 64 (2013).

59. Sandberg, T. E., Salazar, M. J., Weng, L. L., Palsson, B. O. & Feist, A. M. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metab. Eng.* **56,** 1–16 (2019).

60. Gurbatri, C. R., Radford, G., Vrbanac, L., Coker, C., Im, J.-W., Taylor, S. R., Jang, Y., Sivan, A., Rhee, K., Saleh, A. A., Chien, T., Zandkarimi, F., Lia, I., Lannagan, T. R. M., Wang, T., Wright, J. A., Thomas, E., Kobayashi, H., Ng, J. Q., Lawrence, M., Sammour, T., Thomas, M., Lewis, M., Papanicolas, L., Perry, J., Fitzsimmons, T., Kaazan, P., Lim, A., Marker, J., Ostroff, C., Rogers, G., Arpaia, N., Worthley, D. L., Woods, S. L. & Danino, T. Colorectal cancer detection and treatment with engineered probiotics. bioRxiv 2023.04.03.535370 (2023). doi:10.1101/2023.04.03.535370

A1. SLC Characterization Methods and Assays

A1.1 Transformation of Strains

Two different strains were transformed for this project. The SLC contained plasmids

pTD103 and pZA35, both of which were obtained from lab stocks and miniprepped from

overnights using the QIAprep® Spin Miniprep Kit from Qiagen. pTD103 has the colE1 origin of

replication site, superfolding GFP, *luxI* with an LAA degradation tag, and *luxR* driven by *pLuxI*,

and kanamycin resistance. pZA35 contained the p15a ori site with pLuxI driving the E lysis

protein and *luxR*. Additionally, it had the ptac promoter driving *hlyE* production and constitutive

chloramphenicol resistance. The QSC contained pTD103 as well as pJZ13, which had

constitutive RFP expression and chloramphenicol resistance. Plasmid concentrations were

measured via Nanodrop. See plasmid maps below.
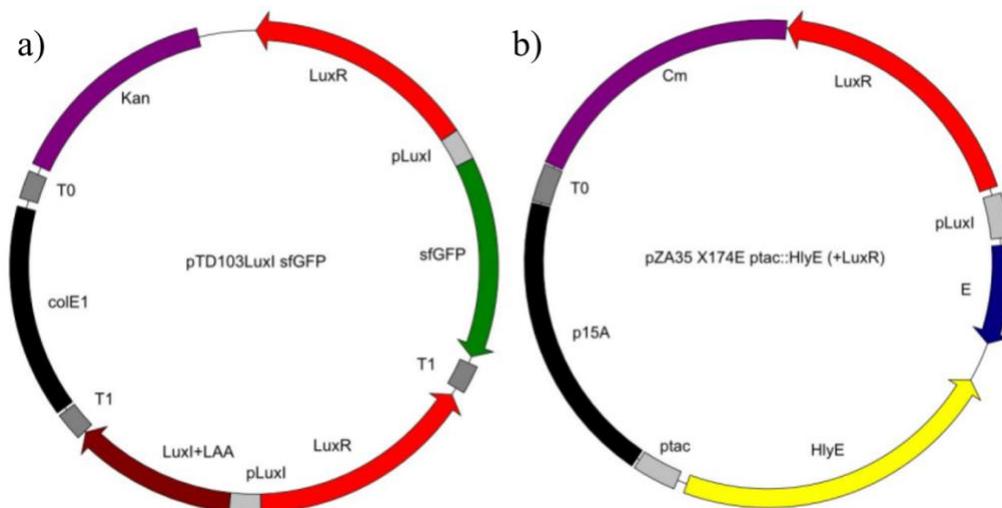
Figure A1.1: Plasmids comprising the lysis circuit[2]**.** The QSC circuit contained the pTD103 plasmid, as well as pJZ13 which simply had constitutive chloramphenicol and RFP expression. Figures were taken from the supplementary material in Din et al.

Heat-competent cell stocks were made from MG1655 7740, obtained from the CGSC. Once thawed, the cells were double transformed via heat shock to make the QSC and SLC strains. The strains received 50 ng of each of the respective plasmids and were left on ice for 15 minutes to acclimate. Heat shock was conducted at 42 °C for 30 seconds, and then recovered on ice for 2 minutes. 500 µL of SOC outgrowth medium (B9020S, NEB) was added and the cells were allowed to recover for 1 hour. Half of the media was plated on 1.5% LB agar plates with 1% glucose, 50 µg mL^{-1} kanamycin, and 34 µg mL^{-1} chloramphenicol. Colonies were picked after overnight incubation at 37 °C.

A1.2 Growth Curves

Using a Tecan microplate reader, circuit behavior was confirmed to follow expected growth curves. 2 µL of overnight culture were pipetted into 200 µL of LB containing the required antibiotics. Samples were run in triplicates and growth conditions in plain LB and LB with 1% glucose were measured. Graphs can be found below for the growth, GFP, and RFP expression.
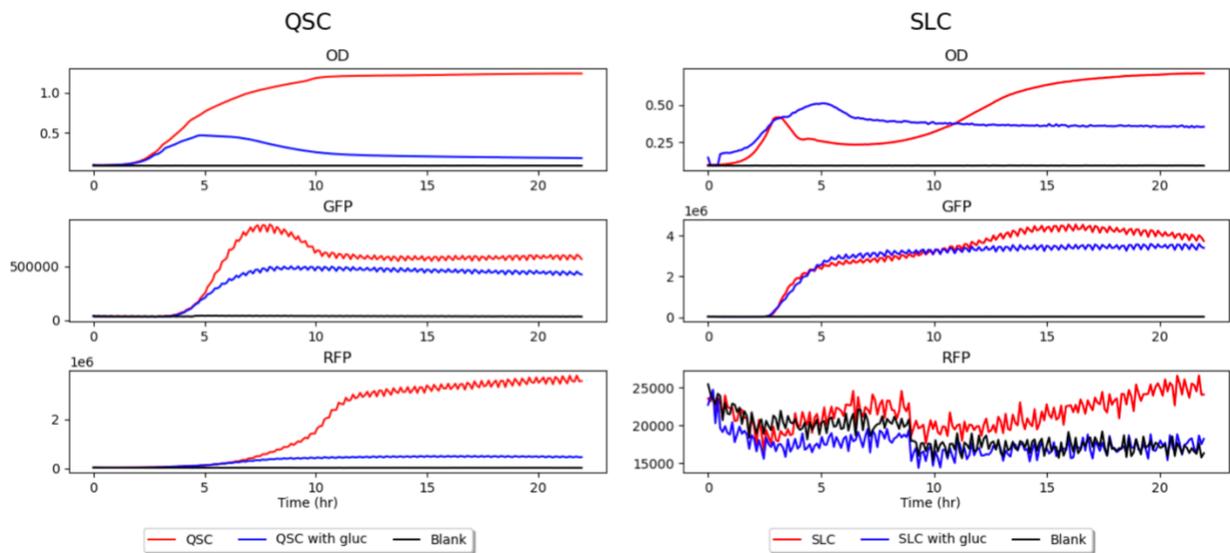
Figure A1.2: QSC and SLC growth (OD), GFP, and RFP expression curves obtained from Tecan plate reader.

Without glucose, the QSC strain grows normally, with a typical logarithmic curve. It appears with glucose there is a die-off at 0.5 OD, perhaps due to an accumulation of acetate. GFP expression peaks as quorum is reached and then decreases. With glucose, the quorum sensing appears to happen at a slower rate, with the GFP increasing at a slower rate. RFP expression follows the OD curve as it is constitutive, though it does appear to be lower with glucose, most likely due to the stress caused by acetate overflow. For the SLC strain, lysis occurs earlier in the non-glucose sample compared to the glucose sample, which is expected. GFP expression is similar in both. RFP expression is close to the blank which is expected since the SLC does not have any plasmids that produce RFP.

A1.3 Plating and Imaging

LB with 0.35% Difco® agar was used to fill petri dishes. Overnight cultures were diluted 100X and allowed to grow in a shaker at 37 °C until an OD of about 0.4-0.6 was reached, which indicates that the bacteria are in exponential phase. 2 µL of the culture was then spotted onto the plate and allowed to dry under a flame from a Bunsen burner. Plates were incubated upside down at 37 °C overnight. The next day they were imaged using the Thermofisher iBright™ FL1500 imaging system.

A1.4 Recreating Flower Shapes

A similar process as described above was used for creating the flower-like patterns. Instead of directly taking the culture and spotting on the plate, however, the culture tube was spun down and brought back up to 1.0 OD in PBS. A 10:1 ratio of QSC to SLC strain was then created and 3 µL was spotted on the 0.35% LB agar plate. The spot was dried, incubated overnight, and imaged the next day.

A1.5 MATLAB Code for Modeling Chemotaxis

Below is the MATLAB code used for developing the chemotactic model.

```
%%%%%%%%%%%%%%%%%%%%%%%% 20240506 logistic growth, 2D
chemotaxis, 2D diffusion

clc;clear;
tic;
nx = 300; dx = 0.1;
ny = nx; dy = dx;
x=nx*dx;
y=ny*dy;
Nt = 1; dt = 0.01;

PERIODIC=1;
NOFLUX=0;
ZEROBC=0;
SAVEPLOTS=1;

idx=1;
filename='chemotaxis.gif';

lambda=0.1;
k=0.2;
chi=5;
Dn = 0.05;
Dc= 0.5;
c0=0.01;
Cmax=1;
Cmin=0.5;
Nmax=0.5;
BoxLength=1;


posx = dx*(1:nx) - x/2;
posy = dy*(1:ny) - y/2;
xini=round(nx/16);
yini=round(ny/16);
n = zeros(nx,ny);

dist=zeros(nx,ny);
for i=nx/2-xini:nx/2+xini
    for j=ny/2-yini:ny/2+yini
        dist(i,j)=sqrt(posx(i)*posx(i)+posy(j)*posy(j));
    end
end
n(nx/2-xini:nx/2+xini,ny/2-yini:ny/2+yini)=cos(dist(nx/2-xini:nx/2+xini,ny/2-
yini:ny/2+yini)*pi/xini/dx/2);
n(n < 0) = 0;
```

```
c = ones(nx,ny);

n2= zeros(nx,ny);
posx = dx*(1:nx) - x/2;
posy = dy*(1:ny) - y/4;
dist=zeros(nx,ny);
xini=round(nx/16);
yini=round(ny/16);
for i=nx/2-xini:nx/2+xini
  for j=ny/4-yini:ny/4+yini
    dist(i,j)=sqrt(posx(i)*posx(i)+posy(j)*posy(j));
  end
end
n2(nx/2-xini:nx/2+xini,ny/4-yini:ny/4+yini)=cos(dist(nx/2-xini:nx/2+xini,ny/4-
yini:ny/4+yini)*pi/xini/dx/2);
n2(n2 < 0) = 0;

  f=figure(1);
    subplot 211
    nn(:,:,1)=n/Nmax;
    nn(:,:,2)=n2/Nmax;
    nn(:,:,3)=0;
    image([-(nx/2-0.5),nx/2+0.5]*BoxLength,[-(ny/2-0.5),ny/2+0.5]*BoxLength,nn);
    title('bacterial density');
    shading interp;

    subplot 212
    cc(:,:,3)=(c-Cmin)/(Cmax-Cmin);
    cc(:,:,1)=0;
    cc(:,:,2)=0;
    image([-(nx/2-0.5),nx/2+0.5]*BoxLength,[-(ny/2-0.5),ny/2+0.5]*BoxLength,cc);
    title('nutrient concentration');
    shading interp;
    drawnow

    if (SAVEPLOTS)
      frame = getframe(f);
      im = frame2im(frame);
      [imind,cm] = rgb2ind(im,256);
      if idx == 1
        imwrite(imind,cm,filename,'gif','Loopcount',inf);
        idx = idx+1;
      else
        imwrite(imind,cm,filename,'gif','WriteMode','append');
        idx = idx+1;
      end
```

```matlab
        end


    figure(2);
        subplot 211
        hold off
        plot(dy:dy:y, n(nx/2,:));
        hold on
        plot(dy:dy:y, n2(nx/2,:));
        title('n');
        %axis([0 x 0 10])
        subplot 212
        plot(dx:dx:x, c(:,ny/2));
        title('c');
        %axis([0 x 0 1.5])
        drawnow



%%%%%%%%%%%%%

if PERIODIC
kx = [0 : nx/2-1, -nx/2 : -1]'*2*pi/(nx - 2)/dx;
ky = [0 : ny/2-1, -ny/2 : -1]'*2*pi/(ny - 2)/dy;

[Fx, Fy] = meshgrid(kx, ky);
k2 = Fx.^2 + Fy.^2;
M1 = exp(-Dn * k2 * dt); %density diffusion
M2 = exp(-Dc * k2 * dt); % glucose diffusion
rekx=0.;
imkx=Fx.*exp(-0.*k2*Dn*dt);
corx=complex(rekx,imkx);
reky=0.;
imky=Fy.*exp(-0.*k2*Dn*dt);
cory=complex(reky,imky);

else
kx = [0 : 2*nx/2-1, -2*nx/2 : -1]'*2*pi/(2*nx - 2)/dx;
ky = [0 : 2*ny/2-1, -2*ny/2 : -1]'*2*pi/(2*ny - 2)/dy;

[Fx, Fy] = meshgrid(kx, ky);
k2 = Fx.^2 + Fy.^2;
M1 = exp(-Dn * k2 * dt); %density diffusion
M2 = exp(-Dc * k2 * dt); % glucose diffusion
rekx=0.;
imkx=Fx.*exp(-0.*k2*Dn*dt);
corx=complex(rekx,imkx);
```

```
reky=0.;
imky=Fy.*exp(-0.*k2*Dn*dt);
cory=complex(reky,imky);
end

count = 0;
SF=0.00001;

for t = 1:100000

  % logistic growth
  dn = lambda * n .* (1 - (n+n2)/Nmax);
  n = n + dn * dt;
  dn2 = lambda * n2 .* (1 - (n+n2)/Nmax);
  n2 = n2 + dn2 * dt;


  % nutrient consumption
  c(c<SF)=c(c<SF).*exp(-k*(n(c<SF)+n2(c<SF))*dt);
  c(c>SF)=c(c>SF)-k*(n(c>SF)+n2(c>SF))*dt;

  %diffusion and chemotaxis
  nf = fft2(n);
  cf = fft2(c);
    % x derivatives
  cre=real(cf).*real(corx)-imag(cf).*imag(corx);
  cim=real(cf).*imag(corx)+imag(cf).*real(corx);
  cfx=complex(cre,cim);
  dcdx=real(ifft2(cfx)); % dc/dx
  flux1x=n./(c+c0).*dcdx;
  ffx=fft2(flux1x);
  cre=real(ffx).*real(corx)-imag(ffx).*imag(corx);
  cim=real(ffx).*imag(corx)+imag(ffx).*real(corx);
  cfx=complex(cre,cim);
  fcx=chi*real(ifft2(cfx));
    % y derivatives
  cre=real(cf).*real(cory)-imag(cf).*imag(cory);
  cim=real(cf).*imag(cory)+imag(cf).*real(cory);
  cfy=complex(cre,cim);
  dcdy=real(ifft2(cfy)); % dc/dy
  flux1y=n./(c+c0).*dcdy;
  ffy=fft2(flux1y);
  cre=real(ffy).*real(cory)-imag(ffy).*imag(cory);
  cim=real(ffy).*imag(cory)+imag(ffy).*real(cory);
  cfy=complex(cre,cim);
  fcy=chi*real(ifft2(cfy));
```

```matlab
    n = real(ifft2(nf .* M1));
    n=n-dt*(fcx+fcy);
    c = real(ifft2(cf .* M2));
    n(n < 0.0000001) = 0;
    c(c < 0.0000001) = 0;

    if(mod(t,100)==0)   % plot the fields
      t * dt
      Nm=max(max(n));

%       count = count + 1;
%       kymo_n(count,:) = n(:,ny/2);
%       kymo_c(count,:) = c(:,ny/2);


     f=figure(1);

      subplot 211
     nn(:,:,1)=n/Nm;
     nn(:,:,2)=n2/Nm;
     nn(:,:,3)=0;
     image([-(nx/2-0.5),nx/2+0.5]*BoxLength,[-(ny/2-0.5),ny/2+0.5]*BoxLength,nn);
     title('n');
     shading interp;

      subplot 212
     cc(:,:,3)=(c-Cmin)/(Cmax-Cmin);
     cc(:,:,1)=0;
     cc(:,:,2)=0;
     image([-(nx/2-0.5),nx/2+0.5]*BoxLength,[-(ny/2-0.5),ny/2+0.5]*BoxLength,cc);
     title('c');
     shading interp;
     drawnow

    figure(2);
     subplot 211
     hold off
     plot(dy:dy:y, n(nx/2,:));
     hold on
     plot(dy:dy:y, n2(nx/2,:));
     title('n');
     xlabel('y');
%    axis([0 x 0 10])
     subplot 212
     plot(dy:dy:y, c(nx/2,:));
```

75

```
        title('c');
        xlabel('y');
   %    axis([0 x 0 1.5])
        drawnow
        %pause

        if (SAVEPLOTS)
            frame = getframe(f);
            im = frame2im(frame);
            [imind,cm] = rgb2ind(im,256);
            if idx == 1
                imwrite(imind,cm,filename,'gif','Loopcount',inf);
                idx = idx+1;
            else
                imwrite(imind,cm,filename,'gif','WriteMode','append');
                idx = idx+1;
            end
        end
    end
end

% figure(3)
% subplot 221
% pcolor(kymo_n);shading flat;colorbar
% xlabel('Posision')
% ylabel('Time')
% title('r')
% subplot 222
% pcolor(kymo_c);shading flat;colorbar
% xlabel('Posision')
% ylabel('Time')
% title('s')

toc
```

A2. Tumor DNA Sensor Methods and Assays

A2.1 Primer Design and Assembly

Primer design was conducted using Benchling and primer3plus.com to automatically pick

ideal primers given the region of interest. Since Golden Gate assembly was used to build the

constructs, the BsaI recognition cut sequence needed to be attached at the front of each primer at

the 5' end: CCCGGTCTCA. Additionally, a 4 base pair sequence was used to match up

76

fragments that needed to be ligated together. The Golden Gate assembly protocol as well as the

PCR protocol were both referenced from NEB. Below are the two constructs I designed.
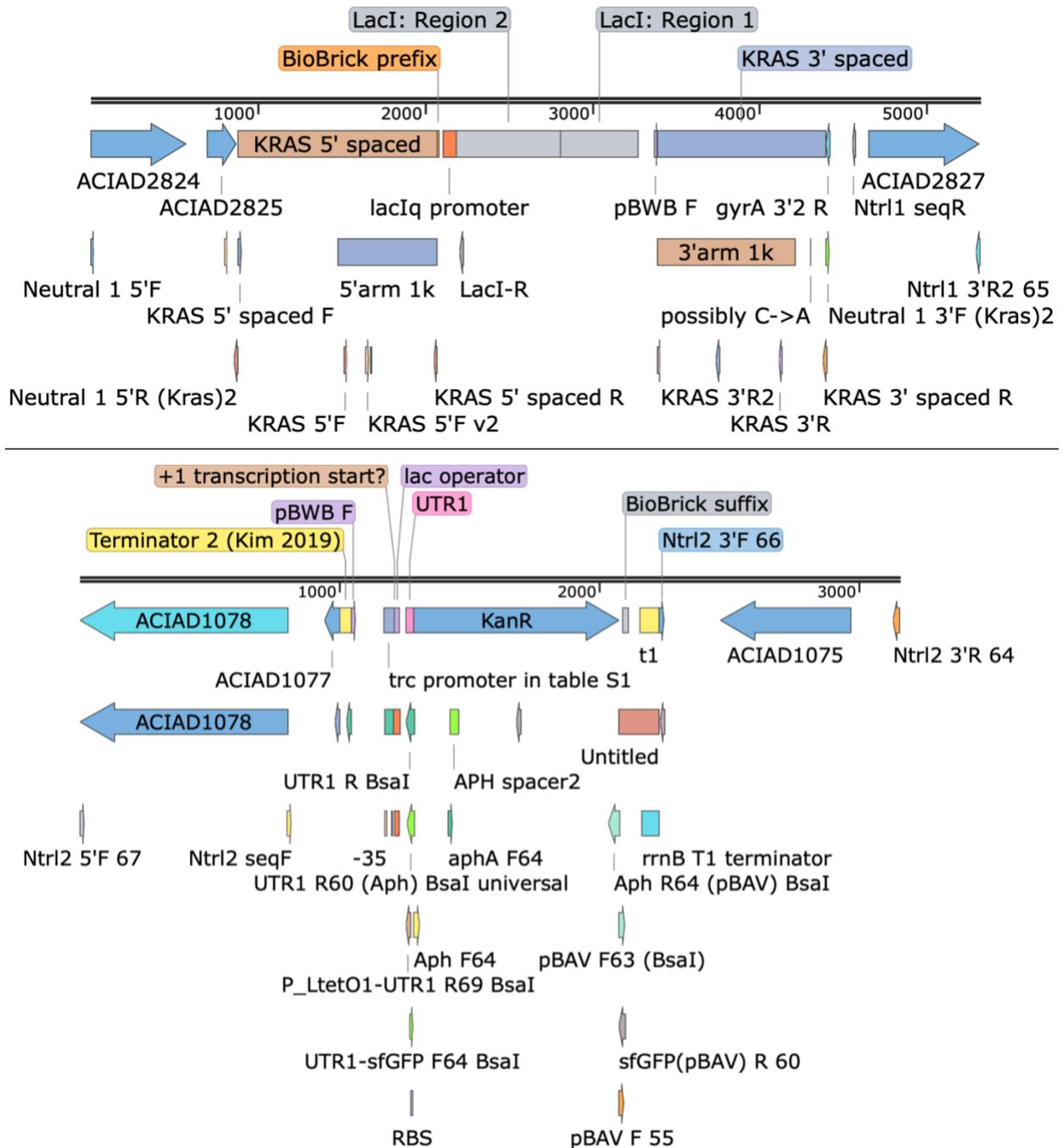
Figure A2.1: DNA constructs designed to transform into ADP1 genome. (top) N1-KRAS-LacI. (bottom) N2-pTrc-KanR.

A2.2 Genome Integration

Genomic integration to embed the DNA constructs within the N1 and N2 regions of the ADP1 strain was conducted in 2 separate ways. For the N2 insertion of the Kan sequence, 2 µL of N1 GFP *A. baylyi* taken from overnight culture was mixed with 2 µL of purified PCR product containing the N2-pTrc-KanR construct. 2 µL of that mixture was spotted onto a 2% LB agar plate without antibiotic. After about 6 hours of growth, the colony was cut out using a sterile blade and suspended into 500 µL of PBS. The solution was sonicated to separate the colony from the agar slab. 50 µL of the resulting solution was plated onto a large petri dish with 2% LB agar containing 50 µg mL^{-1} of Kan, and resulting colonies were picked and screened.

For the N1 region, a different method was used. About 1 µg of the DNA construct was accumulated from the purified PCR product. Cells containing the N2 KanR region were grown overnight and centrifuged the next day. The pellet was washed and diluted 10-fold. 1 mL of this washed cell solution was cultured in a shaking incubator with 1 µg of the DNA construct overnight and 1 mM IPTG. After about 1 hour of shaking, 50 µg mL^{-1} of Kan was added to the solution. The next day, the cells were washed and diluted 10,000-fold. 50 µL of the resulting dilution was plated onto a large petri dish with 2% LB agar containing 50 µg mL^{-1} of Kan and 1 mM IPTG, and resulting colonies were screened for GFP loss and then picked.

### A2.3 Genomic DNA Extraction

Genomic DNA was extracted using the Purelink genomic DNA minikit (Invitrogen; K18200). Following overnight resuspension in DNAse-free water, PCR amplification was conducted on regions of interest.

A2.4 Serial Dilutions

To count transformants, cells were 10-fold serially diluted 6 times, and 2 µl spots were deposited onto selective (100 µg/ml Kan) and non-selective 2% LB agar plates, with 2

measurement replicates at each dilution level, and 3 technical replicates in total. Colonies were then counted after imaging the plates using the iBright FL1500, using imageJ.
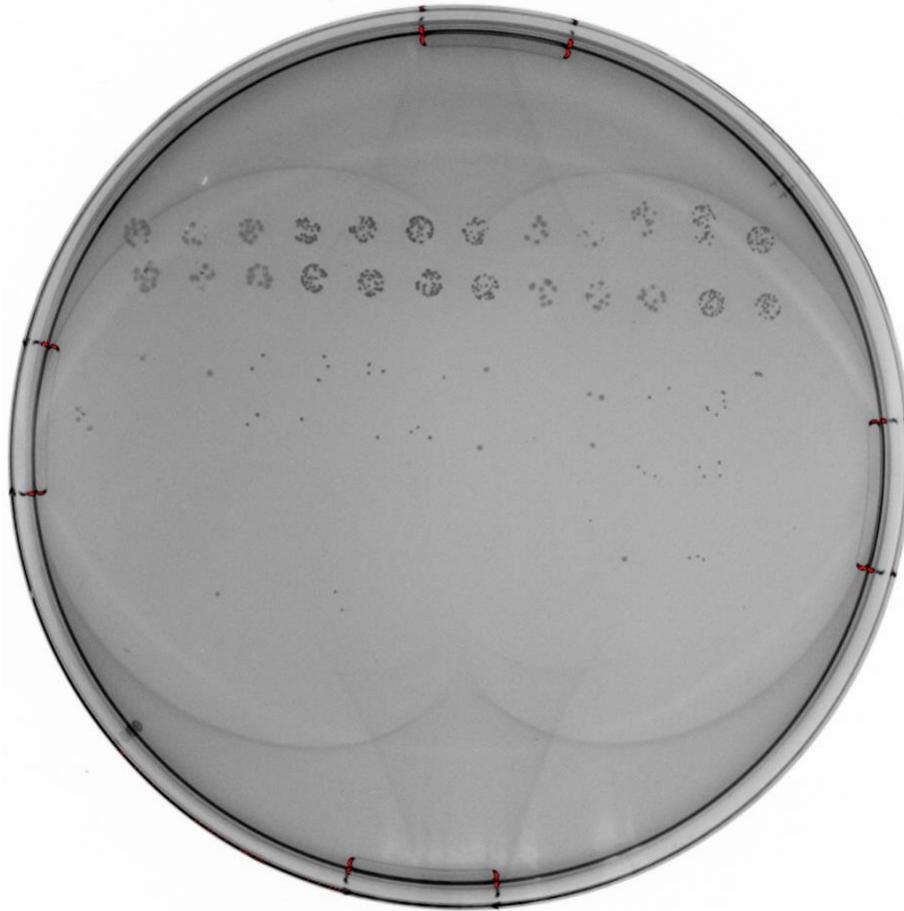
**Figure A2.2: Example plate with serial dilutions.**

Above is an example of the serial dilutions after overnight growth. This was on the LB-only plate with dilution levels 4 to 6. The colonies for levels 4 and 5 were counted to produce the results for HGT rates.

A3. ALE Methods and Assays

A3.1 Transformation of Strains

BOP27 and EcN were the strains used in this project. WT and evolved versions containing the M4 plasmid were acquired from glycerol stocks already made by the publishers of

the paper this project was based on. The M4 plasmid was simple, with *pLuxI* driving GFP, and

constitutive *specR* and *luxR* production. This allowed for AHL induction to begin GFP

production. The other two plasmids used were the pJZ50v2 plasmid with inducible lysis and

constitutive GFP as well as the pLux X174E GFP plasmid which was created using the pJZ50v2

plasmid and removing the GFP promoter, replacing with an RBS. The plasmid constructs can be

seen below.
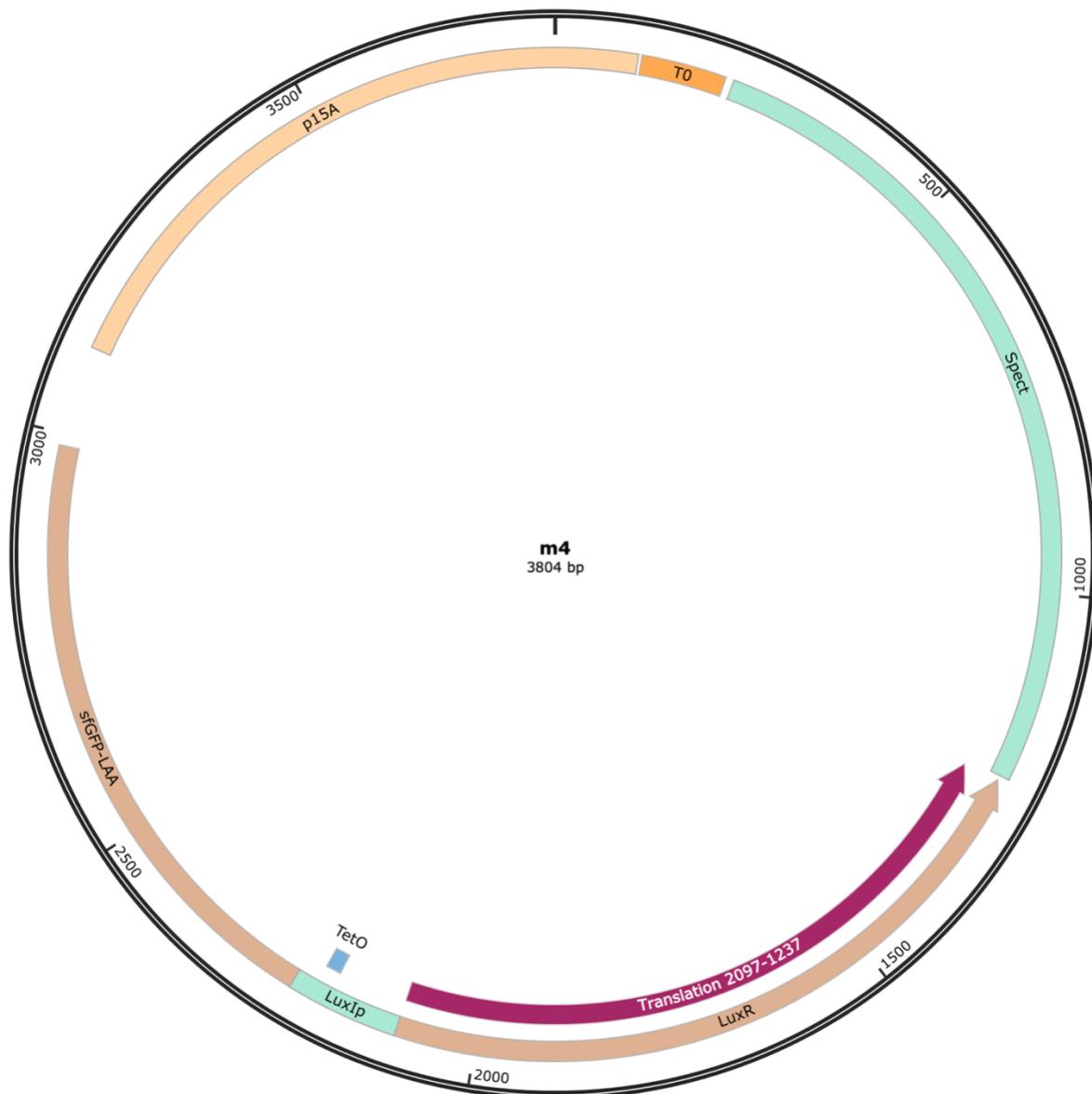
Figure A3.1: M4 plasmid construct used for MG1655 characterization in M9 lactate.
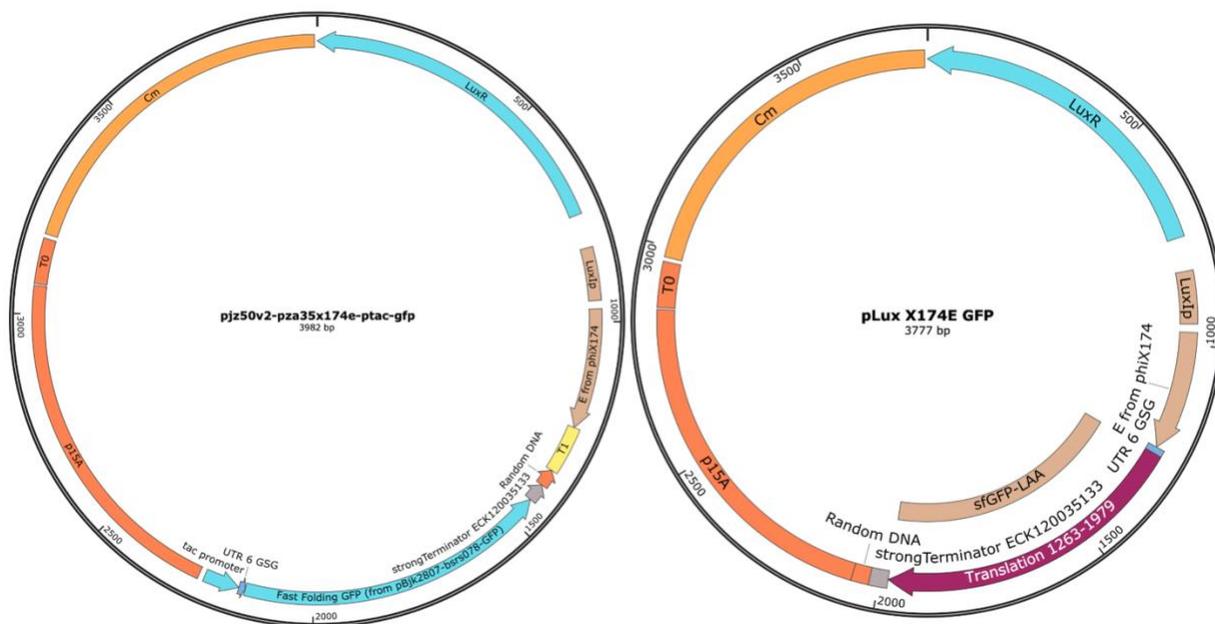
Figure A3.2: Plasmid constructs used for inducible lysis testing. (left) pJZ50v2. (right) pLux X174E GFP.

The heat shock procedure for MG1655 was the same as in Chapter 1. The strains received 50 ng of the respective plasmids and were left on ice for 15 minutes to acclimate. Heat shock was conducted at 42 °C for 30 seconds, and then recovered on ice for 2 minutes. 500 µL of SOC outgrowth medium (B9020S, NEB) was added and the cells were allowed to recover for 1 hour. Half of the media was plated on 1.5% LB agar plates with 1% glucose and 34 µg mL^{-1} chloramphenicol. Colonies were picked after overnight incubation at 37 °C.

EcN was transformed via electroporation. 1 mm electroporation cuvettes (40-103, Genesee Scientific) were used. 60 µL of the electrocompetent cells were mixed with 50 ng of plasmid. After 15 minutes on ice, the mixture was transferred to the cuvette and exposed to a 1.8 kV electric current for approximately 6 milliseconds using an electroporator (1652100, BioRad). Immediately after pulsing, 500 µL of SOC outgrowth medium (B9020S, NEB) was added and the cells were allowed to recover for 1 hour. Half of the media was plated on 1.5% LB agar

plates with 1% glucose and 34 µg mL−1 chloramphenicol. Colonies were picked after overnight incubation at 37 °C.

A3.2 Growth Curves

Using a Tecan microplate reader, circuit behavior was confirmed to follow expected growth curves. 2 µL of overnight culture, which was grown in the evolved media of M9 lactate, were pipetted into 200 µL of evolved media containing the required antibiotics. Induced samples received 10 nM of AHL. Samples were run in triplicates and growth and GFP output were measured.

For lysis induction growth curves, the inducible lysis strain and the pJZ50v2 strain were grown ON in LB with 1% glucose to inhibit lysis. 2 µL of overnight culture were added to 200 µL of LB. Additionally, GFP gain regulation was turned off, so consecutive runs could be concatenated. Once the OD of a triplicate batch hit around 0.3, the plate was taken out and induced with 10 nM of AHL. The plate was then placed back into the Tecan plate reader to continue taking measurements. Analysis was performed by concatenating the sheets of the concurrent data sets and normalizing the data by subtracting the average of the blank readings.