UNIVERSITY OF CALIFORNIA, SANTA CRUZ

THESIS

Toxic Algal Bloom Interference (TABI)

Srikar Bevara[1], Edgar Cruz[1], Blanca Davila Gil[1], Tyler Gaw[1], Ashlyn Huynh[1], Allison Jaballas[2], Varun Kodur[2], James Larbalestier[2], Daniel Li[3], Zokhira Mukhammadyunusova[1], Jordan Nichols[1,4], Nikita Rajesh[1], Julia Saltz[1], McKenna Smith[3,5], Katie Warren[1]

[1] Biomolecular Engineering and Bioinformatics
[2] Molecular, Cellular and Developmental Biology
[3] Biochemistry and Molecular Biology
[4] Neuroscience
[5] Marine Biology

September 1, 2023

Abstract

The prevalence of harmful algal blooms (HABs) is increasing globally due to a combination of factors including warming temperatures and chronic agricultural runoff. HABs threaten crucial water systems with toxic compounds that harm the environment and limit access to freshwater. The blooming cyanobacteria *Microcystis aeruginosa* (*M. aeruginosa*) secretes microcystin, a carcinogenic hepatotoxin responsible for a variety of adverse health effects and the disruption of freshwater ecosystems. Comprehensive remediation of toxic HABs can only be met by strategies based on horizontal gene transfer (HGT), and efficiently transforming *M. aeruginosa* in a complex bacterial community is difficult—in part due to the targeting of foreign DNA by restriction-modification (R-M) systems. To inhibit toxin production at the scale of a HAB, we must establish a method for the efficient, conjugative transformation of *M. aeruginosa*. To do this, we will use the Stealth program to identify underrepresented R-M motifs in the *M. aeruginosa* genome; then, we will use our Chameleon program to remove these motifs from protein-coding sequences through synonymous codon-optimization of a broad-host range, conjugative plasmid (pSPDY). Transformation efficiency of M. aeruginosa with modified and unmodified pSPDY will be quantified through natural and conjugative transformation to validate the use of Stealth and Chameleon in future works. These experiments will assess a viable method for HGT into *M. aeruginosa*, which we will leverage for toxin-disruption in the future. Ultimately, the goal of this project is to provide a foundation for efficiently engineering non-model species in complex bacterial communities.

Acknowledgements

The 2023 UCSC iGEM team, TABI, would like to thank our principal investigator Dr. David Bernick, as well as our mentor Gia Balius, for their invaluable support and guidance throughout our project. We thank Erica Jane Estamo and Lorena Vargas from Hartnell College and Clare Reyes from UCSC for their assistance as interns. We also thank the Baskin School of Engineering for generously providing us with a workspace. We acknowledge the support of the following individuals: Dr. Raphael Kudela and Kendra Negrey (University of California, Santa Cruz) for assistance in project direction and LC-MS testing of microcystin, Dr. Shaun McKinnie (University of California, Santa Cruz) for assistance in gene disruption pathways, Bryan Condy (City of Watsonville Lab Manager) for his key insights on Pinto Lake, Dr. Jen Quick-Cleveland (University of California, Santa Cruz) for ecological input on Pinto Lake, Dr. Rebecca Dubois and Dr. Dianne Hendricks (University of California, Santa Cruz) for supplying us with plates and competent cells, David Kelaita (University of Colorado, Boulder; 2021 UCSC iGEM member) for his support in project design, Dr. Glenn Millhauser and Dr. Kevin Singewald (University of California, Santa Cruz) for guidance in analyzing microcystin, Diego Gelsinger (Columbia University) for guidance on conjugation and sharing lab supplies, Brandy McNulty (University of California, Santa Cruz) for assisting our team with library preparation and genomic sequencing, Audrey Levine and Eric Baugher (Santa Cruz County) for giving guidance on the project. We would also like to express our gratitude to Promega Corporation and Integrated DNA Technologies (IDT) for providing lab materials. TABI is funded by the Student Organization Funding Advisory, John R. Lewis, Porter College, Crown College, Oakes College, the UCSC Physical & Biological Sciences Sciences Division, UCSC Baskin Engineering, the Cynthia Mathews Scholarship, and the iGEM Impact Grant. We would also like to thank everyone who donated to our team during the 2023 UC Santa Cruz Spring Give. Lastly, we are grateful to the 2023 iGEM competition committee and judges for their time and consideration.

Table of Contents

Abstract **2**
Table of Contents **4**
Section 1 ..... **1**
Introduction **1**
Section 2 ..... **4**
Discovery Work 4
2.1 Human Practices 4
2.2 Establishing TABI 6
2.2.1 Project Evolution 6
2.3 Picking a Host 6
2.3.1 *M. aeruginosa* 6
2.3.2 *E. coli* 6
Section 3 ..... **8**
Engineering 8
3.1 Sequencing 8
3.2 Plasmid Design 8
3.3 DNA Transfer 9
3.4 Results 9
3.4.1 *M. aeruginosa* UTEX 2385 Sequencing 9
3.4.2 Sequencing the *mcy* cluster 10
3.4.3 Natural transformation of *M. aeruginosa* 11
Section 4 ..... **12**
Methods 12
4.1 Genomic DNA extraction and purification of *M. aeruginosa* UTEX 2385 12
4.1.1 Long-Read Sequencing 13
4.2 Golden Gate Assembly of pSPDY 14
4.3 Conjugative Delivery of pSPDY 14
4.4 Natural Transformation of *M. aeruginosa* UTEX 2385 15
Section 5 ..... **16**
Computational Methods and Chameleon 16
5.1 Background 16
5.1.1 Stealth Analysis 16
5.2 Chameleon 16
5.2.1 Pipeline Flow 17
Section 6 ..... **18**
Discussion 18
6.1 Future Research 18
6.2 Future Impact and other applications 18

6.3 Ethics and Safety.....19
Bibliography..... 20

List of Figures

FIGURE 1.1: News Reports of Algae Blooms in the U.S. from 2010 to the present day..... 1
FIGURE 1.2: Different possible strategies to mitigate or control Microcystis blooms.....2
FIGURE 3.1: Annotation of pSPDY..... 9
FIGURE 3.2: Bandage genome visualization of metagenomic assembly.....10
FIGURE 3.3: Annotated consensus sequence of *mcy* cluster.....11
FIGURE 4.1. Visualization of sequencing process.....13
FIGURE 4.2: Gel analysis of purified microcystis DNA..... 14
FIGURE 5.1: Overview of The Chameleon Tools Pipeline.....22

Section 1

Introduction

M. aeruginosa is a species of freshwater cyanobacteria that commonly forms dense aggregates in eutrophic water bodies. At the end of its life cycle, *M. aeruginosa* cells lyse and release microcystin, a potent hepatotoxin and carcinogen [1]. The accumulation of microcystin in crops, livestock, potable water, and recreational water has been associated with detrimental health issues. Microcystin damages mammalian livers by binding to a class of enzymes known as protein phosphatases. Protein phosphatases are used in biochemical pathways to remove phosphate from a protein. The buildup of phosphorylated-proteins in the liver causes liver damage [2]. Exposure to the toxin can cause symptoms including abdominal pain, nausea, vomiting, headache, diarrhea, sore throat, pneumonia and in severe cases, liver failure and death [3].

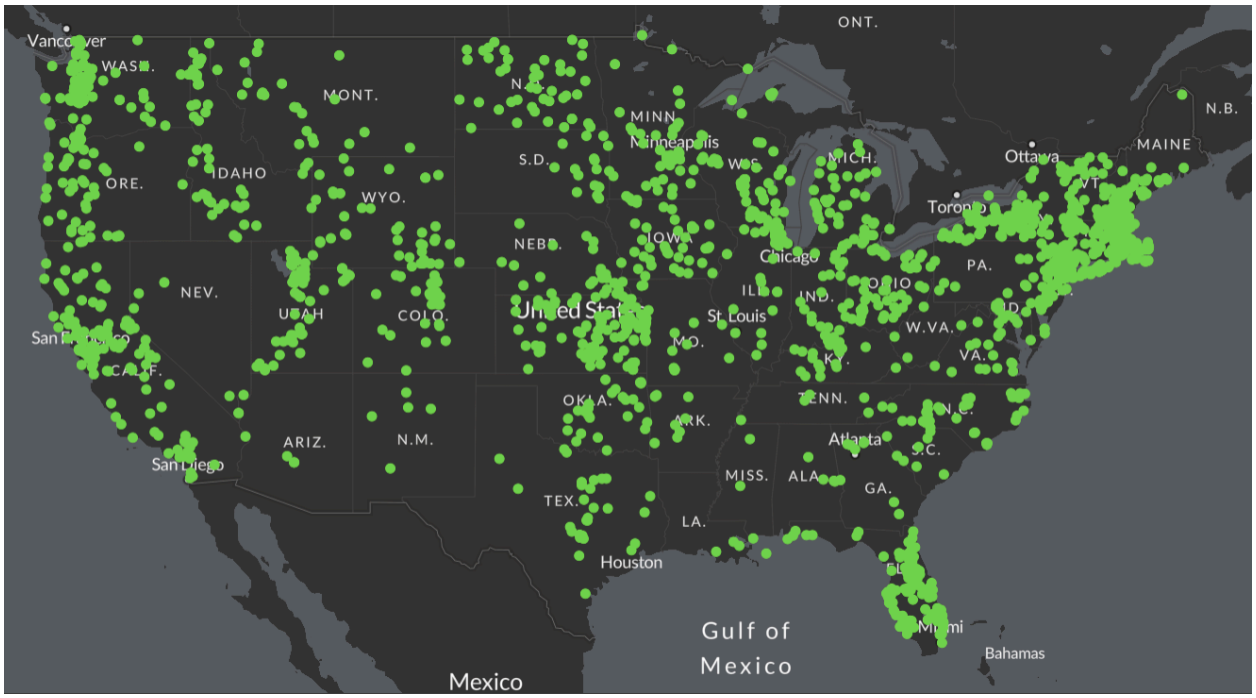

FIGURE 1.1: News Reports of Algae Blooms in the U.S. from 2010 to the present day.

In addition to its adverse health effects, microcystin exhibits allelopathic properties, disrupting local ecosystems by inhibiting growth of *M. aeruginosa*'s competitors, such as green algae [4]. This allows the cyanobacteria to dominate the phytoplankton community, threatening aquatic biodiversity.

HABs have been detected in the 48 contiguous U.S. states, and have been associated with human and animal illness in at least 43 states [4, 5]. A multitude of industries, including fishing, agriculture, and tourism are vulnerable to the impacts of HABs, resulting in annual economic losses of about $4 billion USD [6]. When microcystin-contaminated water is used for crop irrigation, these toxins can diminish nutritional quality, pollute surrounding soil, limit plant growth, and decrease production yield [1].

HABs tend to form due to factors including high surface water temperatures, extreme weather, slow water circulation, wind currents, and water currents [6, 7]. The increasing prevalence of HABs has also been linked to an accumulation of nutrients such as carbon, nitrogen, and phosphorus that often comes from

agricultural runoff and septic system leaks. The buildup of these nutrients provides the ideal habitat for *M. aeruginosa* to grow and thrive. Recent studies have identified 4,400 lakes (Fig. 1.1) in the United States that exceeded the Environmental Protection Agency (EPA) recommended recreational water quality standards for microcystin contamination of 0.8 parts per billion (ppb) [8].

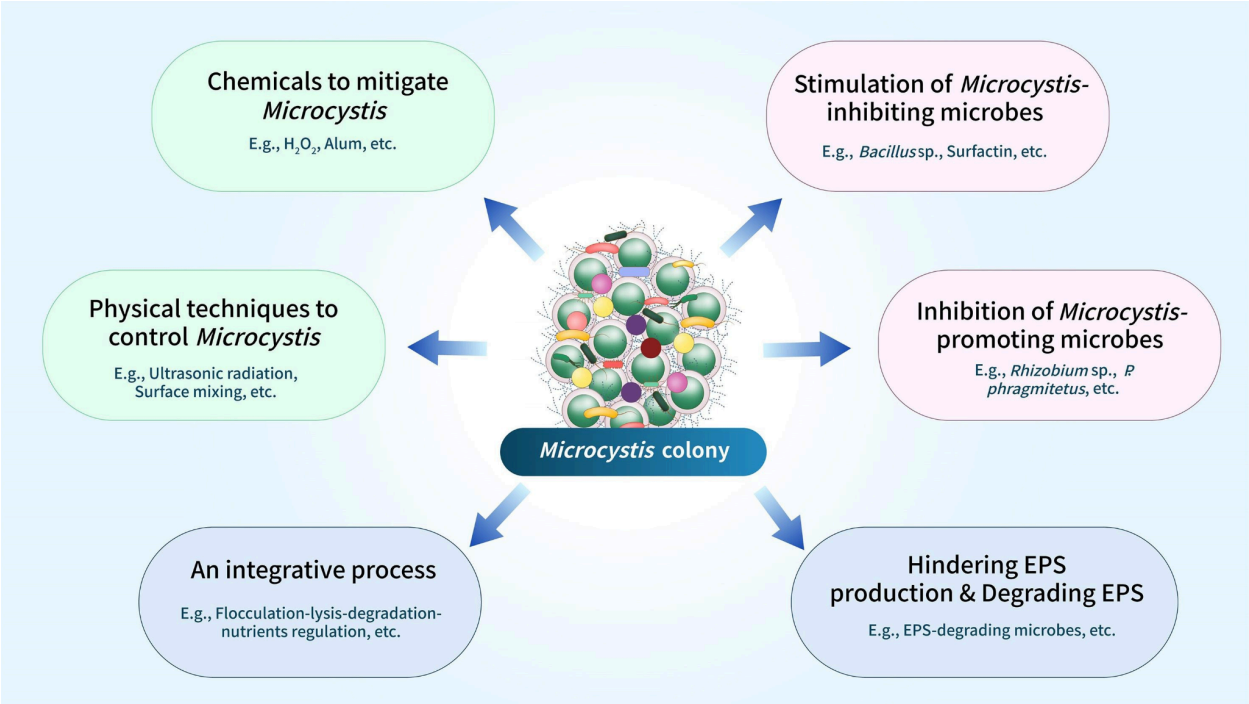

FIGURE 1.2: Different possible strategies to mitigate or control *Microcystis* blooms. Source: Adapted from [8]

Current HAB treatments can be categorized as chemical, physical, integrative processes, competitive microbial, inhibitory microbial, and extracellular polymeric substances (EPS) producing methods (Fig.1.2). Chemical compounds such as aluminum sulfate (alum) bind phosphate, limiting the surrounding nutrients that are necessary for cell growth. Additionally, chemicals like hydrogen peroxide (H_2O_2) and chitosan are employed to damage cells and promote lysis. Ultrasonication and surface mixing are physical disruptions that inhibit bloom formation. Integrative processes such as flocculation, are readily applied as water treatments when drinking water is contaminated by microcystin. An alternate control method involves adjusting microbe balances near algal blooms by introducing microcystin-degrading microbes. Degrading EPS is another HAB treatment option, which involves the use of enzymes in the extracellular space. While not thoroughly established, these methods show potential in mitigating microcystin. However, they are costly, labor-intensive, and environmentally-harmful, limiting their applications in lake restoration.

Our project is motivated by a local case of microcystin contamination in Pinto Lake, located in Watsonville, California. At its peak, microcystin levels in Pinto Lake exceeded 2,893,000 ppb, three hundred thousand times greater than the toxin exposure limit in California [9]. In 2017, Pinto Lake underwent an alum treatment to suppress *M. aeruginosa* blooms. Although initially successful, HABs have returned in recent years. Given the price of alum treatment and financial constraints faced by the City of Watsonville, an alternative treatment is needed.

Initially, we prioritized the creation of a self-sustaining toxin treatment that hijacks natural biological processes. However, with limitations on current cyanobacteria research, we instead focused our efforts on developing a method that increases gene editing efficiency of non-model organisms.

The success of gene editing is limited by the bacteria's R-M system, which degrades foreign genetic material, protecting the host from foreign nucleic acids. These systems are composed of two main parts; restriction enzymes and methyltransferases. Restriction enzymes seek out specific DNA sequences that are unprotected by the methyltransferase and destroy them. This bacterial immune response inhibits the process of altering genetic material and greatly increases the difficulties of using synthetic biology within a non-model organism. We hypothesize that it is possible to bypass a host's R-M system by optimizing the sequences of gene editing material so the patterns recognized by the native restriction enzymes are replaced with ones that will not trigger the natural immune response.

With the use of bioinformatics, we aim to standardize genomic editing of non-model, Gram-negative bacteria like *M. aeruginosa*. This streamlines the production process for cost-effective, sustainable, and propagable HAB treatment alternatives around the globe.

Ultimately, we plan to use our work to develop a high-transformation efficiency plasmid that can effectively neutralize microcystin production in *M. aeruginosa* at a population scale. This plasmid would selectively disrupt microcystin synthetase genes, thereby halting production of the toxin while maintaining cell viability. This approach would curb HAB toxicity with minimal ecological disruption. Given that *M. aeruginosa* is naturally competent, we are also investigating the possibility of delivering naked plasmid DNA to be taken up directly by *M. aeruginosa* cells themselves; this route would enable a treatment without the use of live engineered organisms, though there may be concerns depending on the efficiency of natural transformation (which we are still assessing) and the viability of plasmid DNA in the extracellular environment.

Section 2

Discovery Work

2.1 Human Practices

Project TABI developed through the team's shared desire to address public health crises with synthetic biology. Throughout our discovery work, we prioritized projects that offered practical and effective solutions to real problems facing under-served communities. We considered many projects in the scope of public health; our focus had to adapt with respect to our timeline and through the development of our human practices.

Our discovery work converged on a project in environmental remediation. As a generation raised in the shadow of global warming, we understand that the effects of environmental crises are extensive, and that they disproportionately impact under-equipped communities. We wanted to model a project that demonstrated the potential for synthetic biology to effectively address environmental crises.

With this, we first considered remediation of one of the most prevalent toxins in marine ecosystems: domoic acid. Domoic acid is a potent neurotoxin released by *Pseudo-nitzschia australis* during HABs; the toxin seasonally contaminates marine ecosystems through accumulation in shellfish and sardines. Organisms that ingest contaminated fish are poisoned; the environment and economy suffer when toxic HABs persist. We recognized that the frequency of toxic HABs would only increase with global warming, so we wanted to address this with synthetic biology.

TABI began with a meeting with Professor Raphael Kudela at UCSC, who specializes in phytoplankton ecology. He referred us to multiple researchers that would have insight into how toxic HABs impact marine ecology. This meeting was also when we were introduced to *Microcystis aeruginosa* and microcystin contamination in Pinto Lake.

With Dr. Kudela's referral, we spoke to Monica Thrukall, a UCSD graduate student in the Allen Lab studying *Pseudo-nitzschia australis*. She informed us that domoic acid remediation was unfeasible because *Pseudo-nitzschia australis* blooms were too distributed, and techniques for engineering this species were under-developed. This is the first time we felt frustrated by the lack of generalized techniques for engineering non-model species. We began considering generalized solutions for this limitation, and we shifted our remediation efforts toward a more local environment: microcystin contamination in Pinto Lake.

We had the privilege of meeting with Bryan Condy, the laboratory manager for the City of Watsonville. We discussed issues faced by Pinto Lake, noting that there has been an increase in toxic HABs since the 1980s. Bryan informed us that the warming climate, chronic agricultural runoff, and excessive phosphorus levels from sediment at the bottom of Pinto Lake are the main cause of HABs. Representing a low-income community, Bryan stressed the need to preserve free, safe, and accessible outdoor opportunities in the under-served community of Watsonville.

Our effort towards microcystin remediation directed us to a meeting with Professor Shaun McKinnie of UCSC who specializes in biologically-significant natural products and their related enzymes. Dr. McKinnie introduced our team to the *mcy* gene cluster that synthesizes microcystin, and recommended research into degradation genes within the genomes of *M. aeruginosa* and its competitors. We decided against the expression of microcystin-degrading enzymes in *M. aeruginosa* because we wanted to address the source of the problem: microcystin production.

We decided to pursue targeted mutagenesis of a gene within the *mcy* cluster to selectively disrupt *M. aeruginosa* toxicity while maintaining cell viability. We believed this approach would work to remediate microcystin toxicity in Pinto Lake with minimal ecological disruption. We consulted with several researchers to help draft our protocol. Notably, Dr. Glenn Millhauser and Dr. Kevin Singewald of UCSC provided us with guidance about microcystin analysis, and Dr. Diego Gelsinger of Columbia University aided us in prokaryotic genome engineering.

With this guidance, we designed a plasmid capable of targeted mutagenesis of a gene within the *mcy* cluster. We were excited by the support we had with regards to genomic integration, but we recognized that this approach would only be viable if the plasmid was capable of being delivered by HGT.

Our objective shifted toward developing efficient methods for HGT into *M. aeruginosa*. Dr. Gelsinger supplied us with a conjugative strain of *E. coli* (EcGT2) that could conjugate with other species without requiring a helper plasmid. He informed us that conjugation efficiency was low in general, and that the process was difficult to optimize in the lab—much more so in the environment. It became clear that a generalized, bioinformatic approach aimed at improving transformation efficiency in non-model species was the most impactful contribution that TABI could provide to the iGEM community.

Our principal investigator, Dr. Bernick, introduced us to Stealth (section 5.1.1), a cutting-edge bioinformatic program written to enhance genetic engineering in non-model bacterial species like *M. aeruginosa.* Stealth facilitates an advanced statistical analysis of a specific organism's genome to identify underrepresented motifs that may have faced negative-selective pressure by targeting via endogenous restriction modification (R-M) systems. Resulting data allows for evasion of the host's R-M system by virtue of pattern avoidance.

The *M. aeruginosa* strain we are working with, UTEX 2385, is a non-model organism. To use Stealth, we needed to sequence UTEX 2385's genome. We completed the first round of sequencing using the Oxford Nanopore, but the genome was partially sequenced and the sample included DNA from organisms besides *M aeruginosa*. We then reached out to Brandy McNulty, a Nanopore Production Specialist at the UC Santa Cruz Genomics Institute. She helped us troubleshoot our sequencing protocol and suggested that the lack of *M. aeruginosa* genomic information may be due to poor library preparation. She is now supervising our current sequencing attempt to improve our results.

Future plans include meeting with the Resource Conservation District regarding restoration practices of Pinto Lake. The City of Watsonville received grants to treat the lake, but most of it was used for the short-term alum treatment performed by HAB Aquatics Solutions in 2017 [10]. We are hoping to get insight into their interest in our project and the possibility of releasing it into Pinto Lake. We also want to get the Watsonville community more involved, specifically through dialogue with our project. Currently, we have plans to present both the problem facing Pinto Lake and our proposed solutions to a group of high schoolers interning at the Watsonville Wetland Watch.

2.2 Establishing TABI

2.2.1 Project Evolution

Initially, we investigated several different options for our project; these included: controlling HABs with quorum sensing molecules, chitosan flocculation, phage-mediated delivery of a microcystin-degrading gene cluster, and using a conjugative plasmid to disrupt the *mcy* cluster. The three former options presented various issues with practical use; we decided to pursue the latter option of disrupting the *mcy* cluster with a conjugative plasmid, we recognized that foundational work in improving transformation efficiency and demonstrating conjugation into *M. aeruginosa* would be necessary first. Therefore, our project focuses on improving transformation efficiency in *M. aeruginosa* and demonstrating RP4 conjugation from *E. coli* to *M. aeruginosa*. We decided against flocculation of microcystin with chitosan because of the prohibitive cost of deploying this solution in a large water system despite chitosan's status as a waste product [11]. We decided against controlling HABs with quorum sensing molecules because the only well-documented signaling molecules used by *M. aeruginosa* we could find were acyl-homoserine lactones (AHLs), which promote rather than inhibit bloom formation [12]. We decided against phage-mediated delivery of degradative enzymes because of the lack of established, efficient engineering techniques for working with cyanobacteria-infecting bacteriophages [13].

2.3 Picking a Host

2.3.1 *M. aeruginosa*

Techniques for the engineering of cyanobacteria, including *M. aeruginosa*, are poorly-developed. Despite this challenge, it was necessary to work with *M. aeruginosa* to establish the techniques we would ultimately require to design a conjugative plasmid that halts microcystin production. The only strain of *M. aeruginosa* with a complete, published genome at this time is the PCC 7806 strain, we opted to use the UTEX 2385 strain due to its availability, expression of the *mcy* cluster, and opportunity to contribute a completely sequenced and annotated genome [14].

2.3.2 *E. coli*

We opted to use *E. coli* as the initial host organism for the replication and assembly of our plasmid before transformation (by natural transformation and conjugation) into *M. aeruginosa*. With a 20 minute cell cycle in optimal conditions [15], *E. coli* is a far more efficient vector for the replication and assembly of our conjugative plasmid than *M. aeruginosa*, which is limited by a 4 day cell cycle [13]. *E. coli* is also a well-documented model species, with a type IV conjugation system [17] that is better understood than the endogenous conjugation system in *M. aeruginosa* [18]. Additionally, given the conditions and reagents required to culture *M. aeruginosa* and its status as a Biosafety Level 2 (BSL-2) organism owing to its production of microcystin, we sought to minimize our work with *M. aeruginosa* by working with *E. coli*. The specific strains of *E. coli* we are working with for this project are TOP10, BL21 (DE3), and EcGT2. The former two strains are being used to copy the plasmid and validate expression. The EcGT2 strain is being used to conjugate the plasmid from *E. coli* to *M. aeruginosa* using native conjugation machinery on EcGT2 to avoid having to use a helper plasmid and verify that RP4 conjugation works in *M. aeruginosa*. Since the EcGT2 strain is auxotrophic for diaminopimelic acid (DAP), we also plan to use it to deliver the final plasmid that would be deployed in a field trial after iGEM due to the reduced likelihood of escape.

Section 3

Engineering

We aim to improve transformation efficiency in *M. aeruginosa* and demonstrate the functionality of EcGT2 and the RP4 system in conjugating from *E. coli* to *M. aeruginosa*. We will improve transformation efficiency by removing putative R-M motifs from our plasmid using the Stealth program. The improvement in transformation efficiency will be quantified by comparing the generation of transformants between a version of the plasmid that has not been modified by Stealth and another version of the plasmid that has been modified by Stealth following natural transformation. Conjugation from *E. coli* to *M. aeruginosa* using EcGT2 and the RP4 system will be demonstrated via co-culturing of transformed EcGT2 cells with untransformed *M. aeruginosa* cells followed by selection of *M. aeruginosa* transformants. This will lay the foundations for a future plasmid that disrupts toxin production and spreads across a population of *M. aeruginosa* via conjugation with high transformation efficiency.

3.1 Sequencing

In order to identify restriction-modification motifs that need to be removed from our plasmid, we sequenced the genome of the UTEX 2385 strain of *M. aeruginosa*. While several coding sequences of this strain have been published, a complete genome is not available. To assemble a full genome, we performed Nanopore sequencing on our extracted UTEX 2385 DNA. For results see Section 3.3.1 and for detailed methods, see Section 4.

3.2 Plasmid Design

Our plasmid, pSPDY, combines the broad-host range components of pSHDY with the conjugative components of pSPIN to yield a single plasmid that should be capable of sustained expression and conjugation into *M. aeruginosa*. Components from pSHDY include host-independent replication machinery, an origin of replication, and a chloramphenicol resistance gene. Components from pSPIN include an RP4 mobilization gene and origin of transfer (oriT); these parts are essential for the conjugative delivery of pSPDY by EcGT2 *E. coli*, a strain engineered with a genetically-encoded RP4 conjugation system.

Two pSPDY constructs will be synthesized with Golden Gate assembly [19] to assess changes in transformation efficiency when Stealth-identified R-M motifs are removed. One construct will be unmodified, retaining Stealth-identified R-M motifs targeted by *M. aeruginosa* UTEX 2385's restriction-modification systems. The other construct will be modified by the Chameleon program that our team developed. Chameleon removes Stealth-identified R-M motifs from protein-coding sequences by synonymous codon-optimization. Each of the two plasmids will then be transformed into *M. aeruginosa* by natural and conjugative transformation. Transformation efficiency of each construct will be quantified by CFU/ug plasmid DNA transformed. To avoid confounding variables and to reduce the number of fragments required for assembly, both pSPDY constructs will include the same Stealth-modified pSPIN-derived insert containing the eGFP gene, RP4 mobilization gene and associated oriT.

The pSPDY plasmid includes an eGFP gene to report activity of differential ribosome binding sites (RBS). We will synthesize two modified pSPDY constructs with different RBS associated with the eGFP gene: one will include an *E. coli*-derived RBS, and the other will include a putative *M. aeruginosa* RBS. The putative *M. aeruginosa* RBS was identified bioinformatically from nucleotide sequences directly upstream of the *M. aeruginosa* PCC 7806 ribosome genes, and verified by the 16S rRNA anti-RBS

sequence. Thus, expression of eGFP will allow for validation of the putative *M. aeruginosa* RBS. With this information, we will be able to contribute a potentially-useful RBS for cyanobacteria engineering to the iGEM community.

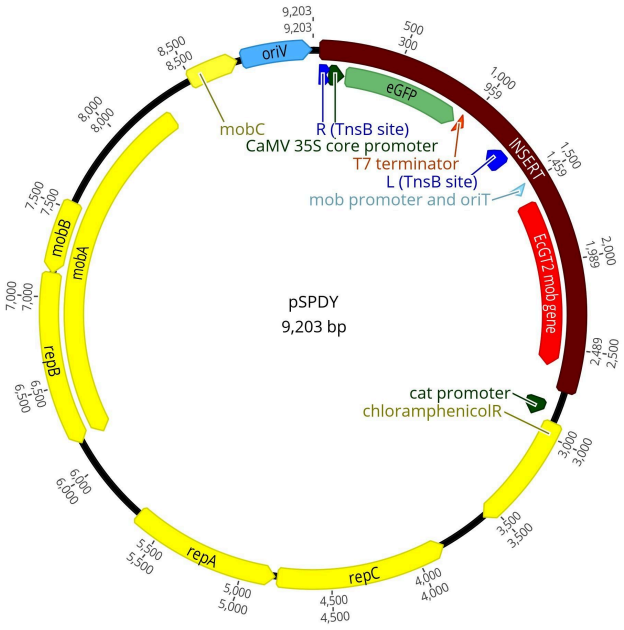

FIGURE 3.1: Annotation of pSPDY. Sequences in yellow indicate ORFs taken from pSHDY. Mobilization genes from pSHDY are unused. Insert represents sequences added to the pSHDY backbone via Golden Gate assembly. The TnsB sites are unused, but would be used in a future version of the plasmid using the enclosed region for an integration payload. The EcGT2 mob gene in red is taken from pSPIN.

3.3 DNA Transfer

Conjugation is a biological process by which bacteria share DNA, commonly occurring through plasmids. The donor cells, already containing the plasmid, synthesize a tubular appendage called a sex pilus. When in contact with the recipient cell, the sex pilus facilitates the formation of a conjugation bridge, a temporary channel connecting the two cells. The plasmid is transferred through the conjugation bridge with the help of specialized proteins.

To enable conjugation from *E. coli* to *M. aeruginosa*, we will be using the EcGT2 strain of *E. coli* as well as the oriV and RP4 mobilization gene taken from pSPIN. The EcGT2 strain supplies the RP4 conjugation machinery as it encodes it genomically. Thus, a helper plasmid is not required to deliver the pSPDY plasmid from EcGT2 *E. coli* to *M. aeruginosa* by conjugation.

3.4 Results

3.4.1 *M. aeruginosa* UTEX 2385 Sequencing

Nanopore sequencing was performed twice using the same DNA library. The first and second runs produced 108.91K and 204.51K reads, respectively. Reads from both runs were combined and assembled with Flye, resulting in 213 contigs.

From the assembled contigs, RNAmmer [20] was used to extract genes that code for 16s ribosomal RNA. Using RNAmmer, we identified eleven contigs that contained ribosomal RNA genes. These contigs were then compared to sequences in the rRNA/ITS databases using NCBI BLASTN. Our metagenomic analysis revealed hits to at least 5 possible genera present in our sample: *Microcystis, Gemmatimonas, Blastomonas, Hydrogenophaga and Stenotrophobacter.* Three contigs suggested hits to *M. aeruginosa.*

Assembled contigs were also visualized with Bandage [21] (Figure 3.2). The average sequencing depth across reads believed to be *M. aeruginosa* ranged from approximately 30X to 40X. The Bandage assembly graph indicated two well-defined clusters. Combined with the data gathered from our BLASTN analysis, we believe these clusters to represent *M. aeruginosa* and *Stenotrophobacter namibiensis*.

The two largest *M. aeruginosa* contigs were approximately 2.9 Mb and 1 Mb, which accounts for at least 60% of the expected genome. Gaps in the assembly, shown as breaks in the Bandage graph, could represent complexities including repetitive regions or regions with low coverage.

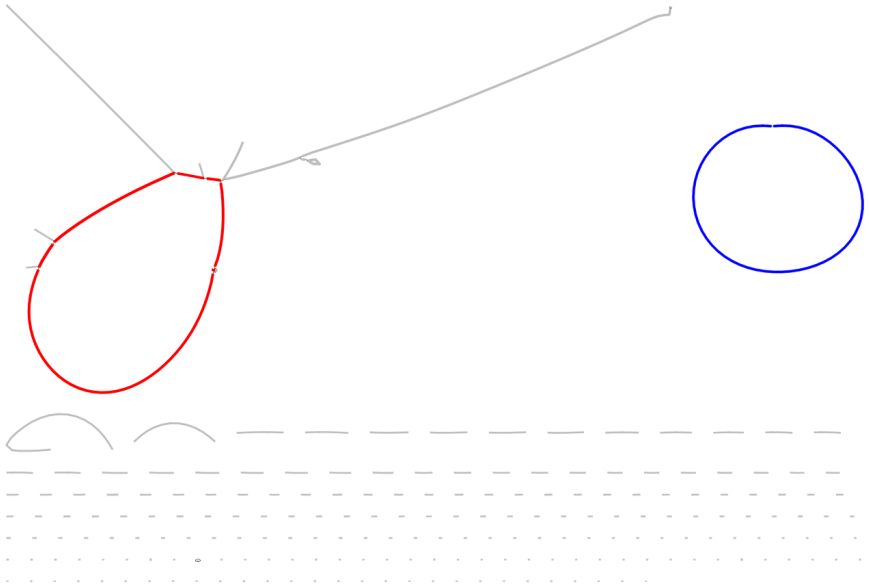

FIGURE 3.2: Bandage genome visualization of metagenomic assembly. Assembled genomes of *M. aeruginosa* shown in red and *Stenotrophobacter namibiensis* in blue.

3.4.2 Sequencing the *mcy* cluster

To enable future work on eliminating microcystin production in *M. aeruginosa* UTEX 2385, we have sequenced a section of this strain's *mcy* cluster to design guide RNAs. The sequenced region (1,131 BP) includes part of the mcyD ORF, the bidirectional promoter region between mcyD and mcyA, and part of the mcyA ORF. There are no previously-published sequences of this region of the UTEX 2385 genome available. This region would be useful for designing guide RNAs because there are multiple parts of this region for which inserting an early transcriptional terminator payload would halt the expression of multiple genes critical for microcystin production. To sequence this region of the *mcy* cluster, we designed

a pair of primers spanning the sequencing region, prepared amplicons from extracted genomic DNA, and submitted the primers & amplicons to the UC Berkeley DNA Sequencing Facility for Sanger sequencing. Using the returned Sanger reads, a consensus sequence was assembled. The mcyD, bidirectional promoter region, and mcyA sequences were annotated using known versions from the *M. aeruginosa* PCC 7806 genome. Although not used within the scope of our iGEM project, this sequenced region of the *mcy* cluster will be useful for us in the future when we have a developed platform for transformation and conjugation of *M. aeruginosa*.

FIGURE 3.3: Annotated consensus sequence of a 1,131 BP region covering part of the mcyD gene, the bidirectional promoter region, and part of the mcyA gene of *M. aeruginosa* UTEX 2385. Annotations were made based on alignment to sequences in the published *M. aeruginosa* PCC 7806 *mcy* cluster. Image rendered in Geneious Prime.

3.4.3 Natural transformation of *M. aeruginosa*

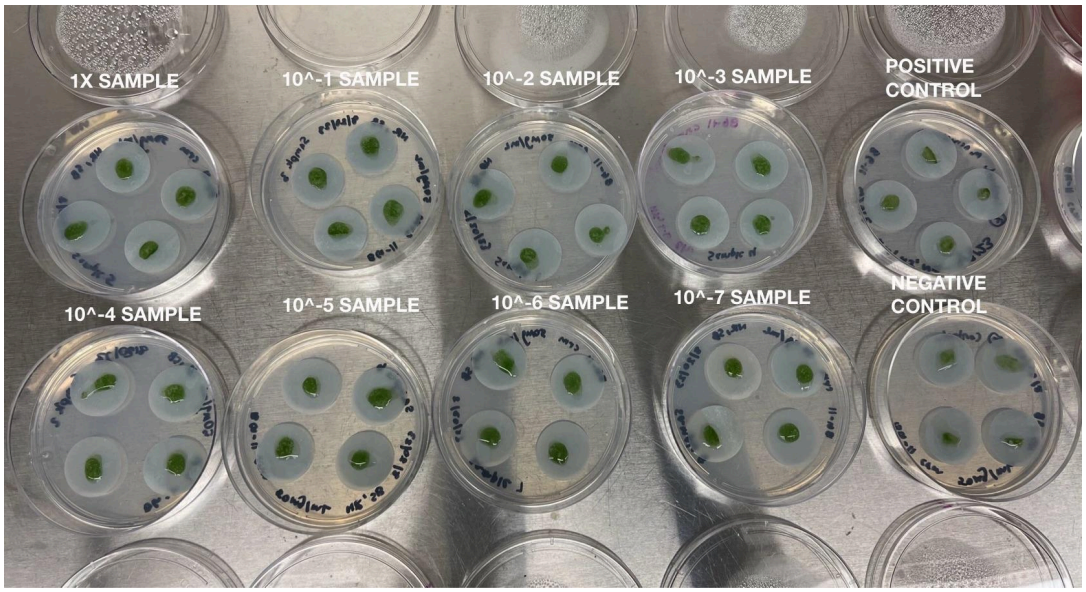

FIGURE 3.4: Serial dilutions and controls from the transformation of the pSHDY plasmid with *M. aeruginosa*

The ability of *M. aeruginosa* to readily take up foreign DNA has been established in previous work with *M. aeruginosa* PCC 7806 [22]. We wanted to confirm the natural competency of the UTEX 2385 strain with pSHDY as well as determine the amount of plasmid needed for successful DNA uptake. The pSHDY plasmid consists of three antibiotic resistance factors: chloramphenicol, kanamycin and streptomycin. We selected for successful transformation of pSHDY with chloramphenicol. Serial dilutions were performed to determine the minimum amount of plasmid required for successful natural transformation. With 8 samples with different plasmid concentrations, we observed successful growth on chloramphenicol BG-11 plates in all 8 samples. These results validated the natural transformation of *M. aeruginosa* UTEX 2385. A review of the samples under the microscope revealed the cell count of *M. aeruginosa* increased as more plasmid was added. The next iteration of natural transformations will include another dilution series of the plasmid. With this step, we aim to confidently quantify transformation efficiency of the UTEX 2385 strain of *M. aeruginosa.*

Pending further quantified assessment of the efficiency of our natural transformations, the doubling time of *M. aeruginosa* UTEX 2385, and the observed number of transformants, there may be evidence of conjugation between *M. aeruginosa* UTEX 2385 cells in our culture. Conjugation has not been established in this strain (nor other strains of *M. aeruginosa*). If conjugation is possible between *M. aeruginosa* cells, it may be possible to leverage their conjugation system to spread our plasmid horizontally beyond the transient *E. coli* to *M. aeruginosa* conjugation. Additionally, if this type of conjugation is possible, we would have greater safety considerations given that the anti-escape function of using DAP auxotrophs to deliver the plasmid would be negated.

Section 4

Methods

4.1 Genomic DNA extraction and purification of *M. aeruginosa* UTEX 2385

The protocol was adapted from the NEB Monarch HMW DNA Extraction Kit for Tissue [23].

Microbial lysis was performed on 3.125 mL samples taken from a large liquid culture of M. aeruginosa to prepare the sample for DNA extraction. The samples were swirled and then centrifuged at 16,000 rcf for 1 minute to form a pellet. The resulting supernatant was removed and discarded. Next, 40 µL of PBS was added to the pellet, and it was resuspended and vortexed. Following that, 1.8 mL of premixed NEB Monarch HMW gDNA Tissue Lysis Buffer and 60 µL of proteinase K were added. The mixture was gently mixed ten times using a 1 mL wide-bore pipette, and the sample was incubated at 56°C in a heat block for 10 minutes. Then, 15 µL of RNase A was added to the incubated sample, which was gently mixed again and incubated at 56°C in a heat block for 10 minutes on a thermomixer set at 650 rpm.

Once the microbial lysis was complete, 900 µL of Protein Separation Solution was added and mixed using a Hula Mixer for 10 minutes, rotating at 10 rpm. The sample was then centrifuged at 16,000 rcf for 10 minutes. The colorless, transparent upper phase, which contains the DNA, was aspirated from the centrifuged sample. Next, three DNA-collecting glass beads and 2.5 mL of 100% isopropanol were added to the aspirated phase and mixed using a Hula Mixer for 20 minutes, rotating at 10 rpm. Subsequently, the mixed sample rested for 1 minute before aspirating the isopropanol from the tube. Then, 2 mL of Wash Buffer was added and subsequently removed by pipetting. The wash step was repeated a second time before the beads were placed in the bead retainer and spun for 1 second to remove all residual Wash Buffer. The beads were then immediately transferred from the bead retainer to 200 µL of Extraction EB and incubated overnight at room temperature. The Extraction EB was separated from the beads by pouring it into a new bead retainer with an Eppendorf tube beneath, and it was spun at 1000 rcf for 1 minute. The sample was then gently mixed ten times to prevent heterogeneity. Extracted DNA was validated using Qubit results. All DNA was subsequently stored at 4°C.

4.1.1 Long-Read Sequencing

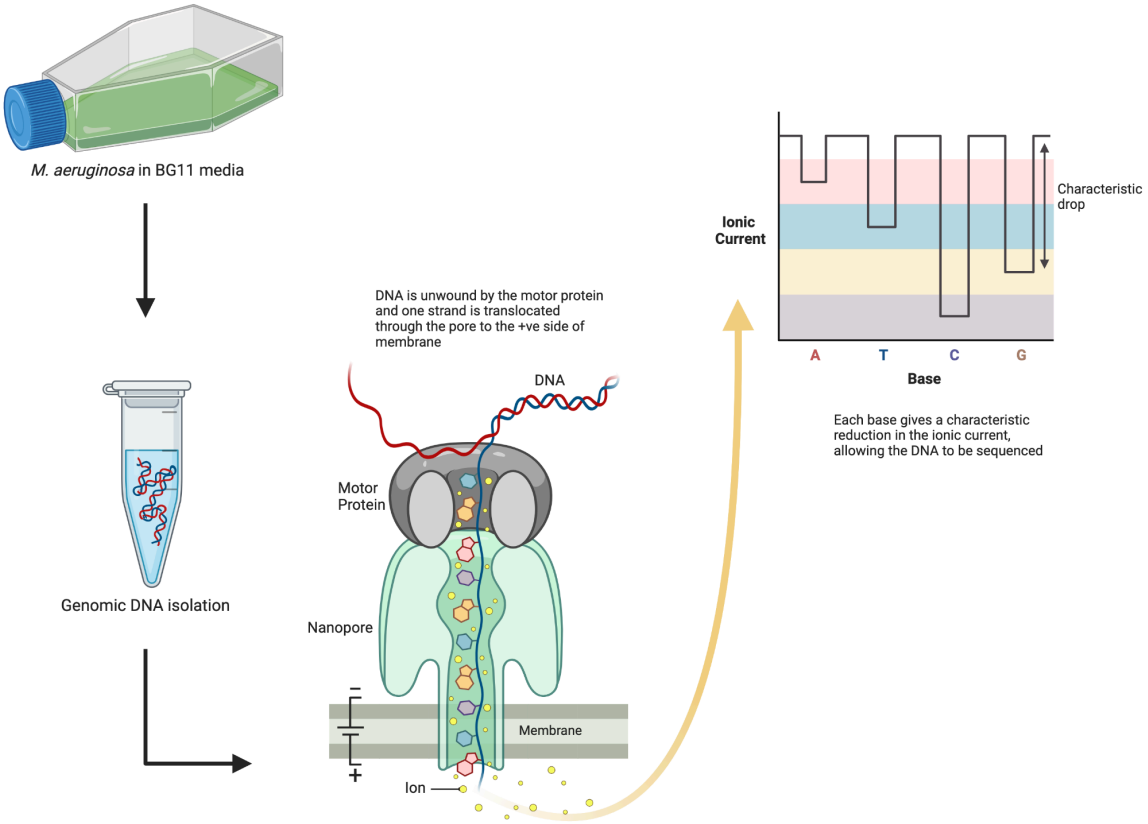

FIGURE 4.1. Visualization of sequencing process. Figure made on BioRender. Nanopore template provided by BioRender.

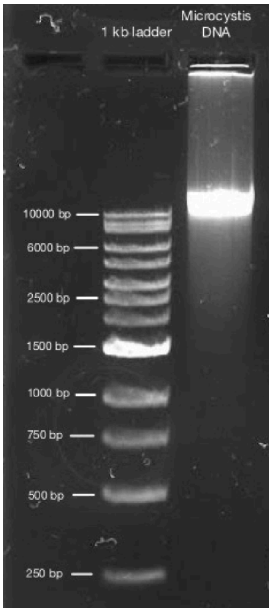

FIGURE 4.2: A gel analysis of purified microcystis DNA confirmed the presence of fragments larger than 10 kb. This suggests sufficiently long DNA sequences for Nanopore MinION sequencing were obtained.

M. aeruginosa UTEX 2385 was sequenced using the Oxford Nanopore MinION™ device (Fig. 4.1). Library preparation was performed on the isolated genomic DNA using the ONT Ligation Sequencing Kit SQK-LSK112. The prepared library was loaded onto an R9 flow cell and run on a 72-hour cycle with a wash and reload step performed after 24 hours to improve pore efficiency. The minimum read length was set to 1000 bp. Reads were basecalled with Dorado 0.3.2 and assembled with Flye. Sequencing and basecalling were conducted twice with the same DNA library, and sequences merged into a single file for assembly. Visualization of the assembly was achieved with Bandage 0.9.0.

4.2 Golden Gate Assembly of pSPDY.

The plasmid will be assembled according to the protocol outlined by a previous iGEM team. The following will be combined to create a 20 uL reaction mixture: 0.5 µL of T4 DNA Ligase, 2 µL of 10X T4 DNA Ligase Buffer, 0.5 µL of Type IIS restriction enzyme, 100 ng of vector with an equimolar amount of inserts, and MilliQ filled up to 20 µL. The reaction mixture will undergo the following cycle parameters: 15-20 cycles at 37°C for 5 minutes, 15-20 cycles at 16°C for 5 minutes, enzyme inactivation at 55°C for 15 minutes, ligase inactivation at 85°C for 20 minutes, and held at 4°C. The resulting reaction product will be transformed into TOP10 *E. Coli* for replication, extracted with miniprep, and verified with gel electrophoresis to determine the successful plasmid assembly.

4.3 Conjugative Delivery of pSPDY

The following conjugation reaction was performed between an EcGT2 strain of *E. coli* and *M. aeruginosa* strain UTEX 2385. 1.5 mL of E. coli donor and M. aeruginosa recipient were pipetted into separate microfuge tubes and centrifuged at 4000 x g for 5 minutes at room temperature. To concentrate the samples, the supernatant from the centrifuged samples were decanted and an additional 1.5 mL of donor and recipient were added to their respective tubes and centrifuged. To wash the donor and recipient, the supernatant from both tubes were decanted and the pellet was resuspended in 1 mL 1X PBS pH 7.4 and centrifuged. The OD_{600} and OD_{730} of the donor and recipient respectively were taken to measure cell count for a desired count of $8*10^8$ to $1*10^9$ cells. Donor and recipient samples were combined in a 1:1, 5:1 and 10:1 ratio of donor to recipient reaction. A positive and negative control of only donor and only recipient respectively were also created. The 1:1, 5:1, 10:1, and control reactions were centrifuged and the supernatant was aspirated from the samples. The pellet was then resuspended in 10 µL 1X PBS pH 7.4. The 10 µL of cell mixtures were plated on separate 2% agar BG11+DAP+2% galactose and dried for 30 minutes. The plates were incubated at 20°C for 24-72 hours.

After plate incubation, transconjugants were selected. Spot reactions were scraped and resuspended in 1 mL 1X PBS pH 7.4 in a microfuge tube. For each sample, a series of serial dilutions from 10^0 to 10^7 were made in a 96 well plate, in which the dilutions consisted of 20 µL of conjugation sample and 180 µL of 1X PBS pH 7.4. 5 µL spots of dilutions for each sample were plated on the following with an 8-multichannel pipette: 2% agar BG11 without antibiotics and 2% agar BG11 + 50 mg/mL kanamycin. In addition, 100 µL of conjugation reactions were plated on a 2% agar BG11 plate + kanamycin to acquire biomass and integration efficiency. The spots were left to dry for 30 minutes and plates were incubated at 20°C for 24-48 hours.

4.4 Natural Transformation of *M. aeruginosa* strain UTEX 2385

The natural transformation protocol was adapted from the 2019 HK_SSC iGEM team. [24]. A sample of 30 mL of a *M. aeruginosa* culture was pipetted into a falcon tube and centrifuged for 10 minutes at 2500 x g. The supernatant was decanted and cells were washed using 10mM NaCl. The sample was centrifuged again and a tenth of the original volume of BG-11 was added to resuspend the pellet. 2 µg of pSHDY was pipetted into the sample and mixed by inverting. The sample was incubated in the dark overnight at 20°C.

After incubating the culture, it was added to 25mm diameter, 0.45 µM pore size millipore filters placed on BG-11 plates. The plates were incubated for 20 hours at 20°C in a 12:12 light dark cycle with a light lux level < 3200 and consistent gas exchange.

The filters were then transferred to 2% agar BG-11 plates containing antibiotics to select for transformants. The antibiotics used were spectinomycin, kanamycin and chloramphenicol at a 15µg/mL concentration. The plates were then incubated for 7-10 days at 20°C.

Section 5

Computational Methods and Chameleon

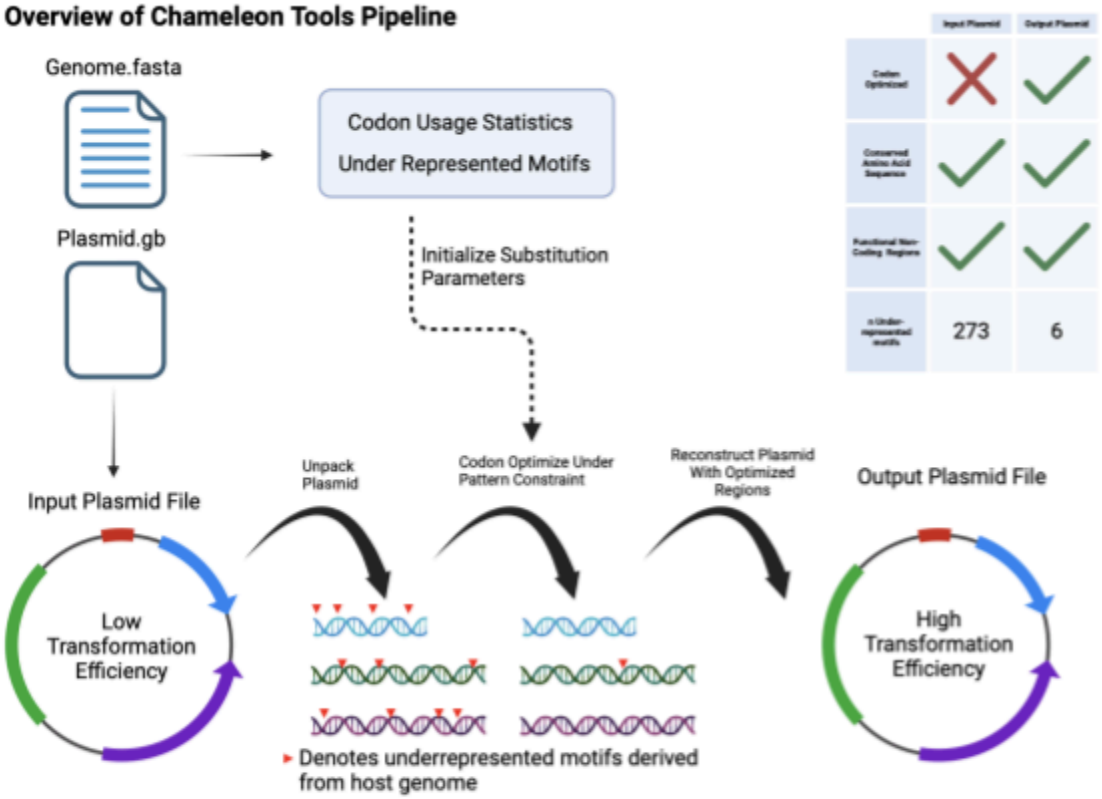

Fig. 5.1. An overview of the Chameleon Tools Plasmid Stealth (pStealth) Pipeline. Required input consists of a fastA format genomic sequence file and a genbank format plasmid file with annotated features. Codon usage statistics and under-represented genomic motifs are identified and collected from the genome file, and genbank annotations are used to deconstruct the plasmid so that protein coding regions can be optimized. The output consists of an optimized plasmid file that closely aligns to the hosts codon usage profile while containing a minimal number of under-represented motifs.

5.1 Background

The presence of Restriction-Modification (R-M) systems is widespread in prokaryotes and presents one of the most significant barriers impeding genetic tractability in non-model bacteria [26]. This barrier is especially evident in *Microcystis aeruginosa*, which possesses an extensive and robust RM system with significant variation between lineages [25].

Evasion of the R-M system may be achieved by ensuring that synthetic DNA does not contain active restriction targeting sites. Previously-successful transformations of *M. aeruginosa* have employed a "mimicry-by-methylation" approach which may yield marginal improvements in transformation efficiency. The goal of this approach is to replicate the methylation profile of the target bacteria in synthetic DNA, rendering it inert to the R-M system, via the *in vitro* pretreatment of synthetic DNA with a methylase cocktail. Another approach, termed "Stealth by engineering", is achieved by synthesizing DNA void of restriction targets, and has been shown to be highly efficient .

All existing approaches hinge on the characterization of the target bacteria's R-M system for insight into actionable evasion, a requirement that remains unmet for many non-model bacteria. Established approaches to characterize R-M targeting profiles include the use of specialized sequencing technologies, such as Single-molecule real-time (SMRT), that reveal methylomic insight, and gene analysis methods that offer limited precision in target predictions. Still, such technologies are not accessible to many researchers [26]. This drastically limits the application of modern genetics in non-model organisms and presents a potentially insurmountable road-block.

Dr. Bernick introduced us to the Stealth program, which capitalizes on the hypothesis that restriction motifs become under-represented in the genome over time. Due to the semi-conservative nature of DNA replication, there is a temporary lapse in the protective methylation of endogenous DNA as nascent DNA strands have not yet interacted with the modification component of the R-M system. This renders the unmethylated DNA vulnerable to interaction with the restriction component of the R-M system, an interaction that will induce a costly DNA damage response in the host bacterium. This results in a selective pressure against the presence of restriction motifs, favoring variants with fewer occurrences of such motifs and leading to their under-representation in the genome over time.

The original version of Stealth assesses the expected frequency of all possible motifs, using a hidden markov model based approach, and compares these values to the actual counts of motifs within genomic sequence. The program then outputs a list of under-represented motifs. The researchers behind Stealth have shown an increase in transformation of over 40,000X by employing the stealth^2 by engineering approach in *H. pylori*, an organism shown to resist plasmid transformation [27].

The Stealth framework provides a strong foundation for the evasion of the R-M system by revealing targeting DNA motifs. We set out to develop the Chameleon pipeline as an end-to-end solution to barriers imposed by the R-M system of a target bacteria, dependent on a basic genomic sequence of a target bacteria. The Chameleon pipeline offers sequence optimization by combining established codon optimization methodologies with the consideration of negatively weighted motifs, yielded from the Stealth program, to determine the optimal sequence. This avoidance of underrepresented motifs should increase the genetic tractability of non-model bacteria with extensive R-M systems, and in turn increase the transformation efficiency of synthetic plasmid into the live cell. Another key consideration in the avoidance of such sequences is an integrative approach to sequence optimization that considers other factors such as codon usage.

5.1.1 Stealth Analysis

From sequencing, the total sequencing data was able to produce multiple contigs of sequencing reads that were confirmed or hypothesized to be the UTEX LB 2385 strain of *M.aeruginosa*. Of the contigs produced, 3 contigs were BLAST confirmed *M.aeruginosa* based on ribosomal RNA genes found within the sequence, and an additional 4 contigs were hypothesized *M.aeruginosa* based on observed read coverage. Stealth V3 calibration was performed for BLAST confirmed contigs, being run with the parameters of 1,200,000 random samples, p-value cutoff of 0.05, a bootstrapping cutoff score of 55 based on a 20% FDR, and a maximum motif size of 8. To account for the undersampling of k-mers of size 4 and 5, another Stealth V3 calibration was performed on the same BLAST confirmed contigs for maximum size 6. Stealth analysis was then performed on the BLAST contigs with options of 250,000 samples, p-value cutoff of 0.05, a bootstrapping cutoff of 65 based on a 20% FDR, and a maximum motif size of 6 instead of 8. Some manual data processing was done to create a union of both Stealth analysis runs which was then processed by the Chameleon pipeline to extract and expand reverse-complement palindrome

degenerate-base motifs of sizes 4 to 8. In total, this process identified 100 unique and fully expanded reverse complement palindrome motifs of all sizes between 4 to 8 inclusive.

5.2 Chameleon

The software contribution of the project offers a complement to the Stealth program in the form of a software pipeline titled Chameleon that takes the input of a synthetic plasmid sequence and outputs a codon-optimized and Stealth motif removed version of the same plasmid sequence. Chameleon aims to accomplish: reading in plasmid sequence and target host genome, performing Stealth analysis to identify target motifs, performing codon frequency analysis for codon optimization, and finally outputting a Stealth motif reduced plasmid that preserves the desired functionality while improving transformation efficiency.
We set out to develop Chameleon Tools as a generalizable solution to restriction site avoidance and codon optimization for researchers working in non-model prokaryotes. In the broader scope, this project opens the design space for work with numerous other non-model bacteria species that would otherwise be unfeasible to engineer; Stealth and Chameleon are highly generalizable programs that could be applied in any case of an R-M system challenging transformation efficiency, which are ubiquitous among bacteria

5.2.1 Pipeline Flow

The initial front end of Chameleon is a python script that takes two inputs: a bacterial host genome file in GenBank or FastA format, and a plasmid with important coding and non-coding sequences well-annotated in GenBank format. The initial python script firstly parses the input genome file to extract coding sequence (CDS) regions for use in codon frequency analysis and saves the full sequence for later use in Stealth. This first script also parses through the plasmid file to find regions of the plasmid that are mutable. Mutable regions are defined as areas of the plasmid within a CDS that do not overlap any annotated non-coding sequence or other CDS that lie on different frames and do not include any CDS start codons. The decision to avoid the start codon of genes was made to prevent altering any alternative start codons which had the potential to break genes that utilized them. These regions are given as Biopython SimpleLocation classes that represent 0-indexed ranges and include position on the forward or reverse strand and are trimmed to lie in frame of the parent CDS they originated from.

The second component of Chameleon past input reading is Stealth analysis and codon frequency analysis which takes in the output of reading in the host genome file from the initial pipeline script. For the purposes of both the pipeline fluidity and having a referential basis, Stealth V1 is used in the publicly available version of the pipeline [27]. Stealth analysis is performed using the host genome and the output is parsed to expand and preserve only the reverse complement palindromes to avoid. Sequentially, codon frequency is assessed from the CDS regions parsed from the host genome file and a frequency usage table is stored for later use.

Lastly, the sequence of the input plasmid is partitioned into sections defined by the mutable regions output by the initial pipeline script. These sequences which represent sections of CDS regions in coding genes are then codon optimized by translating to amino acid, and regenerating a codon sequence by using codon frequency table values to prime a weighted choice method, resulting in a tendency towards recapitulation of the respective host organism codon usage frequency, aligning codon usage with available tRNA pools. From there, any motifs that could be erased by modifying these regions are removed by introducing silent mutations via the ChromatoSeq module (overview shown below). In the event that not all motifs may be avoided through optimization, the modification or sequence that results in the fewest motifs remaining is kept and used. This process is repeated a variable number of times, and the resultant sequence outputs for each plasmid section given are further screened (based on codon adaptation indices and mRNA ΔG) in

order to rationally sort and select an optimized sequence output under the specified constraints. The introduction of layers of non-deterministic motif removal and sequence generation enhances functionality of the pipeline as various other external factors can affect the viability of the final output (eg. IDT complexity score, hairpin formations, etc.) and randomly regenerated sequences have the ability to avoid these factors that deterministic sequence generation simply do not.

Section 6

Discussion

6.1 Future Research

After completing our iGEM project's foundational work in improving transformation efficiency and demonstrating the RP4 conjugation system from *E. coli* to *M. aeruginosa*, we will build our planned *mcy* cluster-disrupting conjugative plasmid (pINTO), ultimately pursuing deployment in Pinto Lake to combat HABs therein. In preparation for this, we are already in contact with Bryan Condy, the City of Watsonville's laboratory manager, who has shown interest in our project's goal of reducing microcystin contamination in Pinto Lake. Before releasing any engineered organism into Pinto Lake, we will have to both conduct more research substantiating our project and seek approval from the relevant environmental authorities. This research will include assessing how the engineered *E. coli* and the plasmid itself will behave in conditions resembling Pinto Lake; this will be done to determine efficacy of our project in a realistic context and identify any unintended consequences. Adjustments to the project may be required before it can be responsibly deployed, such as additional anti-escape measures beyond the *E. coli* strain deployed being a DAP auxotroph. Although the project focuses on *M. aeruginosa*, real-world deployment would also require consideration of effects, both direct and indirect, on other species in Pinto Lake.

The pINTO plasmid will be built by combining the high transformation efficiency and RP4 conjugation system of pSHDY with the INTEGRATE (insertion of transposable elements by guide RNA-assisted targeting) system. The INTEGRATE system is an RNA-guided transposon including the TniQ-Cascade complex, transposase proteins, and a DNA payload. The guide RNA corresponds to a sequence 50 base pairs upstream of the target for the insertion of the DNA payload. We are using the INTEGRATE system because it enables high-accuracy targeting of a specified genomic sequence and the insertion of a large DNA payload. Using our sequenced region of the *M. aeruginosa* UTEX 2385 *mcy* cluster, we have already identified a guide RNA that enables the insertion of a payload into the mcyA gene's start codon. Since the mcyA, mcyB, and mcyC genes all rely on the same promoter for expression, inserting a payload containing an early transcriptional terminator would halt expression of all three genes, thereby halting microcystin production.

To validate the efficacy of pINTO, microcystin production from *M. aeruginosa* transformants must be quantified and shown to be zero. Multiple methods are available to do so, including ELISAs and LC-MS. Through multiple lines of evidence, we will verify both that transformation with pINTO causes *M. aeruginosa* cells to halt microcystin production and that the halting of production is maintained across subsequent generations.

6.2 Future Impact and other applications

Given sufficient proof of the efficacy of pINTO in laboratory conditions resembling Pinto Lake and permission from the City of Watsonville (and other environmental authorities), we will be able to apply our project in the real world, thereby demonstrating a minimally-disruptive and inexpensive solution to reducing HABs of *M. aeruginosa* in a freshwater body. It is our hope that following success in Pinto Lake, we will be able to build a generalizable system for addressing microcystin contamination in other water systems around the world.

Aside from using conjugation to deliver plasmids, it may be possible to leverage the previously-described natural competency of *M. aeruginosa* to treat water systems using naked plasmid DNA. This would assuage some safety concerns regarding escape by avoiding delivering live organisms. To determine whether this would be a viable route instead of conjugation, we would have to reflect on our data regarding the efficiency of natural transformation and determine the amount of plasmid that would need to be delivered in a spot treatment to be effective. We may also have to assess the viability of naked DNA in the natural environment, as it may be vulnerable to degradation by extracellular nucleases.

Beyond addressing HABs of *M. aeruginosa*, our work in empirically validating the Stealth and Chameleon systems (should using them prove to increase the transformation efficiency of pSPDY) will make a valuable contribution to the field of bacterial engineering by presenting a method and toolset for improving otherwise-low transformation efficiencies in cyanobacteria and other non-model species by evading the R-M system.

6.3 Ethics and Safety

iGEM's no-release policy has been adhered to by our project, as we have not released any engineered organisms or plasmid into the natural environment. Though our project's ultimate goal is to combat HABs in the natural environment, we only intend to do so after the iGEM competition and in cooperation with environmental authorities. There are extant safety concerns with the project as it exists, and no release will occur until these concerns have been fully addressed. One concern regards long-term effects on the species composition of the water system following deployment; it is not certain how the wider ecosystem would be affected by the elimination of microcystin production in the water system's *M. aeruginosa* population. Given that there are other cyanobacteria capable of producing other cyanotoxins that may be of concern, it would be important to consider whether the deployment of our project might enable other toxic species instead. Plasmid escape is another safety concern. Although the auxotrophic nature of the EcGT2 *E. coli* strain being used ensures that conjugation from *E. coli* is transient, it is uncertain whether horizontal spread can continue by other means. For example, *M. aeruginosa* is known to possess pilus-like structures and type IV pilus-like genes [18]; although conjugation from *M. aeruginosa* has not been established, unintended spread of the plasmid may occur if this is possible.

In light of our safety concerns, we have held meetings with freshwater ecologists, including Dr. Raphael Kudela and Dr. Marylou Sison-Magnus. From these meetings, it was concluded that concerns for the ecological impacts of the transformation of *M. aeruginosa*, such as its out-competition by other species of cyanobacteria, is of secondary importance as compared to the direct harmful effects of *M. aeruginosa* blooms. Despite some of our safety concerns being alleviated, we have continued to emphasize the criticality of environmental responsibility before releasing any engineered organisms. Careful testing, consultation, and regulatory approval must always precede any release of this kind.

Bibliography

[1] R. Melaram, A. R. Newton, and J. Chafin, "Microcystin Contamination and Toxicity: Implications for Agriculture and Public Health," *Toxins*, vol. 14, no. 5, Art. no. 5, May 2022, doi: 10.3390/toxins14050350.

[2] N. Butler, J. C. Carlisle, R. Linville, and B. Washburn, "MICROCYSTINS: A Brief Overview of their Toxicity and Effects, with Special Reference to fish, Wildlife and Livestock," 2009.

[3] "Facts about Cyanobacterial Blooms for Poison Center Professionals | CDC," Nov. 28, 2022. https://www.cdc.gov/habs/materials/factsheet-cyanobacterial-habs.html (accessed Jul. 26, 2023).

[4] T. Le, "Microcystis: Toxic Blue-Green Algae," *Off. Environ. Health Hazard Assesment*, [Online]. Available: https://oehha.ca.gov/media/downloads/ecotoxicology/fact-sheet/microfactsheet122408.pdf

[5] C. Joab and G. Chetelat, "Harmful Algal Bloom Primer," *Reg. Water Qual. Control Board Cent. Val. Reg.*.

[6] N. O. and A. A. US Department of Commerce, "Why do harmful algal blooms occur?" https://oceanservice.noaa.gov/facts/why_habs.html (accessed Jul. 26, 2023).

[7] Sacramento Environmental Commission, "Cyanobacteria in Sacramento Region Waterways," Technical Report. [Online]. Available: https://emd.saccounty.gov/SEC/Documents/Final%20Cyanobacteria%20Report.pdf

[8] V. V. Le, A. Srivastava, S.-R. Ko, C.-Y. Ahn, and H.-M. Oh, "Microcystis colony formation: Extracellular polymeric substance, associated microorganisms, and its application," *Bioresour. Technol.*, vol. 360, p. 127610, Sep. 2022, doi: 10.1016/j.biortech.2022.127610.

[9] Charles Montoya, Emma Kluzniok, Steve Palmisano, Ben Heinstein, and Robert Ketley, "Pinto Lake Restoration Project Final Report," City of Watsonville, Grant Report. Accessed: Jul. 26, 2023. [Online]. Available: https://www.waterboards.ca.gov/rwqcb3/water_issues/programs/tmdl/docs/pinto_lake/final319grant_report_pinto_lake2018.pdf

[10] "Pinto Lake Restoration Efforts | Watsonville." https://www.watsonville.gov/728/Pinto-Lake-Restoration-Efforts (accessed Jul. 26, 2023).

[11] S. Rakesh, S. Saxena, D. Dhar, R. Prasanna, and A. Saxena, "Comparative evaluation of inorganic and organic amendments for their flocculation efficiency of selected microalgae," *J. Appl. Phycol.*, vol. 26, Feb. 2014, doi: 10.1007/s10811-013-0114-4.

[12] G. Lamas-Samanamud *et al.*, "The Role of Quorum Sensing in the Development of Microcystis aeruginosa Blooms: Gene Expression," *Microorganisms*, vol. 11, no. 2, Art. no. 2, Feb. 2023, doi: 10.3390/microorganisms11020383.

[13] "Genetic engineering of marine cyanophages reveals integration but not lysogeny in T7-like cyanophages | The ISME Journal." https://www.nature.com/articles/s41396-021-01085-8 (accessed Jun. 05, 2023).

[14] L. Frangeul *et al.*, "Highly plastic genome of Microcystis aeruginosa PCC 7806, a ubiquitous toxic freshwater cyanobacterium," *BMC Genomics*, vol. 9, no. 1, p. 274, Jun. 2008, doi: 10.1186/1471-2164-9-274.

[15] B. Gibson, D. J. Wilson, E. Feil, and A. Eyre-Walker, "The distribution of bacterial doubling times in the wild," *Proc. R. Soc. B Biol. Sci.*, vol. 285, no. 1880, p. 20180789, Jun. 2018, doi: 10.1098/rspb.2018.0789.

[16] A. E. Wilson, W. A. Wilson, and M. E. Hay, "Intraspecific Variation in Growth and Morphology of the Bloom-Forming Cyanobacterium Microcystis aeruginosa," *Appl. Environ. Microbiol.*, vol. 72, no. 11, pp. 7386–7389, Nov. 2006, doi: 10.1128/AEM.00834-06.

[17] K. Wallden, A. Rivera-Calzada, and G. Waksman, "Type IV secretion systems: versatility and diversity in function," *Cell. Microbiol.*, vol. 12, no. 9, pp. 1203–1212, Sep. 2010, doi: 10.1111/j.1462-5822.2010.01499.x.

[18] K. Nakasugi and B. A. Neilan, "Identification of Pilus-Like Structures and Genes in Microcystis

aeruginosa PCC7806," *Appl. Environ. Microbiol.*, vol. 71, no. 11, pp. 7621–7625, Nov. 2005, doi: 10.1128/AEM.71.11.7621-7625.2005.

[19] C. Engler, R. Kandzia, and S. Marillonnet, "A One Pot, One Step, Precision Cloning Method with High Throughput Capability," *PLoS ONE*, vol. 3, no. 11, p. e3647, Nov. 2008, doi: 10.1371/journal.pone.0003647.

[20] K. Lagesen, P. Hallin, E. A. Rødland, H.-H. Stærfeldt, T. Rognes, and D. W. Ussery, "RNAmmer: consistent and rapid annotation of ribosomal RNA genes," *Nucleic Acids Res.*, vol. 35, no. 9, pp. 3100–3108, May 2007, doi: 10.1093/nar/gkm160.

[21] R. R. Wick, M. B. Schultz, J. Zobel, and K. E. Holt, "Bandage: interactive visualization of *de novo* genome assemblies," *Bioinformatics*, vol. 31, no. 20, pp. 3350–3352, Oct. 2015, doi: 10.1093/bioinformatics/btv383.

[22] F. Nies, M. Mielke, J. Pochert, and T. Lamparter, "Natural transformation of the filamentous cyanobacterium Phormidium lacuna," *PLOS ONE*, vol. 15, no. 6, p. e0234440, Jun. 2020, doi: 10.1371/journal.pone.0234440.

[23] ZYMO RESEARCH, "Quick-DNA HMW MagBead Kit." https://www.zymoresearch.com/products/quick-dna-hmw-magbead-kit (accessed Jul. 26, 2023).

[24] "Team:HK SSC/Improve - 2019.igem.org." https://2019.igem.org/Team:HK_SSC/Improve (accessed Jul. 27, 2023).

[25] L. Zhao, Y. Song, L. Li, N. Gan, J. J. Brand, and L. Song, "The highly heterogeneous methylated genomes and diverse restriction-modification systems of bloom-forming Microcystis," *Harmful Algae*, vol. 75, pp. 87–93, May 2018, doi: 10.1016/j.hal.2018.04.005.

[26] C. D. Johnston *et al.*, "Systematic evasion of the restriction-modification barrier in bacteria," *Proc. Natl. Acad. Sci.*, vol. 116, no. 23, pp. 11454–11459, Jun. 2019, doi: 10.1073/pnas.1820256116.

[27] S. Hu, S. Giacopazzi, R. Modlin, K. Karplus, D. L. Bernick, and K. M. Ottemann, "Altering under-represented DNA sequences elevates bacterial transformation efficiency," University of California, Santa Cruz, California, Aug. 06, 2023. doi: 10.1128/mbio.02105-23.