2022

CPF1-BASED CRISPR GENOME EDITING IN THE CYANOBACTERIUM N. PUNCTIFORME

Soohan Woo
University of the Pacific

CPF1-BASED CRISPR GENOME EDITING IN THE CYANOBACTERIUM *N. PUNCTIFORME*

By

Soohan Woo

A Thesis Submitted to the

Graduate School

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF ARTS

College of the Pacific
Biological Sciences

University of the Pacific
Stockton, CA

2022

CPF1-BASED CRISPR GENOME EDITING IN THE CYANOBACTERIUM *N. PUNCTIFORME*

By

Soohan Woo

APPROVED BY:

Dissertation Advisor: Douglas Risser, Ph.D.

Committee Member: Geoffrey Lin-Cereghino, Ph.D.

Committee Member: Craig Vierra, Ph.D.

Department Chair: Eric Thomas, Ph.D.

Dean of Graduate School: Thomas H Naehr Ph.D.

CPF1-BASED CRISPR GENOME EDITING IN THE CYANOBACTERIUM *N. PUNCTIFORME*

By

Soohan Woo

DEDICATION

This thesis is dedicated to my friend Enling, who constantly checked in on my progress on the thesis and encouraged me as I struggled through writing it. She would always ask about my health to make sure that I was eating well and that I wasn't neglecting my well-being while doing my research. Without her friendship, I would have had a more difficult time balancing my work on this thesis with the other aspects of my life. Thank you, Enling, for your selflessness and heart to care for others, from which I have benefited greatly.

ACKNOWLEDGMENTS

I am grateful for Dr. Risser's patient guidance despite the limited time and the other constraints to my research that came with adapting to a global pandemic. The training that I received under his tutelage and his expertise will aid me as I take these skills beyond the education system.

Geoff Lin-Cereghino's constant encouragement was always uplifting and much desired throughout this graduate program; his words will always be remembered with fondness.

CPF1-BASED CRISPR GENOME EDITING IN THE CYANOBACTERIUM *N. PUNCTIFORME*

Abstract

By Soohan Woo

University of the Pacific
2022

CRISPR systems have been growing in their utility and their application throughout the biological field as researchers continue to grow in their understanding of the relatively novel genome editing technology. However, despite the potential of CRISPR as a genome editing tool, the complexity of applying this technology to a specific organism calls for custom modifications to the system to improve its success rate. In this project, a CRISPR-Cpf1 system that can be effectively employed in the cyanobacterium *Nostoc punctiforme* was designed, with a focus on the hormogonium development of this species. Multiple plasmids containing the CRISPR system and targeting different genes were constructed using a Gibson-based rapid assembly cloning method, and then were tested by introduction into *Nostoc punctiforme* via conjugation. Plasmids were constructed to mutate 7 different genes in *N. punctiforme* with 4 of the 7 successfully mutating their target genes. For one of the genes where the plasmid failed to produce mutants, the usage of a larger homology repair template (HRT) was found to enhance the efficiency of gene editing, allowing the gene to be knocked out. Thus, the length of the HRT appears to be a critical factor in designing successful constructs. The system developed in this project aims to make CRISPR a more viable tool in studying *Nostoc* cyanobacteria, and more specifically to aid in understanding the mechanisms behind hormogonium development in the

studied species. This system may have a wider application for studying the *Nostoc* genus and related organisms, such as *Anabaena*.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1: *NOSTOC PUNCTIFORME* AND CRISPR

Nostoc punctiforme and Its Relevance in the Scientific Community

Cyanobacteria belong to a phylum of prokaryotes that are oxygenic and photosynthetic, and are known to play a significant role in the global nitrogen cycle (Henson et. al., 2004). They are found in most aquatic and terrestrial environments, and have been noted to be one of the largest sources of fixed nitrogen in the world. Amongst the species of nitrogen-fixing cyanobacteria, some of them participate in symbiotic relationships with certain plants, providing the host plant with a directly supply of reduced nitrogen (Henson et. al., 2004).

Nostoc punctiforme Cell Differentiation

Nostoc punctiforme is a species of filamentous cyanobacteria, with cells forming a long, thin line. During the life cycle of *N. punctiforme*, its vegetative cells differentiate into heterocysts, akinetes, and hormogonia (Moraes et. al., 2017). Heterocysts are the species' nitrogen-fixing cells, and allow for the oxygen-sensitive nitrogenases in the cyanobacteria to function effectively. Akinetes are spore-like cells, granting the species endurance through harsh environmental factors such as cold and desiccation (Campbell et. al., 2007). Hormogonia, which will be the cell type of interest in this paper, are filaments that provide motility and are able to sense and respond to the environment, which is crucial for a photosynthetic, phototactic species like *N. punctiforme* (Meeks et. al., 2002).

N. punctiforme and Symbiotic Relationships

Studies done on the symbiotic relationships between plants and *N. punctiforme* have shown that, in some cases, the host plant tends to have a positive effect on the frequency of heterocyst and hormogonium differentiation (Wong et. al., 2002). This is beneficial for the host

plant, because the heterocysts provide fixed nitrogen for plant growth, while the hormogonia act as infectious agents to form the symbiotic relationship between the two species. The advantages that *N. punctiforme* provides for its host plants have been additional reasons to highlight this cyanobacterium as an organism to be studied.

CRISPR and Its Application to Cyanobacteria

CRISPR-Cas systems are adaptive immune systems found naturally in bacteria and archaea, but have been engineered into a genome editing technique more recently (Gong et. al., 2021). These systems have created greater potential for genetic editing to be done at a significantly faster speed across a broad range of organisms. CRISPR-Cas systems utilize a crRNA to direct a Cas protein, a type of RNA-guided endonuclease, to a target DNA site identified by a protospacer that lines up with the spacer and a protospacer-adjacent motif (PAM) in order to induce a break in the double-stranded DNA. The most well-known CRISPR-Cas system is CRISPR-Cas9, which was first found in *Streptococcus pyogenes*. CRISPR-Cas9 is characterized by these features: ribonuclease III activity allows for crRNAs to mature from a pre-crRNA array; the Cas9 protein is directed to a target DNA site by the crRNA in complex with a tracrRNA; Cas9 recognizes the target with a 3' G-rich PAM site and proceeds to cut the DNA, resulting in blunt ends (Gong et. al., 2021). These ends can be rejoined by non-homologous end joining, which can result in small insertions or deletions known as indels (Su et. al., 2016). In the case that a homologous repair template (HRT) is provided, homologous repair can be used to fix the breakage in the DNA, allowing for some form of alteration, insertion, or deletion of the DNA to occur between the homology arms (Savic et. al., 2018).

CRISPR-Cpf1 as an Alternative to CRISPR-Cas9

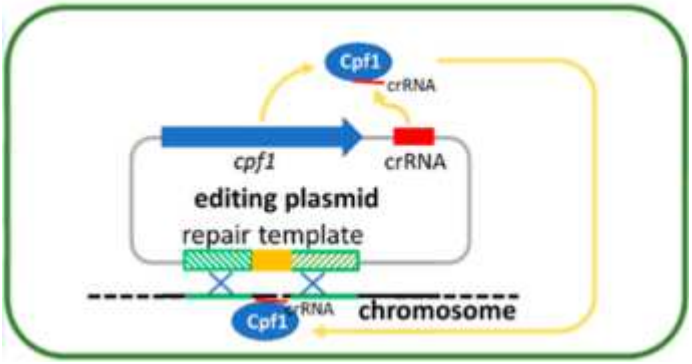

Figure 1. A diagram of the editing plasmid containing the CRISPR-Cpf1 system, displaying the following components: a *cpf1* gene, crRNA, and a homologous repair template (Niu et. al., 2018).

CRISPR-Cpf1 differs from CRISPR-Cas9 in multiple ways: the Cpf1 protein recognizes the target DNA site with a 5' T-rich PAM; the system does not require the aid of tracrRNA for the target DNA site to be cut; the Cpf1 cuts the double-stranded DNA to produce staggered ends, not blunt ends (Niu et. al., 2018). Cpf1-containing CRISPR-Cas loci in *Francisella tularensis subsp. Novicida U112* have been found to encode effective CRISPR-Cpf1 systems as a defense mechanism for the bacterium (Zetsche et. al., 2015). By studying these loci, researchers have been able to characterize Cpf1-containing CRISPR systems and show that Cpf1 is a single RNA-guided endonuclease, as shown in Figure 1. Cpf1 has been shown to significantly differ from Cas9 when it comes to structure and function. Because Cpf1 does not require tracrRNA to process crRNA arrays or to mediate interference, the design for a CRISPR-Cpf1 system is potentially simpler by a notable degree than for a CRISPR-Cas9 system (Deltcheva et. al., 2011). As an example, the roughly 42-nucleotide long crRNA used by Cpf1 is both easier and cheaper to produce than the roughly 100-nucleotide long guide RNA used by systems that utilize Cas9

(Zetsche et. al., 2015). Additionally, the Cas9 protein is known to cut close to the PAM site, while the Cpf1 protein is known to cut away from the PAM site in the 3' direction (Gong et. al., 2021). Thus, it is important that homologous repair is implemented into the CRISPR system, because indels created by non-homologous end joining will not remove the target site. If the target site is not removed, the chromosome will be cut continuously, leading to a lethal result for the cell.

CRISPR-Cpf1 in Cyanobacteria, as Opposed to CRISPR-Cas9

Studies have shown that Cas9 is highly toxic to some organisms, and thus can not be used for genome editing in certain species (Jiang et. al., 2014). In cyanobacteria, Cas9 was also found to be toxic past a certain dosage level, and studies done to remedy the solution found that Cpf1 was significantly less toxic and thus more applicable to cyanobacteria (Ungerer et. al., 2016). The usage of CRISPR-Cpf1 is still under development due to the lack of knowledge surrounding this system in comparison with CRISPR-Cas9, but research is ongoing to make this system one that is efficiently able to knock out genes in cyanobacteria.

CRISPR-Cpf1 has been designed using the cyanobacterium *Anabaena* PCC 7120, which is closely related to our model organism, *N. punctiforme* (Niu et. al., 2018). A system was formed using a vector containing components of the replicon RSF1010 which allows for replication in cyanobacteria, the resistance marker *nptII*, the *Francisella novicida cpf1* gene, an *Aar*I-*lacZ'*-*Aar*I site flanked by CRISPR direct repeats for cloning the spacer sequence of crRNA, and a *Sal*I-*Kpn*I site for the cloning of the homologous repair template (Ungerer et. al., 2016). Two separate editing plasmids were formed for the deletion of *hetR* and *patS*, both genes necessary for heterocyst formation (Niu et. al., 2018). Although the system was effective for the deletion of *hetR*, as shown by the mutant strain being unable to form heterocysts on media with

no combined nitrogen, the system was unable to successfully delete *patS*, as shown by colonies still retaining the wild-type copies of *patS* (Niu et. al., 2018). The CRISPR-Cpf1 system's editing efficiency was improved by applying a "two-spacer" strategy, where two editing plasmids, as opposed to one, were employed with the same homologous repair templates but different spacer sequences (Niu et. al., 2018). This new strategy, when applied to the deletion of 26 different genes in *Anabaena*, showed promising results, as some of the genes were deleted with 100% or close to 100% efficiency (Niu et. al., 2018). The goal of this project overall is to successfully employ the Cpf1 system in *N. punctiforme*, which would open up possibilities for *N. punctiforme* to be genetically engineered for several different means, such as commercially viable biofuel production.

CHAPTER 2: MATERIALS AND METHODS

Targeted Genes Involved in Hormogonium Development

Genes were identified based on previous research that they are or may be involved with hormogonium development in *N. punctiforme*. The gene *hmpF*, which has been identified as being involved in motility by transposon mutagenesis (Cho et. al., 2017), was listed as one of the genes to knock out with the developed CRISPR-Cas system. The second gene that was targeted was *pilB*, which has been shown to be essential to motility, because deletion of this gene led to the complete abolishing of motility in *N. punctiforme* (Khayatan et. al., 2015). Another gene that was targeted was *hfq*, which encodes a putative RNA chaperone essential for hormogonium motility (Harwood et. al., 2021). Four additional genes, that have not been given formal gene names, were targeted as well, and given the following numerical IDs: 456, 459, 875, and 3486. These four genes have been found by a previous study on *N. punctiforme* to be essential for motility (Khayatan et. al., 2017).

Plasmid Construction

In order to construct a plasmid encoding the Cpf1 protein and containing the customized crRNA and homologous repair template, as shown in Figure 1, the 5' and 3' regions of homology were amplified with PCR using Accuzyme Mix (Bioline Inc.). The gene-specific 5'-F and 5'-R or 3'-F and 3'-R primers used to amplify the 5' and 3' regions of homology in each plasmid are listed in Table 1. The thermocycler settings for the PCR were as follows: 98^{o}C for 3 minutes for initial denaturation, 98^{o}C for 15 seconds for denaturation, 55^{o}C for 15 seconds for annealing, 72^{o}C for 2 minutes per kb in length of the DNA for extension, and 72^{o}C for 7 minutes for final extension. The denaturation, annealing, and extension steps cycled 30 times.

The spacers targeting the genes were chosen based on the criteria that: 1) the targets had the PAM 5'-KTTV-3' (K is either T or G, while V can be C, G, or A); 2) the spacers were 22 base pairs long; 3) the spacers did not have more than eight continuous AT or GC pairs; 4) the sites were unique on the chromosome (Niu et. al., 2018). The gRNA arrays were then designed using gene synthesis technology from GeneScript, ordered online, and amplified via PCR. The primers gRNAarray-F and pCpf1b-rep/trm-R listed in Table 1 were used for the PCR, and the gRNA sequences are listed in Table 2. The enzyme used and the conditions were the same as for the PCR reactions for the HRT.

The plasmid backbone, pCpf1b, or pAM504-Cpf1 were also amplified via PCR using Ranger Mix (Bioline Inc.), and the primers pCpf1b-BamHI-F and pCpf1b-R listed in Table 1. The thermocycler settings were: 95 ºC for 1 minute, 98 ºC for 10 seconds, and 65 ºC for 14 minutes, with the cycle being repeated 30 times. The primers for all PCR products contain 20-30 bp adapters with appropriate complementarity to allow for subsequent Gibson-based assembly.

Table 1

Primers for PCR Amplification of Homology Regions

Sequence Name	DNA Sequence
Npun_F0875-5'-F	gaaattcgatatctagatctcatCATCCAAAATTCTGCTAGGTG
Npun_F0875-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacCTCAGACACCAGAAATCAAAG
Npun_F0875-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAGGCAAGTAGAGAATGGGGAATG
Npun_F0875-3'-R	gttgttgccattgcggatccCAAGTTGCATAGGTTCTTAG
Npun_F0875-5'-F-con	CTAAATCCATATCGCCTTC
Npun_F0875-3'-R-con	CTAGTGCATACTAAAGATGC
Npun_F3486-5'-F	gaaattcgatatctagatctcatGATGGTACGACGTTGCGAC
Npun_F3486-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacGGTCATATTTTTATCCTCTGTG
Npun_F3486-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAGCCGCATAACAATGTACTAG
Npun_F3486-3'-R	gttgttgccattgcggatccTAGGTTTGGGGCTTTTTG
Npun_F3486-5'-F-con	GAAGACCCCCACAAACAG
Npun_F3486-3'-R-con	GGAATTTACCACGCTTTTC
Npun_F0459-5'-F	gaaattcgatatctagatctcatCAGTTCAAGTTACCGTTAG
Npun_F0459-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacGGTCATGGTCGCTTCGCTC
Npun_F0459-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAAGCAAAAACTGATATCCCAATAG
Npun_F0459-3'-R	gttgttgccattgcggatccGGAACTATCAATCAGTCTG
Npun_F0459-5'-F-con	CTTTTTCCCAGATCCAATC
Npun_F0459-3'-R-con	TCACAGGAAGAATCTTGG
Npun_R0118-5'-F	gaaattcgatatctagatctcatCCAAAAACTTGAGATGACC
Npun_R0118-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacGTAAGTCATAAATGAGCGAATG
Npun_R0118-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAACGGTTTTAAAACTAATAGTCATTGG
Npun_R0118-3'-R	gttgttgccattgcggatccCCAAAGTTTGACTAAATACC
Npun_R0118-5'-F-con	CGTCAATTGAGGTCGAAAATAC
Npun_R0118-3'-R-con	GCCATCACCCGACCATATTC
Npun_F5230-5'-F	gaaattcgatatctagatctcatCGTAGTAAGGACTTTAGTC
Npun_F5230-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacGAATTTACTAAGTATCATGGTC
Npun_F5230-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAACAAGTTAGAAATAGAATTTTGCTTGC
Npun_F5230-3'-R	gttgttgccattgcggatccGAGGAATTCGGCTGCTAC
Npun_F5230-5'-F-con	TAATCCCCCCATACAATC
Npun_F5230-3'-R-con	GTGGAGCCGTCCTAGATATG
NpF0456-5'-F	gaaattcgatatctagatctcatCAATTATTCCTGTACATCTCTATG
NpF0456-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacGAGTGCCATGATTTTTATTGG
NpF0456-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAGAAAGTTAGAACTATTGGGCATTG
NpF0456-3'-R	gttgttgccattgcggatccGCTAAAGACCAAGATGTTTCTAC
NpF0456-5'-F-con	GCAAAGGCAATTACGCCTC
NpF0456-3'-R-con	CTGCCATCAATGTACTTTG
hmpF-5'-F	gaaattcgatatctagatctcatCTCGAAAGTGTTTAATTGTGG
hmpF-5'-R	TTTTAACAGTGGCCTTATTAAATGACTTCTCcaaacATACAGCACACCGTTAGTCC
hmpF-3'-F	gtttgGAGAAGTCATTTAATAAGGCCACTGTTAAAAGCGTCTTAATAAGTTAGGAG
hmpF-3'-R	gttgttgccattgcggatccGATTTAGGTATTGCTCACTGC
hmpF-con-F	CACCTGCTTCGTAATTTTTG
hmpF-con-R	GCGGACAAGATTCTTTGTG
NpR0118-2kb-5'-F	gaaattcgatatctagatctcatTTTGACCATTGCATGGCGTAG
NpR0118-2kb-3'-R	gttgttgccattgcggatccGTATCCATCAAAAATTGCGTTAGAG
NpR0118-2kb-5'-con-F	GTAGAAGGTGTATCTTCCTTTG
NpR0118-2kb-3'-con-R	CGAAAACCTTCGGACTTCTC
pCpf1b-BamHI-F	ggatccgcaatggcaacaac
pCpf1b-R	ccgtttttgcctaaatcagcgctag
gRNAarray-F	ctagcgctgatttaggcaaaaacgg
pCpf1b-rep/trm-R	atgagatctagatatcgaatttctg

Table 2
gRNA Sequences for CRISPR Plasmids

Plasmid Name	gRNA Sequence
Npun_F0875	GCGGCTTCATTTACAAGTAGTA
Npun_F3486	GGGTGATGGCAGTCCCACCCGC
Npun_F0459	CAAACCCAACGGGAGACGCTGC
Npun_R0118	GATAACTCCTCCACCCAACTGG
Npun_F5230	ACACCACTTTACCCAGTATTAG
hmpF gRNA	GAACTGGAATTAACAGATGAAA
NpF0456 gRNA	TGGTTTAAAGGAATAGCAGGTT

Plasmid Assembly

To assemble a plasmid containing the CRISPR-Cpf1 system targeting the gene of interest, 1 microliter of each PCR reaction for the gRNA, the pCpf1b template, the 5' HRT, and the 3' HRT was added to 6 microliters of ultra-purified water and 10 microliters of Hifi assembly mix (New England Biolabs, Inc.) and incubated at 50°C for 1 hour.

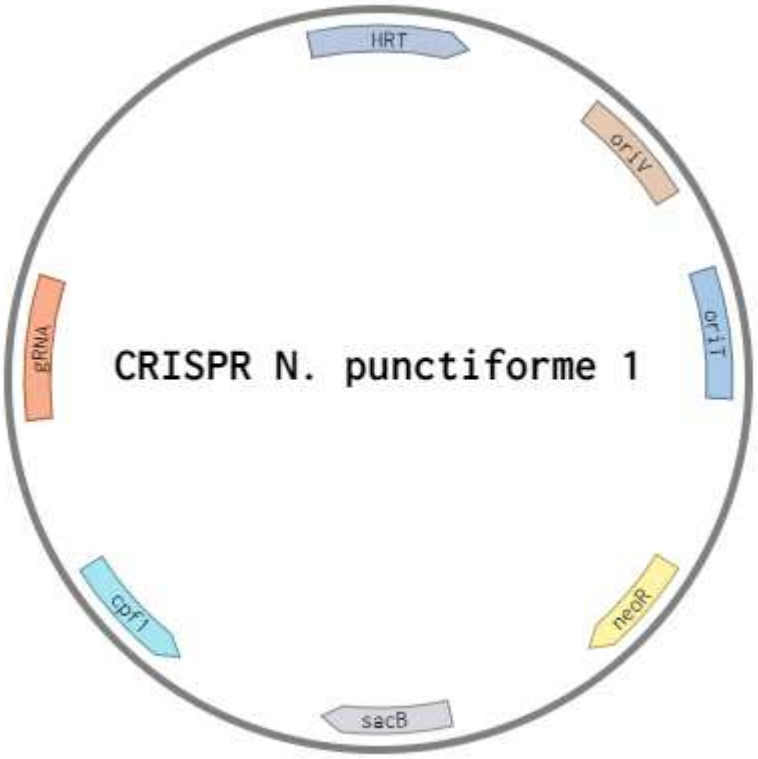

Figure 2. A diagram of the CRISPR construct containing the CRISPR-Cpf1 system, displaying the following components: HRT (homologous repair template), oriV (origin of replication for vector), oriT (origin of transfer), neoR (gene for expressing neomycin resistance), sacB, cpf1, and gRNA (guide RNA).

The final construct shown in Figure 2 involved the main components displayed in Figure 1 (HRT, cpf1, and gRNA (i.e., crRNA)), with some additional parts. An oriV, or origin of replication for vector, was included in the design so that the plasmid would be replicated within the host organism. An oriT, or origin of transfer, was also included in the design so that the plasmid could be transferred from the donor organism (*E. coli*) to the recipient (*N. punctiforme*) via conjugation. The gene neoR would confer neomycin resistance, so that the cells containing the plasmid could be selected for. The gene sacB converts sucrose into a toxic metabolite,

killing the cells (Cai et. al., 1990), so that sucrose could later be added as a means of counter selection, removing the plasmid from cells that contain it.

Transformation of Hifi Assembly Into Cells

Frozen stocks of NEB 5-α (New England Biolabs, Inc.) *E. coli* cells were thawed on ice, before 2 microliters of the assembly reaction were mixed into the cells. The mix was kept on ice for 30 minutes, and then heat-shocked at 42°C for 30 seconds. The tubes of cells were then transferred to ice for 2 minutes before 950 microliters of room-temperature SOC media were added to each tube. The tubes were incubated at 37°C for 60 minutes in a shaker. Afterwards, the cells were centrifuged for 1 minute at 5,000 rcf, and the pellet was resuspended in 100 microliters of the supernatant. Each 100 microliters of cells were then plated onto Lysogeny Broth (LB) agar plates containing 50 μg/mL kanamycin, and incubated overnight at 37°C.

Overnight Culture for Colonies

The following day, overnight cultures containing 2 mL of Terrific Broth (TB) mixed with 2 μL kanamycin (50 mg/mL) were set up for each test tube. The cultures were inoculated and then incubated in the shaker at 37°C overnight.

Plasmid Prep

1.5 mL of cells from each culture were transferred to a 1.5 mL Eppendorf tube, centrifuged at 14,000 rpm for 1 minute, and decanted to leave behind the cell pellets. The cells were resuspended in 1 mL of 0.5 M NaCl, centrifuged at 14,000 rpm for 1 minute, and decanted. Three separate solutions were then added to the cell pellets: 150 μL of Solution #1, which contains 50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, 300 μL of Solution #2, which contains 0.2 M NaOH and 1% SDS, and 225 μL of Solution #3, which contains 3 M potassium acetate and 2 M acetic acid. The tubes were inverted 7-8 times and flicked to mix the cells with

the solutions in between each addition. The tubes were centrifuged for 5 minutes at 14,000 rpm. Afterwards, 2 µL of Rnase A (10 mg/mL) were added to new tubes, and the supernatant from the centrifuged tubes was added to the new tubes containing Rnase A and incubated at 55°C for 10 minutes. After the incubation, the tubes were centrifuged for 5 minutes at 14,000 rpm, and then the supernatant was transferred to new tubes. 330 µL of isopropanol was added to the supernatant, and then the tubes were vortexed for 3 seconds, allowed to incubate at room temperate for 1 minute, and centrifuged for 5 minutes at 14,000 rpm. Being careful not to disturb the pellet and DNA, the supernatant was removed, and 1 mL of 70% EtOH was added. The tubes were inverted once and then centrifuged at 14,000 rpm for 5 minutes. The EtOH was removed by decanting, and after briefly centrifuging the tubes for roughly 8 seconds, the residual EtOH was removed using a 200 µL pipette. The tubes were then placed with the cap open on the grate of the biosafety cabinet with the blower on for 15 minutes so that the EtOH would evaporate. Finally, 50 µL of ultrapure water was added to the DNA and mixed by vortexing.

Confirmation of the Plasmid

The purified plasmid was then confirmed using restriction digest and PCR: the restriction enzymes BamHI and SpeI were added to the plasmid to both cut out the HRT region and linearize the sequence, while CutSmart buffer was also added to aid in the linearization of the plasmid. The entire mixture (3 µL plasmid, 0.5 µL BamHI, 0.5 µL SpeI, 1 µL CutSmart buffer, 5 µL ultrapure water) was incubated at 37°C for 1 hour, and then run on a gel to test for the presence of the linearized plasmid. The purified plasmids were also confirmed using PCR: the 5'-F primer and 3'-R primer for HRT in Table 1 were used to confirm the plasmid. The enzyme used and the conditions were the same as for the PCR reactions for amplifying pCpf1b.

Growing *N. punctiforme* with the Plasmid

Conjugation of *E. coli* with *N. punctiforme*

The confirmed plasmids were transformed into UC585 *E. coli* cells, each tube of UC585 containing 25 µL of cells. The cells were plated on LB agar plates containing kanamycin, ampicillin, and chloramphenicol, and the cells were allowed to grow overnight at 37°C. The transformants that grew were used to inoculate a 2 mL culture of TB containing 2 µL kanamycin (50 mg/mL), 1 µL ampicillin (50 mg/mL), and 1 µL chloramphenicol (15 mg/mL) overnight. The next day, 20 mL of LB with 20 µL kanamycin (50 mg/mL), 10 µL ampicillin (50 mg/mL), and 10 µL chloramphenicol (15 mg/mL) was poured into a 50 mL conical for each strain. 250 µL of the overnight cultures was added to each conical, and the conicals were allowed to shake at 37°C for 3 hours.

Meanwhile, the absorbance of a pre-made culture of *N. punctiforme* was measured using a spectrophotometer. 1 mL of the culture was added to a 1.5 mL tube and centrifuged for 1 minute at maximum speed. The supernatant was decanted, and the cell pellet was resuspended in 1 mL of 100% methanol. The tube was then allowed to incubate in a dark area for 5 minutes to allow the methanol to extra the chlorophyll-a (chl a) from the cells. The tube was then vortexed, and then centrifuged again for 1 minute at maximum speed. After blanking the spectrophotometer with methanol, the supernatant, containing the chl a, had its absorbance measured at a wavelength of 665 nm. Multiplying this value by 12.7 gave us the concentration of the culture in µg chl a/mL. The culture was then centrifuged and the appropriate amount of supernatant was removed so that when the cell pellet was resuspended, the culture was at the desired concentration of 125 µg chl a/mL.

After the 3-hour incubation of the *E. coli* cultures, the conicals were centrifuged at 37°C for 10 minutes at 4000 rpm. The supernatant was decanted, and 4 mL of LB was added to each conical and used to resuspend the pellet. The conicals were centrifuged for an additional 10 minutes at 37°C at 4000 rpm. After decanting the supernatant, 400 µL of LB was added to each conical and used to resuspend the pellet. 400 µL of the *N. punctiforme* culture was then added to each *E. coli* conical and mixed by pipetting, resulting in 800 µL of the mixture in each conical. 400 µL of the mixture was plated onto an AA/4 agar plate containing 0.5% LB, 10mM sucralose, and NH_4^+ MOPS, and this was repeated for each 400 µL of the mixture left. These plates were then left to grow overnight under light.

The following day, 1 mL of AA/4 was pipetted onto each plate, and a sterilized spreader was used to spread the AA/4 and scrape the sample off of the plate. A pipette was then used to transfer the liquid containing the cells to an AA/4 agar plate containing NH_4^+ MOPS, and to spread the cells across the plate. The plates were then placed under light at room temperature with the agar-side facing upwards.

Selection for *N. punctiforme* Containing the Plasmid

After 3 days, 10 mL of AA/4 with 10 µL neomycin (50 mg/mL) was added to each plate. The plates were placed under light at room temperature. After an additional 3 days, the condensation on the lid of the plate was shaken off into a bucket before being placed onto the plate again. Condensation would continue to be shaken off the lid across the next 2-4 days until the plate was dry, at which point the plate would be sealed with electrical tape and placed agar-side up under light at room temperature.

Testing for Mutants

Counting the Numbers of Colonies

After allowing the *N. punctiforme* colonies to grow for 1-2 weeks, the number of colonies on each plate would be counted and recorded. The colonies would then be picked and streaked onto new plates containing neomycin before using visual inspection and PCR verification to ascertain the presence of mutants. A successful mutant should not have been able to grow hormogonia, since all of the targeted genes should be essential for hormogonium development.

PCR Verification of Mutants

Colony PCR was used to verify whether a colony was a mutant. A colony PCR tube for each sample consisted of 12.5 µL Mango mix, 0.5 µL of each primer (5'-F-con primer and 3'-R-con primer for HRT in Table 1), and 11.5 µL of ultra-purified water, to make up a total volume of 25 µL. A piece of the colony would be added to the tube afterwards as template. The PCR conditions for the colony PCR were: 5 minutes at 94°C for initial denaturation, 30 seconds at 94°C for denaturation, 30 seconds at 55°C for annealing, and 4 minutes at 72°C for extension, for 35 cycles. The setting for the final extension, after the cycles ended, was 7 minutes at 72°C.

Chapter 3: Results

Confirmation of Plasmid Assembly Using PCR and Restriction Digest

The plasmid components, consisting of the gRNA, plasmid backbone, and the 5' and 3' regions for the HRT, were assembled via Hifi assembly. This was significantly easier and more efficient than having to separately clone the gRNA and HRT regions using traditional restriction digests and ligation. The assembled plasmid was then confirmed using restriction digest. Every type of plasmid, each of which targeted a different gene, was assembled twice. Both copies of 3 digested plasmids were run on a gel as shown in Figure 3, and only one of the plasmids (the second copy of the plasmid targeting gene 459) did not show a band at the appropriate size for the pCpf1b backbone. Because pCpf1b is a low copy number plasmid it was too difficult to visualize the present of the HRT insert, with only a faint band at the appropriate size (1.8kb) for the HRT visible in lane 3 (Fig. 3).

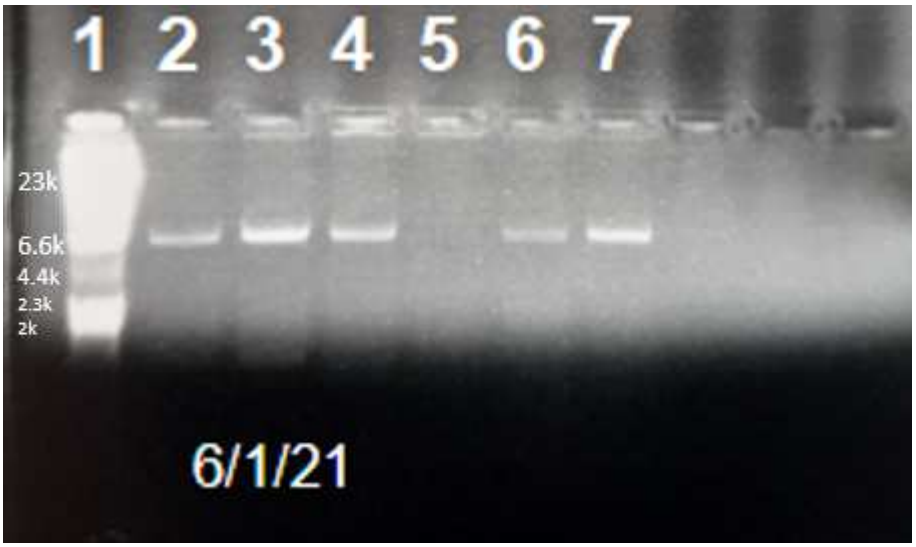

Figure 3. A gel is run to test for the presence of 2 copies each of 3 different linearized plasmids following restriction digest, each plasmid targeting a different gene. When BamHI and SpeI

(Figure 3 Continued)
were used to cut the plasmid, the expectation was that the HRT region, which should be around 1.8kb, would be excised from the plasmid, but the HRT region was difficult to visualize because the plasmid was a low copy number plasmid. **Lane 1:** Ladder. **Lanes 2-3:** Linearized plasmids targeting gene 456. **Lanes 4-5:** Linearized plasmids targeting gene 459. **Lanes 6-7:** Linearized plasmids targeting gene 875.

To further confirm the present of the HRT region in each plasmid, colony PCR was employed. Figure 4 gives an example of a successful colony PCR. Two copies of each plasmid were assembled, and a colony was taken to represent both copies of the Δ*pilB*, Δ456, Δ459, and Δ875 plasmids. Colony PCR was performed using pieces of those colonies, with a negative control and a positive control for each type of plasmid. The positive control used wild-type *N. punctiforme* chromosomal DNA as template instead of a piece of a colony, while the negative control contained no template DNA. The expected size for the HRT region was 1.8kb, and the expected size for the positive control with chromosomal DNA as template was 1.8kb plus the size of the gene of interest. Overall, the colony PCR was successful, as every plasmid can be shown to be giving product in lanes 1, 2, 5, 6, 9, 10, 13, and 14. The second copy of the Δ459 plasmid gives noticeably less product than the other plasmids, which is consistent with the results from Figure 3.

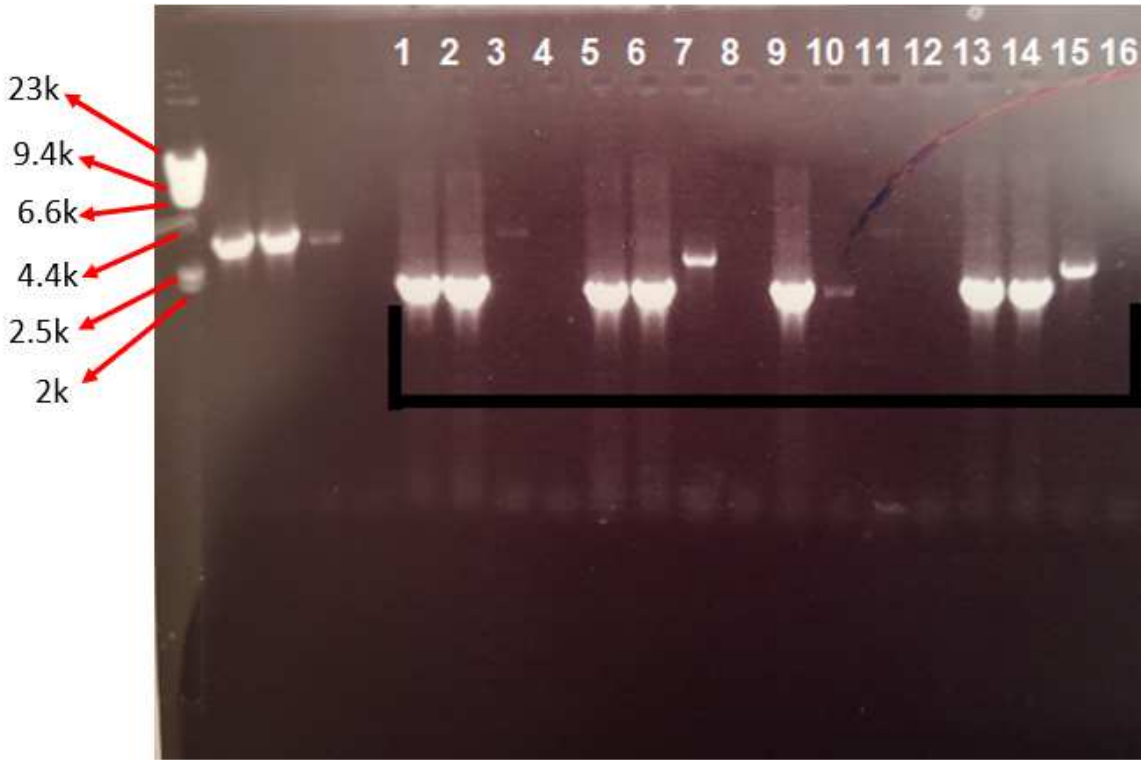

Figure 4. A gel is run after performing colony PCR for 4 of the assembled plasmids: Δ*pilB*, Δ456, Δ459, and Δ875. The relevant lanes are outlined by the black lines. The first lane of the outlined area will be referred to as Lane 1. **Lanes 1-2:** The colony PCR results are shown for two copies of the Δ*pilB* plasmid. **Lanes 3-4:** The positive and negative controls, respectively, are shown for the Δ*pilB* plasmid. **Lanes 5-6:** The colony PCR results are shown for two copies of the Δ456 plasmid. **Lanes 7-8:** The positive and negative controls, respectively, are shown for the Δ456 plasmid. **Lanes 9-10:** The colony PCR results are shown for two copies of the Δ459 plasmid. **Lanes 11-12:** The positive and negative controls, respectively, are shown for the Δ459 plasmid. **Lanes 13-14:** The colony PCR results are shown for two copies of the Δ875 plasmid. **Lanes 15-16:** The positive and negative controls, respectively, are shown for the Δ875 plasmid.

Initial Success of the CRISPR-Cpf1 System

In a previous study that tested the efficacy of 52 spacers by attempting to make deletion mutants of 26 *Anabaena* genes using a CRISPR-Cpf1 system, the editing efficiencies varied greatly depending on the gene that was targeted, ranging anywhere from 0% to 100% efficiency (Niu et. al., 2018). The results displayed in this paper with Cpf1 in *N. punctiforme* were similar

in that they gave varying levels of success depending on the gene that was targeted by the CRISPR-Cpf1 system.

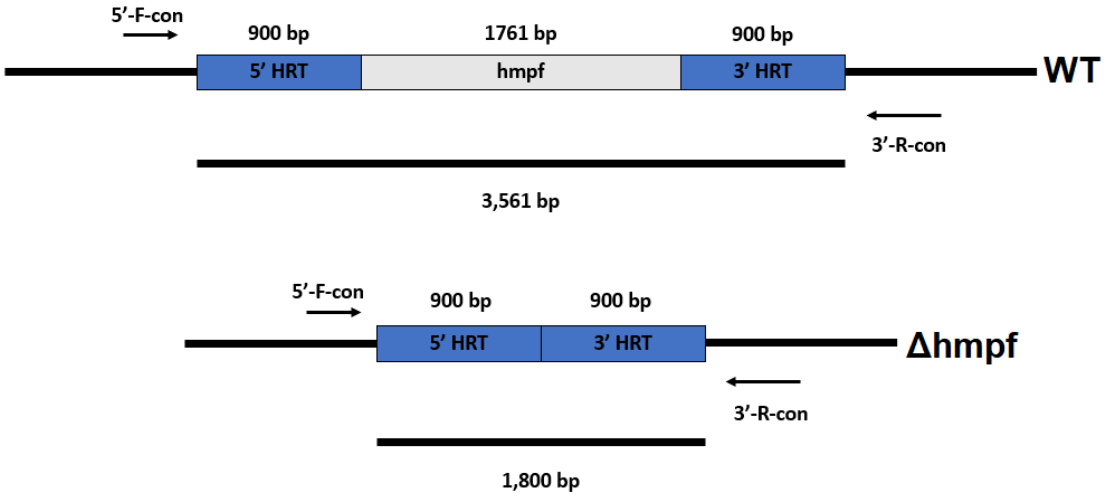

Figure 5. A gene map is shown for the relative positions of the confirmation primers when performing a colony PCR to confirm a mutant. Δ*hmpF* is used as the example mutant in this diagram. **WT:** This diagram shows a gene map for a wild-type colony. The *hmpF* gene is shown in grey, with two blue boxes representing the HRT regions on either side of the *hmpF* gene. The length of each region is displayed above the respective box. The 5'-F-con and 3'-R-con primers are shown as arrows, pointing in their respective directions. The overall length of the region of interest is given below. **Δ*hmpF*:** This diagram shows a gene map for a Δ*hmpF* colony. The *hmpF* gene is no longer present, while the two blue boxes still represent the HRT regions. The length of each region is displayed above the respective box. The 5'-F-con and 3'-R-con primers are shown as arrows, pointing in their respective directions. The overall length of the region of interest is given below.

In Figure 5, two gene maps, one for a wild-type colony and one for an expected mutant Δ*hmpF* colony, are shown displaying the regions of interest. In a colony PCR to confirm the presence of a mutant, the confirmation primers, 5'-F-con and 3'-R-con, would anneal outside of the HRT regions away from the gene of interest. The confirmation primers would not anneal to

the HRT itself on the editing plasmid: they were designed to anneal just outside of the HRT regions. The 5'-F-con primer would anneal next to the 5' HRT region away from the gene of interest, while the 3'-R-con primer would anneal next to the 3' HRT region away from the gene of interest. Both primers would cause the polymerase enzyme to amplify DNA towards the gene of interest, as indicated by the arrows. In this example, a colony that retained *hmpF* would produce a PCR product of 3,561 bp, while a colony with *hmpF* deleted would produce a PCR product of 1,800 bp.

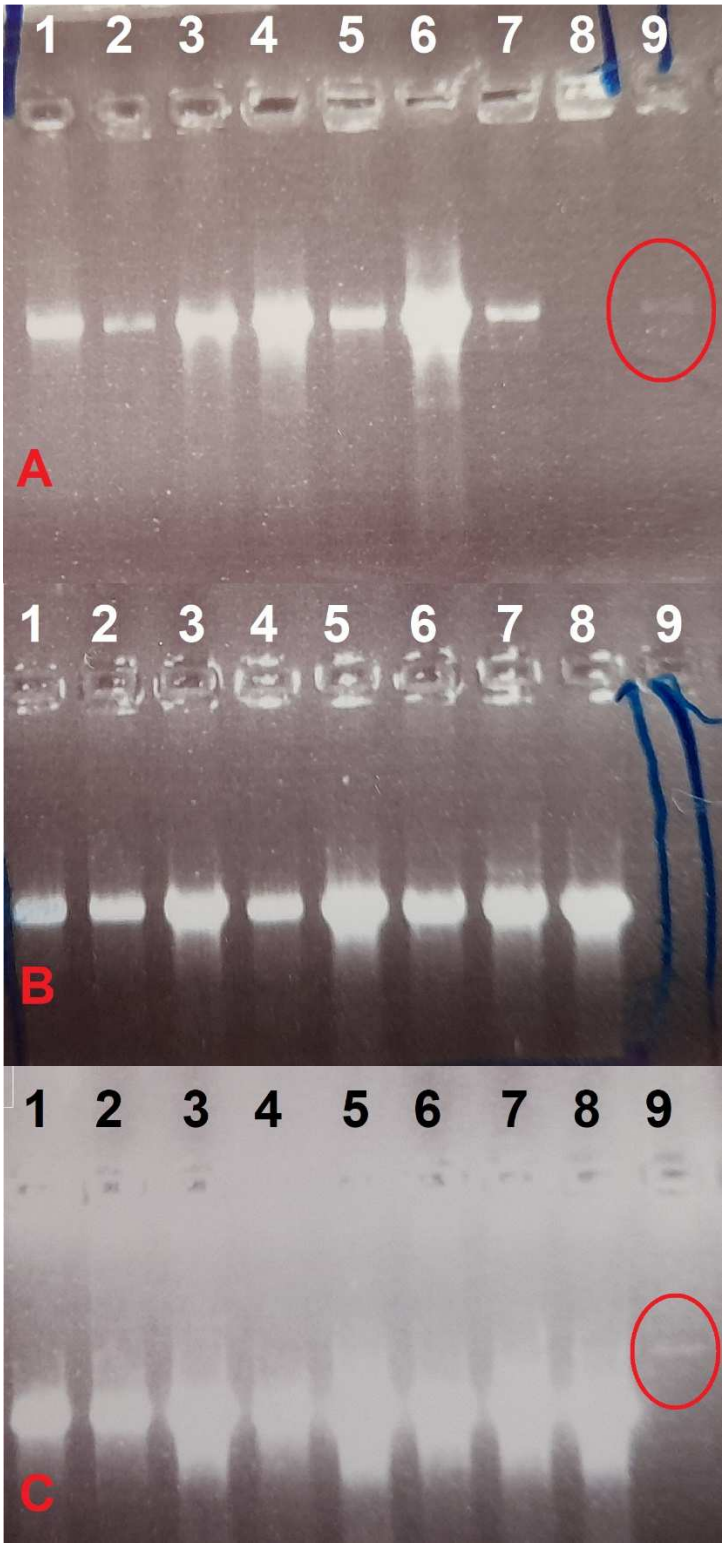

Figure 6. A gel is run after performing colony PCR to confirm mutants for Δ3486 and Δ*hmpF*. The gene-specific 5'-F-con and 3'-R-con primers listed in Table 1 are used for confirmation. **A.** In lanes 1-8, the colony PCR results are shown for Δ3486

(Figure 6 Continued)
using the 5'-F-con and 3'-R-con primers for Δ3486. In lane 9, a negative control is run for Δ3486, circled in red, using wild-type *N. punctiforme*. **B.** In lanes 1-8, the colony PCR results are shown for Δ*hmpF* using the 5'-F-con and 3'-R-con primers for Δ*hmpF*. In lane 9, a negative control is run for Δ*hmpF*, using wild-type *N. punctiforme*. **C.** This is the same image as the one shown in B but at a higher exposure. The negative control band in lane 9 is visible within the red circle; it was previously more difficult to see in the image taken in B.

Figure 6 gives an example of colony PCR. Mutant colonies were visibly verified for Δ*hmpF*, but none were verified for Δ3486. 8 colonies from each were used to perform colony PCR using the gene-specific 5'-F-con and 3'-R-con primers listed in Table 1. A negative control is also run for each gene-knockout, with wild-type *N. punctiforme* substituting the mutant colony. In Figure 6A, bands are visible for the first 7 lanes of colony PCR for Δ3486. They all match up in length with the band for the wild-type negative control in lane 9, confirming that the colonies for Δ3486 are not actually mutants. In Figure 6B, bands are visible for all 8 lanes of colony PCR for Δ*hmpF*. The image in 6B is retaken in 6C at a higher exposure, so that the band in the negative control lane could be seen, circled in red. The wild-type band is clearly larger than the bands for the colony PCR of the Δ*hmpF* colonies, indicating a successful confirmation of mutants for Δ*hmpF*. The low levels of PCR product for the wild-type are most likely due to inefficient amplification of a larger product.

Table 3
Targeted Genes and Presence of Mutants

Gene ID	Number of Colonies	Mutants (Y/N)
hmpf-1	143	Y
hmpf-2	279	Y
pilB-1	0	N/A
pilB-2	0	N/A
hfq-1	32	Y
hfq-2	21	Y
456-1	0	N/A
456-2	0	N/A
459-1	200	Y
459-2	580	Y
875-1	3	Y
875-2	0	N/A
3486-1	368	N
3486-2	956	N

Of the 7 genes targeted in Table 3, only 4 of the genes were successfully mutated by the plasmids: *hmpF*, *hfq*, 459, and 875. Successful mutations were verified visibly, since mutants would lack the motility phenotype, and by colony PCR, where mutants would be lacking the gene of interest in contrast with wild-type chromosomal DNA. In general, there appears to be a correlation between the number of colonies and the success of the CRISPR system in forming mutants, with plasmids producing less colonies successfully resulting in genome editing, while those that produce more colonies do not. The colonies of Δ459 were not initially mutants, but after several rounds of re-streaking these colonies onto plates containing neomycin, mutant colonies were eventually isolated. Some plasmids, such as Δ3486 and the initial plates of Δ459, produced many colonies, but no mutants. This is most likely due to the gRNA being inefficient, and thus ineffectively leading the Cpf1 protein to the target sites. The plates with systems aiming to knockout the genes *pilB*, 456, and 875 all initially yielded no colonies. Across the

span of 1 to 2 weeks, the plates with Δ875 eventually yielded a few colonies that were verified to be mutants.

Lack of Colonies Indicates Different Possible Issues with the CRISPR-Cpf1 System

The Δ*pilB* and Δ456 plates continued to yield no colonies in repeated conjugation attempts, indicating a significant issue with the CRISPR-Cpf1 system. Three hypotheses were developed to address the issue. One possible hypothesis was that off-target editing is occurring. With the rise in popularity of targeted mutagenesis as a genome editing tool, concerns have also been raised about off-target effects causing collateral damage in the genome (Cho et. al., 2014). A second hypothesis is that there is a gene in the HRT region that, when expressed, produces a protein that is lethal to the cells. And the final hypothesis is that the homologous repair mechanism of the *N. punctiforme* is inefficient. Unlike CRISPR-Cas9 systems, CRISPR-Cpf1 systems would continue to perform targeted cutting in such a way that non-homologous end joining would not be able to prevent further cutting. Homologous recombination is crucial to the repair of DNA when there are any breakages in the genome, and if the efficiency of the homologous repair falls short of the efficiency of the Cpf1 protein-gRNA complex, the DNA will remain broken and lead to the death of the cell (Li et. al., 2008).

Homologous Recombination is a Limiting Factor in Cpf1-based Genome Editing in *N. punctiforme*

In order to distinguish between the first two hypotheses and the third, the plasmid containing the CRISPR-Cpf1 system targeting *pilB* was introduced into wild-type *N. punctiforme* and a previously generated Δ*pilB* mutant. If off-target effects were the main issue preventing colonies from growing for the plasmid targeting *pilB*, neither the wild-type nor the Δ*pilB* mutant would be able to grow, since the off-target effects would continue to damage the genome and

prevent cell growth. If a gene on the HRT was expressing a lethal protein, neither wild-type nor the *pilB* mutant would be able to grow, since the lethal protein would still be expressed in both strains. However, if many colonies are present for the Δ*pilB* mutant but no colonies are present for the wild-type, this would most likely indicate inefficient homologous repair. This is because the CRISPR-Cpf1 system would have no *pilB* gene to cut in the Δ*pilB* mutants, leading to colonies growing on the Δ*pilB* plates, but the system would cut the *pilB* gene in the wild-type cells. If the third hypothesis is true that the homologous repair is inefficient, the wild-type cells would die due to the unrepaired DNA breakage.

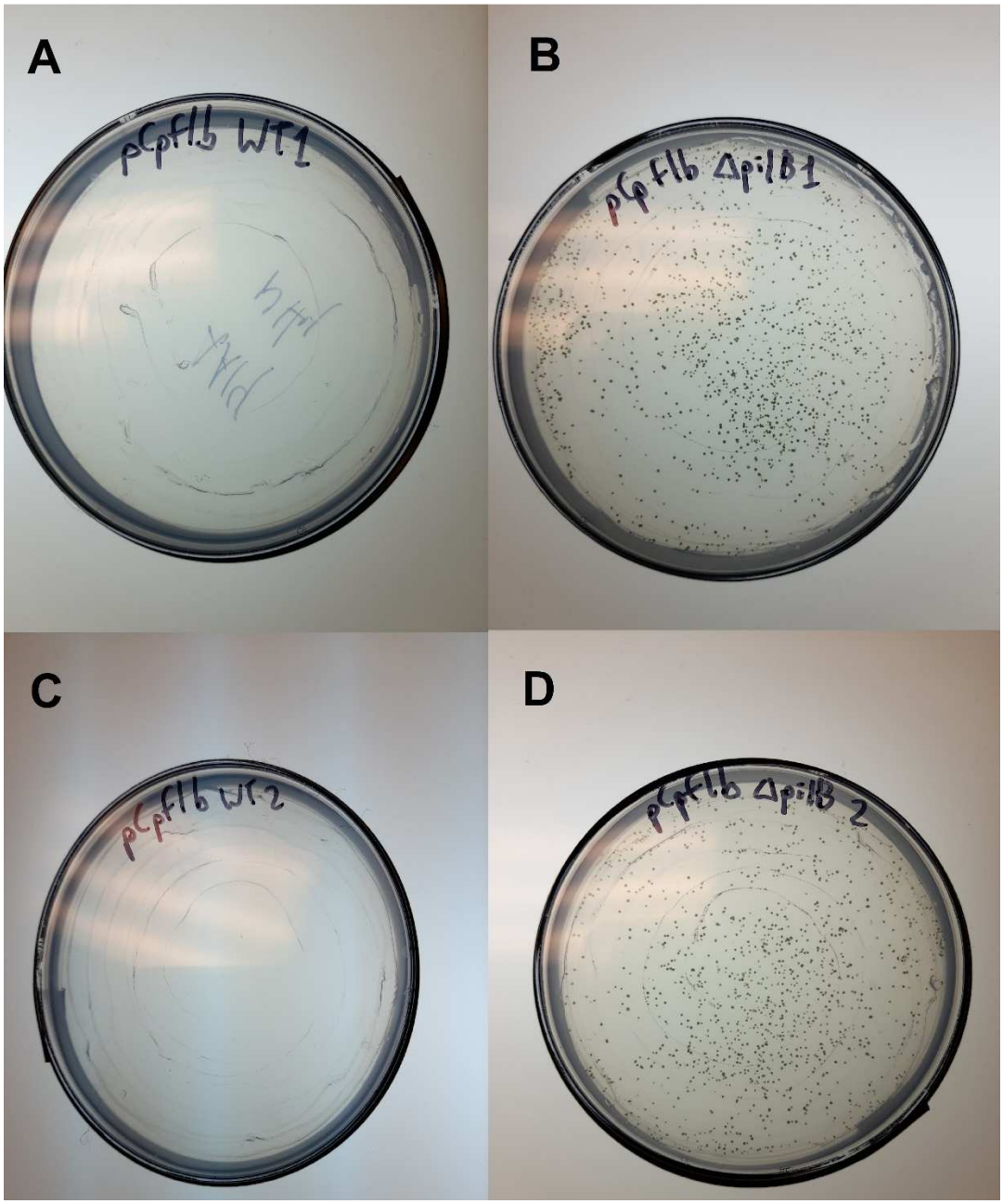

Figure 7. Plates comparing the results of transforming the plasmid containing the CRISPR-Cpf1 system into wild-type *N. punctiforme* cells, and the results of transforming the plasmid into Δ*pilB* mutants. **A.** Picture showing that no colonies grew for wild-type *N. punctiforme* cells on AA/4 + NH_4^+ MOPS plates. **B.** Picture showing that colonies grew for Δ*pilB* mutants on AA/4 + NH_4^+ MOPS plates. **C.** Picture showing that no colonies grew for wild-type *N. punctiforme* cells on AA/4 + NH_4^+ MOPS plates. **D.** Picture showing that colonies grew for Δ*pilB* mutants on AA/4 + NH_4^+ MOPS plates.

Two plates were grown for wild-type *N. punctiforme* cells that had the CRISPR-Cpf1 system targeting *pilB* conjugated into them, and two plates were grown for Δ*pilB* mutants that had the same system conjugated into them, as shown in Figure 7. The results show that no colonies grew for the wild-type, while many colonies grew for the Δ*pilB* mutants. This therefore indicates that inefficient homologous repair is responsible for the Δ*pilB* plasmid being unable to produce colonies.

Improving the Homologous Repair Template

There were two approaches taken to improving the homologous repair capability of the CRISPR-Cpf1 system. The first approach focused on making a high-copy version of the Δ*pilB* plasmid. In addition to the pCpf1b plasmid backbone that was previously used, and is considered to be a low copy plasmid, pAM504, which has a higher copy number oriV, was used to form new plasmids to test the idea that having more copies of the Δ*pilB* plasmid would improve the homologous repair capability of the CRISPR-Cpf1 system due to the increased availability of the HRT.

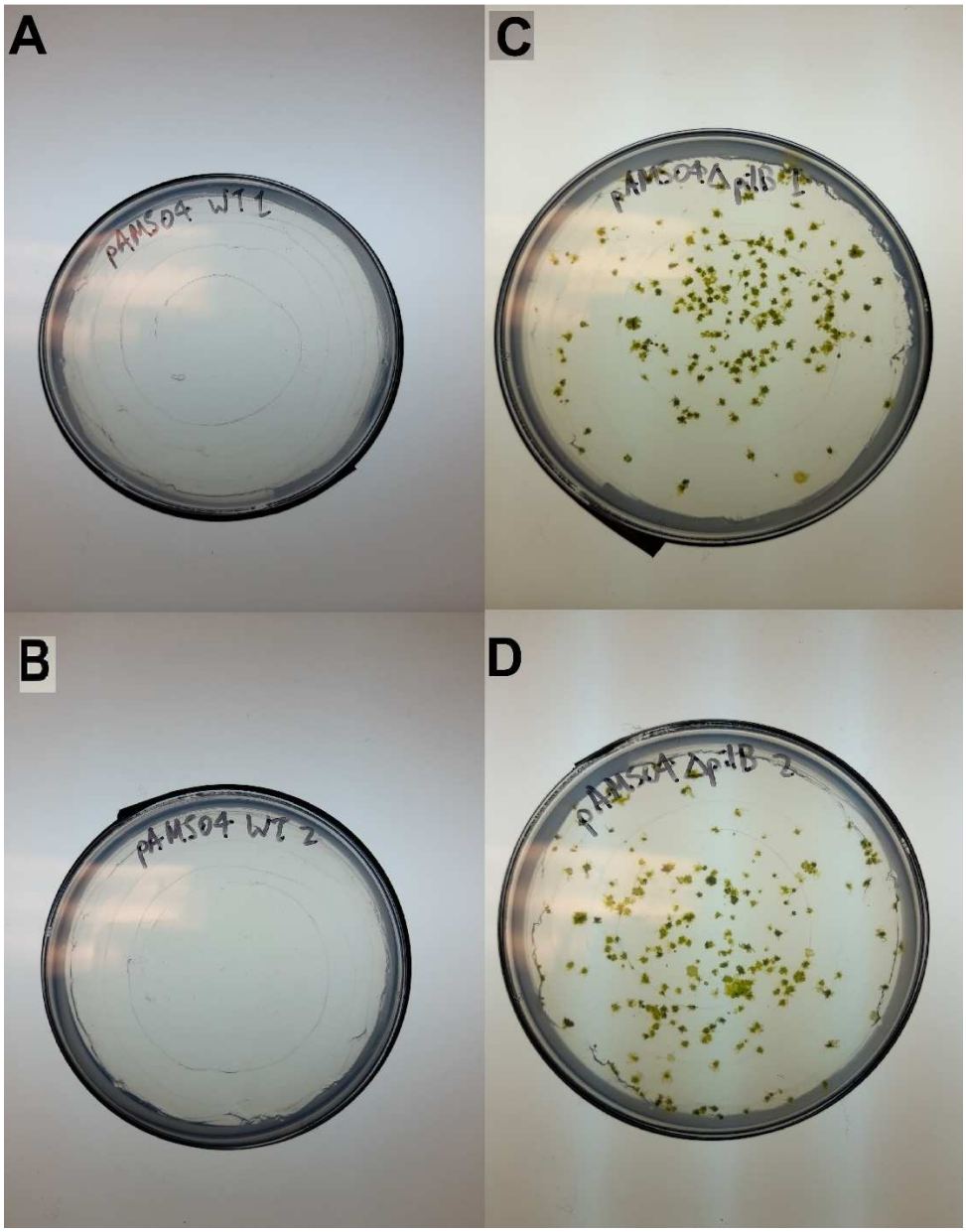

Figure 8. Wild-type *N. punctiforme* cells and Δ*pilB* mutants are grown on AA/4 + NH_4^+ MOPS plates after having CRISPR-Cpf1 plasmids containing the pAM504 backbone transformed into them. **A-B.** Picture showing that no colonies grew for wild-type *N. punctiforme* cells with pAM504 plasmids transformed into them on AA/4 + NH_4^+ MOPS plates. **C-D.** Picture showing that colonies grew for Δ*pilB* mutants with pAM504 plasmids transformed into them on AA/4 + NH_4^+ MOPS plates.

The cells containing the plasmids were plated onto AA/4 + NH_4^+ MOPS agar, and allowed to grow for several days. As shown in Figure 8, plates A and B showed that no colonies grew for wild-type *N. punctiforme* cells containing the pAM504 plasmid while plates C and D showed that colonies grew for Δ*pilB* mutants containing the pAM504 plasmid, indicating that the solution was not to introduce a higher-copy version of the Δ*pilB* plasmid.

This result led us to pursue the second approach, which was to make a larger HRT region. This approach was based on the premise that a larger HRT would increase the efficiency of homologous repair. While the original plasmids had HRT regions that were 900bp in length, new primers were ordered to form new, larger HRT regions 2kb long. This plasmid was then conjugated into *N. punctiforme* cells.

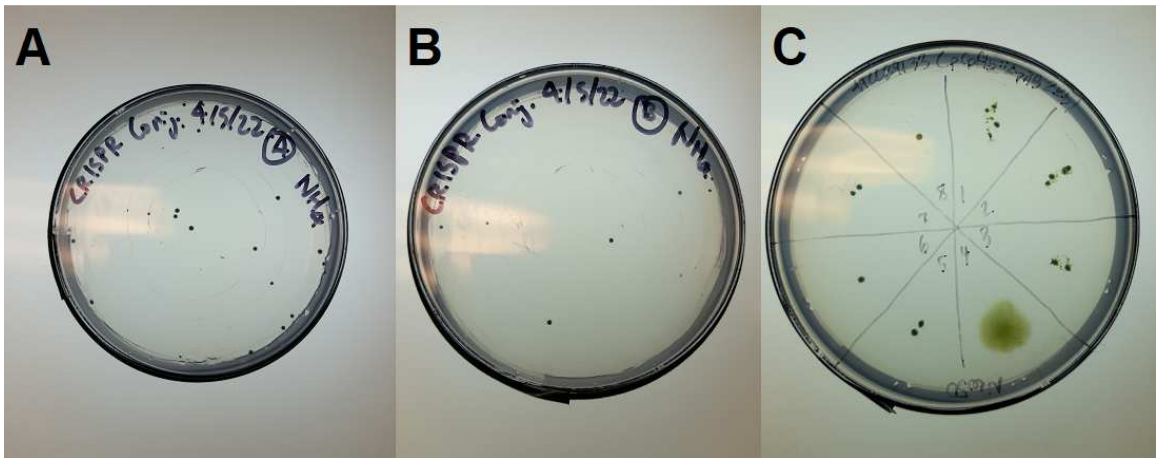

Figure 9. AA/4 + NH_4 plates used to grow *N. punctiforme* colonies containing the new CRISPR construct aiming to knockout the gene *pilB*. **A-B.** Colonies are shown to have grown for both plates used to grow *N. punctiforme* cells containing the Δ*pilB* plasmid. **C.** 8 colonies from plates A and B were picked onto an AA/4 + NH_4 plate containing neomycin (50 µg/mL) to more clearly visualize mutants. Colonies 1-3 and 5-8 were visibly mutants, with no hormogonium development, while colony 4 clearly grew hormogonia.

The *N. punctiforme* cells were plated onto two AA/4 + NH_4 plates (Figure 9). After the colonies were allowed to grow, the plates were inspected for the presence of mutant colonies. Colonies grew on both plates, and were later visibly verified to be mutants by the lack of hormogonia. 8 colonies total were picked from both plates onto an AA/4 + NH_4 plate containing neomycin (50 mg/mL), and 7 out of the 8 failed to show signs of colony spreading, indicating they are likely to be *pilB* mutants. This indicated that lengthening the HRT region helped to resolve the issue with the mutated colonies not growing in the previous experiments.

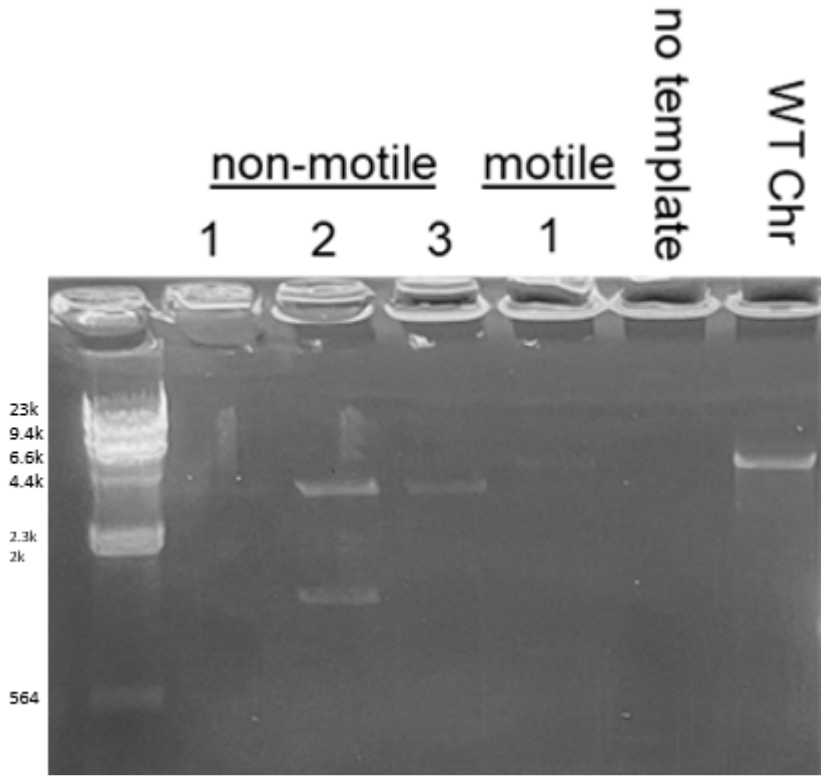

Figure 10. Colony PCR is performed on 4 of the 8 Δ*pilB* colonies, including 3 non-motile colonies and 1 motile colony. **Non-motile:** 3 lanes are shown, each representing a non-motile colony from the Δ*pilB* colonies. **Motile:** This lane represents the colony that was visibly verified to be motile. **No template:** This lane represents the negative control, where no template was added to the PCR mixture. **WT Chr:** This lane represents the positive control, where wild-type chromosomal DNA was added to the PCR mixture.

A colony PCR was done using 4 of the 8 colonies that were picked from the conjugation, and run on a gel (Figure 10). 3 of the 7 non-motile colonies were used, and the 1 motile colony also underwent colony PCR. A negative control consisting of the same PCR mixture but without template was run on the gel, and a negative control with wild-type chromosomal DNA instead of a colony was also run on the gel. The expected size for the mutant colonies was around 4kb, and the expected size for the wild-type colonies was 6kb. The results show bands of the appropriate size for the non-motile colonies, indicating that they are indeed mutants. A second, smaller band is shown in lane 2 of the non-motile colonies that is most likely an artifact, resulting from the primers annealing to an unintended site on the DNA and leading to an additional PCR product. A faint band is seen for the motile colony, matching up with the length of the wild-type chromosomal DNA, confirming that the motile colony was not mutated. This final colony PCR confirms the overall success of the new Δ*pilB* plasmid.

CHAPTER 4: DISCUSSION

The results presented in this study indicate that the CRISPR-Cpf1 system that was designed and tested here is effective in knocking out targeted genes in *N. punctiforme*. Initially, only 3 of the 7 genes were successfully knocked out, and after re-streaking some of the motile colonies, 4 of the 7 plasmids were confirmed to have mutated the target gene. Evidence for this success is shown in the presence of mutants lacking hormogonium development in the colonies that were grown after the system was transformed into the cyanobacteria.

Two major limitations were observed based on the initial results. The first involved the Δ3486 colonies and the initial colonies for Δ459: many colonies grew on the plates, but there were no mutants. This was most likely due to the gRNA being ineffective or inefficient, and thus not properly guiding the Cpf1 protein to the target site to cut the gene of interest. A potential solution would be re-designing the CRISPR system with two different gRNAs that target the same sequence. This two-spacer strategy has been shown to be highly effective at knocking out genes in the *Anabaena* genus of cyanobacteria (Niu et. al., 2018).

The second issue was recognized after noticing that the Δ*pilB* and Δ456 plasmids did not produce any colonies, leading to the hypothesis that the gRNA was targeting the sites in the chromosome other than the target gene, and thus cutting the DNA at a location that was not getting repaired by the HRT. The plasmid containing the CRISPR construct was introduced into both a Δ*pilB* mutant strain of *N. punctiforme* and wild-type *N. punctiforme* to test this hypothesis, and colonies grew for the Δ*pilB* mutant strain while none grew for the wild-type strain. This indicated that the issue with the CRISPR system was not off-target editing in this particular case, since the off-target editing should have been equally as lethal to both the Δ*pilB*

mutant strain and the wild-type strain. In future applications of this system, if there is a problem with off-target editing by the gRNA, a potential solution would be designing multiple gRNAs. This way, if there is concern that a gRNA is off-target editing, there are other gRNAs that we can incorporate instead into the plasmid to properly guide the Cpf1 protein to the target gene. After verifying that the issue was not off-target editing, the conclusion was that although the gRNA may be editing efficiently, the HRT region was simply not large enough, and thus that the homology repair enzyme system was not able to repair the genome at the same rate that the Cpf1 protein was cutting the target sequence. As a potential solution, a larger, 2kb HRT region was designed and assembled into the CRISPR system as a substitute for the previous, smaller HRT region that was 900 bp long. The new system was employed into *N. punctiforme* with the aim of knocking out the gene *pilB*, an attempt which had previously resulted in no colonies growing. The mutated cells ended up growing into colonies, indicating that the new CRISPR system with a larger HRT region was successful at both cutting the gene and repairing the genome so that the cells would survive. Homology-directed repair has been an integral component in an effective CRISPR system, but has worked at varying efficiencies depending on the design of the HRT (Zhang et. al., 2017). In future applications of this system, employing larger regions for the HRT seems to be an effective, potential solution. However, while the application of the larger HRT region in the Δ*pilB* plasmid was successful, to further verify the legitimacy of this solution, it will need to be tested with multiple genes, and with various sizes of the larger HRT region to determine the limitations of this approach.

Designing and developing successful CRISPR systems in different species of cyanobacteria has significant benefits for understanding the genome of the cyanobacteria in the future. Cyanobacteria have been a topic of interest for several reasons, including their ability to

produce biofuels. Resources have been poured into attempts to genetically engineer cyanobacteria to be commercially-relevant biofuel-producing agents in the past several years (Savakis et. al, 2014). Because CRISPR is a powerful tool for genome editing, creating a system that is able to intentionally mutate a species opens up many windows of possibility for engineering the species in a desired way, such as genetically modifying a species of cyanobacteria to produce biofuels more efficiently. Additionally, the success of the CRISPR system designed in this paper gives greater precedent to the notion that this system may be effective in related species of *N. punctiforme* as well, since some components of this system were based off of a system that was effective in *Anabaena*, a genus that is closely related to *Nostoc*.

REFERENCES

Cai Y.P., Wolk C.P. (1990). Use of a conditionally lethal gene in Anabaena sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. Journal of Bacteriology, 172(6), 3138-3145. Doi: 10.1128/jb.172.6.3138-3145.1990

Campbell E.L., Summers M.L., Christman H., Martin M.E., Meeks J.C. (2007). Global Gene Expression Patterns of *Nostoc punctiforme* in Steady-State Dinitrogen-Grown Heterocyst-Containing Cultures and at Single Time Points during the Differentiation of Akinetes and Hormogonia. Journal of Bacteriology. Doi: 10.1128/JB.00360-07

Cho S.W., Kim S., Kim Y., Kweon J., Kim H.S., Bae S., Kim J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Research, 24(1), 132-141. Doi: 10.1101/gr.162339.113

Cho Y.W., Gonzales A., Harwood T.V., Huynh J., Hwang Y., Park J.S., Trieu A.Q., Italia P., Pallipuram V.K., Risser D.D. (2017). Dynamic localization of HmpF regulates type IV pilus activity and directional motility in the filamentous cyanobacterium *Nostoc punctiforme*. Molecular Microbiology, 00(00), 00-00. Doi: 10.1111/mmi.13761

Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature, 471(7340), 602-607. Doi: 10.1038/nature09886

Gong C., Huang S., Song R., Qi W. (2021). Comparative Study between the CRISPR/Cpf1 (Cas12a) and CRISPR/Cas9 Systems for Multiplex Gene Editing in Maize. Agriculture, 11(5), 429. Doi: 10.3390/agriculture11050429

Harwood T.V., Zuniga E.G., Kweon H., Risser D.D. (2021). The cyanobacterial taxis protein HmpF regulates type IV pilus activity in response to light. PNAS, 118(12), e2023988118. Doi: 10.1073/pnas.2023988118

Henson B.J., Hesselbrock S.M., Watson L.E., Barnum S.R. (2004). Molecular phylogeny of the heterocystous cyanobacteria (subsections IV and V) based on *nifD*. International Journal of Systematic and Evolutionary Microbiology 54, 493–497. Doi: 10.1099/ijs.0.02821-0

Jiang W., Brueggeman A.J., Horken K.M., Plucinak T.M., Weeks D.P. (2014). Successful transient expression of Cas9 and single guide RNA genes in Chlamydomonas reinhardtii. Eukaryotic Cell, 13(11), 1465-1469. Doi: 10.1128/EC.00213-14

Khayatan B., Bains D.K., Cheng M.H., Cho Y.W., Huynh J., Kim R., Omoruyi O.H., Pantoja A.P., Park J.S., Peng J.K., Splitt S.D., Tian M.Y., Risser D.D. (2017). A Putative O-Linked β- N-Acetylglucosamine Transferase Is Essential for Hormogonium Development and Motility in the Filamentous Cyanobacterium Nostoc punctiforme. Journal of Bacteriology, 199(9), e00075-17. Doi: 10.1128/JB.00075-17

Khayatan B., Meeks J.C., Risser D.D. (2015). Evidence that a modified type IV pilus-like system powers gliding motility and polysaccharide secretion in filamentous cyanobacteria. Molecular Microbiology, 98(6), 1021-1036. Doi: 10.1111/mmi.13205

Li X., Heyer W.D. (2008). Homologous recombination in DNA repair and DNA damage tolerance. Cell Research, 18(1), 99-113. Doi: 10.1038/cr.2008.1

Meeks J.C., Campbell E.L., Summers M.L., Wong F.C. (2002). Cellular differentiation in the cyanobacterium Nostoc punctiforme. Archives of Microbiology, 178, 395-403. Doi: 10.1007/s00203-002-0476-5

Moraes L.E., Blow M.J., Hawley E.R., Piao H., Kuo R., Chiniquy J., Shapiro N., Woyke T., Fadel J.G., Hess M. (2017). Resequencing and annotation of the *Nostoc punctiforme* ATTC 29133 genome: facilitating biofuel and high-value chemical production. AMB Express 7, 42. Doi: 10.1186/s13568-017-0338-9

Niu T.C., Lin G.M., Xie L.R., Wang Z.Q., Xing W.Y., Zhang J.Y., Zhang C.C. (2018). Expanding the Potential of CRISPR-Cpf1-Based Genome Editing Technology in the Cyanobacterium Anabaena PCC 7120. ACS Synthetic Biology, 18;8(1), 170-180. Doi: 10.1021/acssynbio.8b00437

Savakis P., Hellingwerf K.J. (2014). Engineering cyanobacteria for direct biofuel production from CO2. Current Opinion in Biotechnology, 33, 8-14. Doi: 10.1016/j.copbio.2014.09.007

Savic N., Ringnalda F., Lindsay H., Berk C., Bargsten K., Li Y., Neri D., Robinson M.D., Ciaudo C., Hall J., Jinek M., Schwank G. (2018). Covalent linkage of the DNA repair template to the CRISPR-Cas9 nuclease enhances homology-directed repair. eLife 7:e33761. Doi: 10.7554/eLife.33761

Su T., Liu F., Gu P., Jin H., Chang Y., Wang Q., Liang Q., Qi Q. (2016). A CRISPR-Cas9 Assisted Non-Homologous End-Joining Strategy for One-step Engineering of Bacterial Genome. Scientific Reports, 6, 37895. Doi: 10.1038/srep37895

Ungerer J., Pakrasi H.B. (2016). Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. Scientific Reports, 6, 39681. Doi: 10.1038/srep39681

Wong F.C.Y., Meeks J.C. (2002). Establishment of a functional symbiosis between the cyanobacterium Nostoc punctiforme and the bryophyte Anthoceros punctatus requires

genes involved in nitrogen control and initiation of heterocyst differentiation. Microbiology, 148(1), 315-323. Doi: 10.1099/00221287-148-1-315

Zetsche B., Gootenberg J.S., Abudayyeh O.O., Slaymaker I.M., Makarova K.S., Essletzbichler P., Volz S., Joung J., Oost J., Regec A., Koonin E.V., Zhang F. (2015). Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. Cell, 163(3), 759-771. Doi: 10.1016/j.cell.2015.09.038

Zhang J.P., Li X.L., Li G.H., Chen W., Arakaki C., Botimer G.D., Baylink D., Zhang L., Wen W., Fu Y.W., Xu J., Chun N., Yuan W., Cheng T., Zhang X.B. (2017). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biology, 18, 35. Doi: 10.1186/s13059-017-1164-8