DELFT UNIVERSITY OF TECHNOLOGY

Generating CRISPR-dCas9 System In *Clostridium difficile* Bacteria

Shabnam Hossein-Javaheri
Faculty Mechanical, Maritime and Materials Engineering
Department of Biomedical Engineering
February 2016

"It is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change."

Charles Darwin

DELFT UNIVERSITY OF TECHNOLOGY

Abstract

Shabnam Hossein-Javaheri
Faculty Mechanical, Maritime and Materials Engineering
Department of Biomedical Engineering
February 2016

The increasing incidence of *Clostridium difficile*-associated disease, and its resistance to a number of antibiotics, highlight the need for additional countermeasures. The main issue that hampers the design of antibiotics is a lack of appropriate tools to study the molecular biology of *C.difficile* in great detail. Insight into the function of DNA replication genes and the protein they encode might eventually lead to the development of specific inhibitors that can be exploited as drugs. By means of CRISPR-dCas9 system, we designed and engineered a construct that is capable of targeting specific replication genes and silence those genes. Our design is composed of a nuclease-deactivated Cas9 (dCas9) protein and a customized single-guided RNA (sgRNA) with a 20-basepair complementary region to target replication genes in *C.difficile*. Co-expression of dCas9 and sgRNA can efficiently block transcription by interfering with DNA replication process. We developed a set of conjugative plasmids that carries appropriate Pveg-sgRNA targets to repress CD0001 (dnaA) and CD1214 (spo0A) genes in *C.difficile*. We also built a construct with dCas9 and its inducible promoter (Pcpr) that can be conjugated into *C.difficile*. Once both constructs are transformed into *C.difficile* 630Δerm, we will be able to analyze and detect potential phenotypes expressed in the presence of dCas9.

Acknowledgements

Though only my name appears on the cover of this thesis report, a great many people have contributed to its production. This project consumed huge amount of work, research and dedication. Still, implementation would not have been possible if I did not have a support of many individuals. Therefore, I would like to extend my sincere gratitude to all of them.

First and foremost I wish to thank my supervisor, Dr. Wiep Klass Smits. Despite my lack of microbiology background, he invested his trust and belief in me. I have been amazingly fortunate to have an adviser who gave me the freedom to explore on my own, and at the same time the guidance to recover when my steps faltered.

I am grateful to Dr. Jeroen Corver for his encouragement and tremendous help during this project. I am also thankful to him for reading my reports, commenting on my views and helping me understand and enrich my ideas.

I would like to acknowledge Annemieke Friggen and Tim Dalebout for their assistance with lab works and being great friends. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship and I deeply appreciate their belief in me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

There is a prevailing belief that human beings dominate earth when it comes to cleverness and we always proudly wonder why have other species failed to evolve human-like intelligence. Our privilege is reflected by our innovations that changed the world and our lives forever. We launched rockets into space, we designed smart phones to transmit understandable signals, we created a virtual reality called the Internet, and we even made weapons to assassinate our own kind. But when it comes to the co-existence of human and prokaryotic microorganisms, the story appears to be different.

The population of microbial communities associated with the human body is known to be at least equal to the total number of our somatic and germ cells. The adult human intestine is home to an inconceivable number of microorganisms [1]. Many bacterial species live in our teeth, skin, and respiratory tracts. The question then arises: are humans an advanced form of bacteria? Evolutionary biologists generally agree that humans are descended from bacteria-like ancestors but branched off about two billion years ago [2].

The human-bacteria interactions is far from a mutualistic relationship. On the one hand bacteria play an influential role in maintaining human health and the global ecosystem, but on the other hand they can cause infectious diseases. It is about both the companionship and the competition for survival.

We develop antibiotics to treat and prevent bacterial infection and as a comeback, bacteria become resistant to antibiotics due to a permanent alteration of the nucleotide sequence in the bacteriums DNA, called “Mutations”. A bacterium with a mutation can survive the antibiotic and reproduces millions more with the same resistance within the space of a day [3]. Other than by descent in which information travels through the generations, horizontal gene transfer, or movement of genetic material between bacteria

is another way to spread genes that encode for antibiotic resistance. Horizontal gene transfer is a highly significant phenomenon and a dominant force that drives prokaryote evolution. Perhaps this can explain why bacteria have the ability to develop antibiotic resistance so quickly [4].

1.1 The crisis of antibiotic resistance

Considering how quick bacteria can acquire antibiotic resistance compare to a decade of human dedication making a specific one, we should reassess our current approach in combating infections. Improper use of antimicrobial medicines creates a favorable opportunity for resistant microorganisms to emerge, spread and persist [5]. Upon the underused, overused or misused of antibiotics, the process of antibiotic resistance is increased. The longer the duration of antibiotic exposure, the greater the risk of developing drug resistance.
Another consequence of inappropriate consumption of antibiotics is the disruption of a healthy equivalent microbiota in the human body. Under imbalanced circumstances, harmful bacteria become capable of colonizing and initiating infection. One of the worst intestinal infections is caused by a bacterium called *Clostridium difficile*. The overgrowth of bacteria results in release of toxins that attack the lining of the intestines, causing a condition called *Clostridium difficile* colitis. The incidence of *C.difficile* infections (CDI) epidemic has recently been estimated to be 6.0 per 10.000 patient days with a total in-hospital mortality of CDI cases of 12% [6].

1.1.1 The gut bacterium *Clostridium difficile*

Clostridium difficile, a spore-forming, toxin-producing anaerobic bacterium, is the leading cause of nosocomial diarrhea in the developed world. Colonization of humans by *C.difficile* can cause symptoms such as diarrhea, colitis, pseudomembranous colitis, and death. The main factors influencing the chance of developing *C.difficile* infections include old age, antibiotic use, and long-term hospitalization. In addition, CDI has also been increasingly found outside hospitals, raising concerns that it is rather a complicated infectious disease. CDI associated economic burdens are estimated to be $4.8 billion/ year in the United states and €3 billion/year in EU, including healthcare, nursing home, and indirect socio-economic costs.
C.difficile strains are capable of surviving in unfavorable conditions such as an environment without oxygen by forming dormant spores. A spore is a rounded resistant form of *C.difficile* that allows the bacterium to survive for long amounts of time in environments with limited availability of nutrients. The spores are resistant to heat, low pH, and can

survive for as long as five months [7]. The capability of *C.difficile*to sporulate poses a challenge for hospitals.

1.2 Changing perspective: A new look at old problems

Regarding the current crisis of antimicrobial resistance, it is essential to develop novel methods for the treatment and prevention of CDI. To answer the question of how a different perspective leads to a creative solution, we need to understand more about essential *C.difficile* genes, such as DNA replication genes and what happens in the cell when they start colonizing.
The duplication of genetic material during growth is an essential process, hence the replication genes are crucial for the existence of bacteria. By repressing or abrogating their expression, either the bacterial growth stops or the cell dies. For that reason, gaining an insight into the function of DNA replication genes might eventually lead to the development of specific inhibitors and new generation of drugs.

1.2.1 Biologically inspired engineering

Biologically inspired engineering is a new scientific discipline that applies biological principles to develop new engineering solutions for medicine, industry, and the environment. Bio-inspired engineering involves deep exploration into the way that living cells and organisms grow, protect, attack, and adapt to their environment. Nature offers much to observe that can spark ideas; we discover smart constructions, unusual mechanisms and clever sensing of the biological organisms, which introduce pathways to bio-inspired designs. One of the mechanisms that has gained considerable attention in the field of genetic engineering is bacterial and archaeal defense strategies against foreign viral infections, the so called CRISPR-Cas immune system.

1.3 Overview of CRISPR-Cas9 system

Clustered regularly interspaced short palindromic repeats (CRISPR) together with Cas (CRISPR-associated) are essential genes in adaptive immunity in bacteria and archaea, enabling the organisms to recognize and eliminate foreign genetic materials. These repeats were initially discovered in the 1980s in *E. coli* [8], but their function was unknown until 2007 by Barrangou, who demonstrated that the *S. thermophilus* can store a genome fragment of an infectious virus into its CRISPR locus [9]. CRISPR array consist of short and highly conserved DNA repeats, usually 21 to 48 bp, repeated up to 250 times [10].

Repeats are separated by similar length of nucleotides known as spacers. New spacers can be added rapidly to decode viral or plasmid DNAs (Figure 1.1).

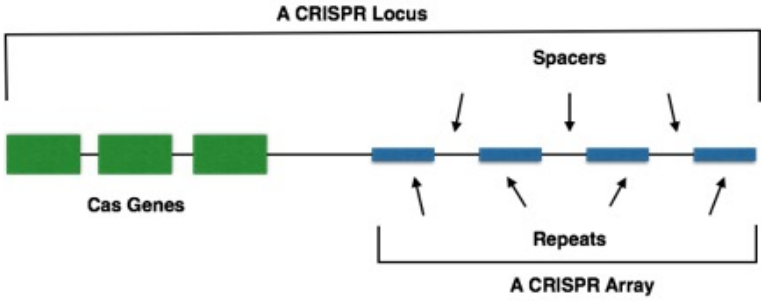

FIGURE 1.1: CRISPR locus in a bacterial chromosome

A CRISPR array is transcribed as a single RNA transcript (pre-crRNA) and generates small CRISPR RNAs (crRNAs), each containing a spacer and repeat fragments [11]. These crRNAs bind to their partially complementary piece of RNA called trans-activating RNA (tracrRNA) [12]. The Cas endonuclease complexed together with the comibnation of a crRNA/tracrRNA snips DNA at sites containing a 20-nucleotide crRNA complementary sequence (Figure 1.2).

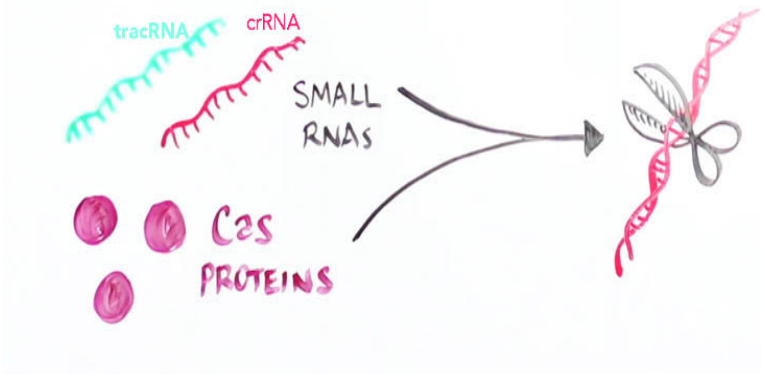

FIGURE 1.2: An overview to CRISPR-Cas9 system

Scientist then wondered if they could engineer one RNA molecule that can possibly mimic the structure of the crRNA and tracRNA bound together that would guide Cas to cut DNA in a specific location. Therefore, based on the type II CRISPR system and its signiture Cas gene, dCas9, a simplified two-component system by combining trRNA and crRNA into a single synthetic single guide RNA (sgRNA) was developed. The sgRNA recruits Cas9 proteins and detects foreign nucleic acids by complementary base pairing. Once a foreign genetic code is matched the current sgRNA sequence, the targeted nucleic acids are cleaved. Introducing mutations in catalytic domains of Cas9 protein results in the production of a catalytically inactive Cas9, or dead Cas9 (dCas9)

protein capable of repressing the gene expression in a site-specific manner (Figure 1.3) [13].

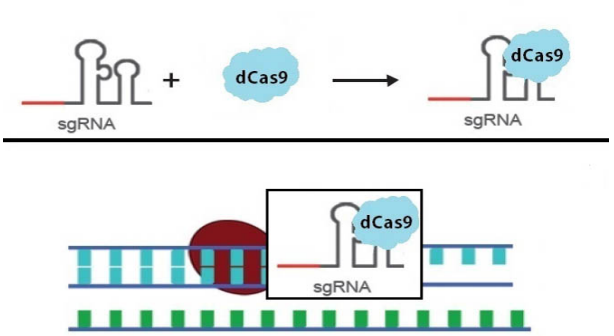

FIGURE 1.3: Top: sgRNA and dCas9 bind together. Bottom: dCas9 pairedwith sgRNA interfere with transcription and terminate the expression of the target gene.

1.3.1 Silencing replication genes with CRISPR-dCas9 system

In order to stop bacterial growth, we need to silence responsible genes that are encoding essential proteins, such as those involved in DNA replication. Bacteria reproduce by binary fission. During this process, a single cell bacterium divides into two identical daughter cells, and that requires a complete copy of genetic information. Most of the genetic information is stored on a circular bacterial chromosome that contains a circular DNA molecule. DNA replication proceeds in three enzymatically catalyzed and coordinated steps: initiation, elongation and termination.

The replication of chromosomal DNA in many types of bacteria is initiated at a specific sequence of nucleotides called *oriC*. Such a fixed initiation point is a key factor in the control of chromosome replication in most bacteria. Initiation of replication involves the binding of the replication initiator (DnaA), loading of the replicative helicase at the origin, and assembly of the rest of the DNA synthesis machinery to recruit the replisome [14, 15]. Loading of helicase is done differently by different organisms. In *B.subtilis*, three proteins, DnaB, DnaD and DnaI load helicase at *oriC*. The assembly of the helicase-loading complex at *oriC* in *B.subtilis* starts with binding of DnaA to *oriC* followed by collaboration of DnaD with *oriC*. During the second stage, DnaB associates with *oriC* through a direct interaction with DnaD [15]. Once this complex is assembled, another complex made of DnaI-helicase together with DnaC interacts with DnaB, and therefore helicase is assembled (Figure 1.4).

As we explained before, initiation of replication and assembly of the helicase depend on DnaA, DnaD, DnaB, and DnaI in *B.subtilis*. DnaA associates with *oriC* independently of any of the other replication initiation proteins. DnaD with *oriC* was dependent on DnaA, but not on DnaB or DnaI, and association of DnaB with *oriC* was dependent

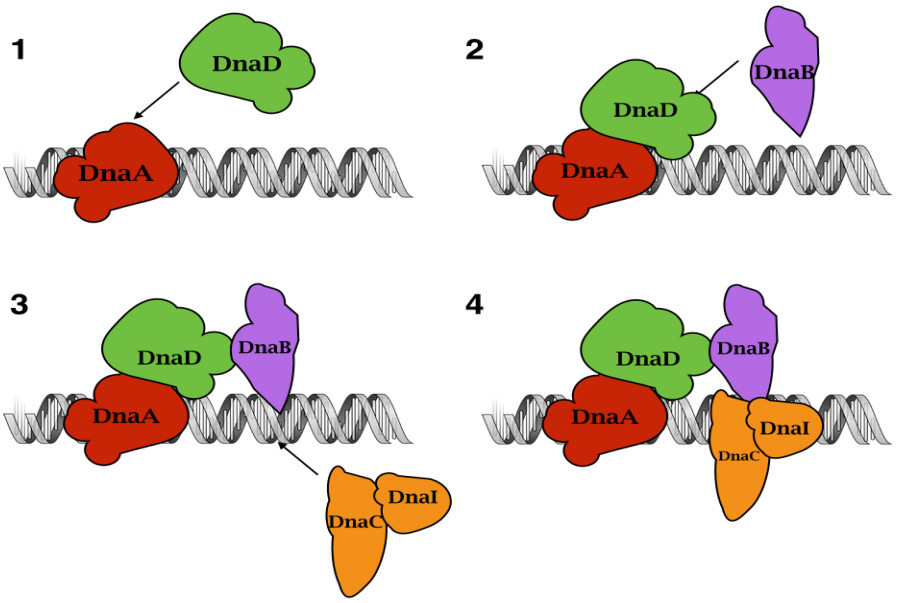

FIGURE 1.4: Schematic of four stage model for the association of helicase and helicase loader with *oriC*

on DnaA and DnaD, but not on DnaI. Therefore, targeting different replication genes may inhibit the growth in different ways. For example, by targeting DnaA, we can only prevent a new replication from taking place. In other words, if we target DnaA after the replication being initiated, we won't be able to stop the progress immediately, but we can only avoid future initiations. If we target DnaC at any point of initiation, the process will be stopped almost instantaneously. This is because when there is no active helicase, the unwinding of DNA strands will not occur and the replication process remains incomplete.

The CRISPR-dCas9 system is a proper starting point for designing our novel method and silencing replication genes in *C.difficile*. However, implementation of new constructs in this bacteria is complex. *C.difficile* is a strict anaerobe and extremely sensitive to even a low amount of oxygen in the environment, therefore it is notoriously challenging to work with. In order to avoid potential difficulties, we initially started investigating the functionality of CRISPR-dCas9 system in *Bacillus subtilis*, which is a genetically tractable bacterium. Moreover, *B.subtilis* has a close homology to *C.difficile*. Homology describes the relationship between genes and how they are inherited from ancestors. A homologous gene is a gene inherited in two species by a common ancestor. In this case, *B.subtilis* and *C.difficile* replication genes descent from a common ancestral DNA sequence as show in (Table 1.1). We showed that the CRISPR-dCas9 system can be used to inhibit the replication genes and to partially stop bacterial growth in *B.subtilis* (Figure 1.5).

***B.subtilis* protein**	***C.difficile* protein**	**Query coverage**
DnaA	CD0001	99%
DnaB	Not Identified	-
DnaC	CD3657	94%
DnaD	CD3653	47%
DnaI	CD3654	46%

TABLE 1.1: Homologs of *B. subtilis* replication proteins can be identified in *C.difficile*

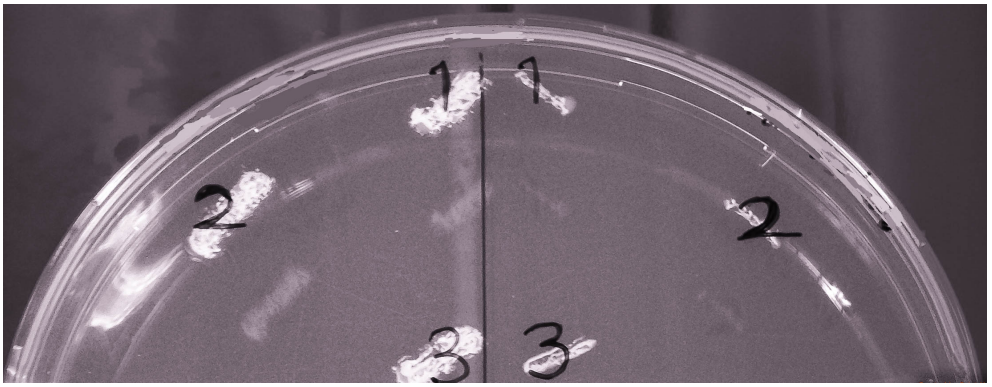

FIGURE 1.5: Left: Inactive CRISPR-dCas9 show regular bacterial growth (thicker lines). Right: Activated CRISPR-dCas9 (thinner lines) signify partial growth inhibition.

1.4 Project goals

DNA replication, the duplication of genetic information during growth, is an essential process for all living organisms. The main goal of this project is to engage the CRISPR-dCas9 system, design a construct that is capable of targeting specific replication genes, and silent those genes. The functionality of the CRISPR-dCas9 has already been confirmed in *B.subtilis*. Since the genetic organizations responsible for activating and deactivating genes in *C.difficile* differs from those of *B. subtilish*, the new design should be adjusted in such away that it can be recognized and transcribed by responsible genetic factors in *C.difficile*. For that matter we need to:

- Design sgRNA to target the sequence of genes of interest in *C.difficile*
- Make a dCas9 construct that is expressible in *C.difficile*

Chapter 2

Materials and Methods

2.1 Design and cloning of sgRNAs

We explained earlier that repression by CRISPR/dCas9 system depends on base pairing between a short segment of the sgRNA and the DNA target, therefore, new DNA targets can be specified simply by modifying the sgRNA sequence. The sgRNA design consists of 102nt chimeric non-coding RNA of which a 20-nt base pairing region, followed by a 42nt Cas9-binding, and a 40-nt transcription terminator (Figure 2.1). This construct has previously been made at the Gross lab at UCSF and it was shown that pJMP2 plasmid, carrying a constitutive (Pveg) promoter, can produce the sgRNA.

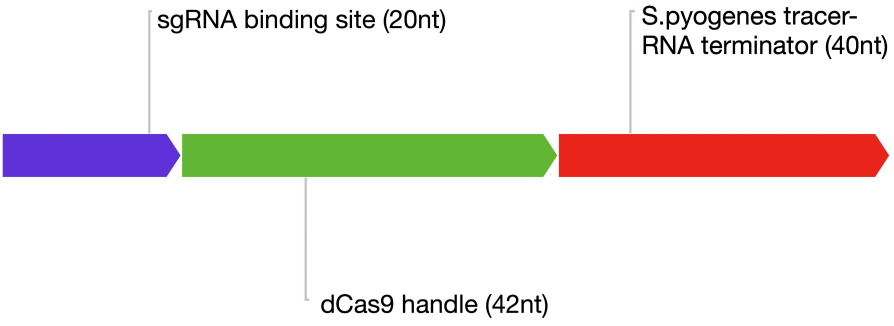

FIGURE 2.1: Schematic of 102nt chimeric non-coding RNA:a 20-nt base-pairing region (purple), a 42-nt dCas9 binding handle (green), and a 40-nt *S.pyogenes* transcription terminator (red).

The original protocol for efficient guide RNA design is done algorithmically by Jason Peters at USCF. In this section, we provide instructions on sgRNA design and plasmid construction for repression of the gene(s) of interest.

2.1.1 Design of sgRNAs for targeted gene repression

The 20-nt base pairing region can be freely customized, allowing broad flexibility in targeting genes for repression. For better repression efficiency, the base pairing region should bind to the non-template DNA strand of the coding region for the gene to be repressed. If the template DNA strand of the coding sequence is chosen in the design, it will result in no or mild repression [16]. In addition to the base-pairing of sgRNA-DNA, the CRISPR-(d)Cas9 system is based on an NGG or NAG (N as any nucleotide) protospacer adjacent motif (PAM). Thus, the selection of targetable sites within the genome is determined by the 20-nt region and NGG PAM motif [16].

Example:
In order to target a site on the template DNA strand, search for N_{20}-N(G/A)G. N_{20} is the sequence of the base pairing region of sgRNA; to target a site on the non-template strand, search for C(C/T)N-N_{20}, and the reverse complementary sequence of N_{20} is the sequence of the sgRNA base pairing region [16]. For example, we want to target the following sequence: **AGACCGCTAACTGAAAGTT**. The sgRNA base pairing sequence is the reverse complementary: **AACTTTCAGTTTAGCGGTCT** (see Figure 2.2).

FIGURE 2.2: An example sequence for designing an sgRNA. The PAM sequence GGG and the base pairing sequence are shown on both non-template and template DNA strands

2.1.2 Mutagenesis PCR

Out of 102-nt desired chimeric non-coding RNA in the pJMP2 plasmid, the dCas9 handle, and transcriptional terminator remains unchanged and only the 20-nt of the sgRNA binding site must be modified via mutagenesis PCR. The primers must be designed in such a way that the final PCR product matches the entire sequence of our initial plasmid except the 20nt of sgRNA binding site (Figure 2.3).

In order to generate a new sgRNA expression vector we need:

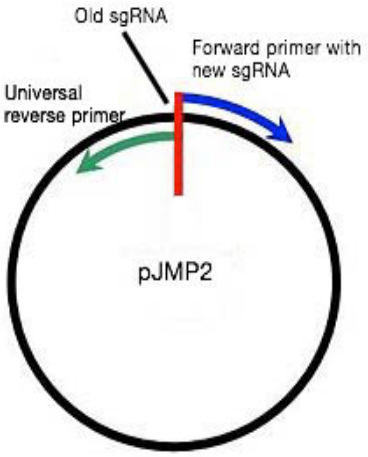

FIGURE 2.3: Schematic of primer binding in the process of sgRNA mutagenesis PCR

- *E. coli* sgRNA expression plasmid, **pJMP2** with ampicillin resistance
- Customized forward sgRNA primer, which targets a replication gene in *C.difficile*
- Universal reverse primer, **oSgr001**

For this project, we selected 4 different genes in *C.difficile* for the design of sgRNA (Table 2.1). However, we continued with oSgr004 for the purpose of this study since silencing CD3657 is expected to lead to a fast and complete growth inhibition .

Primer		Targeted gene
oSgr001	ACATTTATTGTACAACACGAGC	Reverse Universal
oSgr002	GTCACCTTTTATTAATTGTAGTTTTAGAGCTAGAAATAGCAAG	CD0001
oSgr003	TCATTAAGATGTATTTTTAAGTTTTAGAGCTAGAAATAGCAAG	CD0001
oSgr004	TCTGATTCTACACTATGAGGGTTTTAGAGCTAGAAATAGCAAG	CD3657
oSgr005	GCCTCTTTATAGAAATCATCGTTTTAGAGCTAGAAATAGCAAG	CD3657
oSgr006	ACATCTAGTATTAATAAGTCGTTTTAGAGCTAGAAATAGCAAG	CD1214
oSgr007	ACACCTAATCCATCTAGATGGTTTTAGAGCTAGAAATAGCAAG	CD1214
oSgr008	AAATATATCTATAAAGATATGTTTTAGAGCTAGAAATAGCAAG	CD3653
oSgr009	ACGTATAACCCGTCCGCCATGTTTTAGAGCTAGAAATAGCAAG	CD3653

TABLE 2.1: List of primers designed for this project

The mutagenesis PCR on sgRNA is done as the following:

1. Use **Q5** High-Fidelity DNA Polymerase with an error rate 100-fold lower than that of Taq DNA Polymerase and 12-fold lower than that of Pyrococcus furiosus (Pfu) DNA Polymerase. Since the pJMP2 plasmid is approximately 7.5kbp, Q5 High-Fidelity DNA Polymerase is ideal for cloning and can be used for long or difficult amplicon. Amplify new sgRNA vectors using Q5 polymerase:

2. Perform PCR using the following thermal cycler program:

10mM dNTPs	1µL
5X Q5 Buffer	10µL
10µg Reverse Primer	0.5µL
10µg Forward Primer	0.5µL
Q5 polymerase	0.5µL
Template DNA	0.05µL
ddH_2O	37.45µL
Total	50µL

Step	**Temperature°C**	**Duration**
Initial denaturation	98	30 seconds
25 cycles	98	10 seconds
	65	30 seconds
	72	4.20 minutes
Final extension	72	2 minutes

3. Purify completed PCR reaction by removal of remaining dNTPs, primers, Taq, and Mg ion.

4. Treat the purified PCR products with DpnI to degrade the PCR template plasmid.

DpnI restriction endonuclease	0.5µL
10 CutSmart Buffer	5µL
Purified PCR product	25µL
ddH_2O	19.5µL
Total	50µL

Incubate at 37°C for 1 hour. Heat inactivate at 65°C for 20 minutes.

5. Phosphorylate the blunt-end for the next step. For ligation to occur, at least one of the DNA ends (insert or vector) should contain a 5' phosphate. Primers are usually supplied non-phosphorylated; thus, the PCR product will not contain a 5' phosphate. To phosphorylate use:

100ng/µL *E. coli* sgRNA plasmid template	40µL
T4 PNK	0.5µL
10XT4 Ligation Buffer	1µL
ddH_2O	8.5µL
Total	50µL

Incubate the mixture at 37°C for 30 min, and then heat-inactivate the reaction at 65°C for 20 min.

6. Circularize the PCR products by ligating the ends. Add 1.25µL T4 ligase to the mixture and leave it overnight at room temperature (16°C).

7. Transform 10µL of the ligation reaction to 100µL *E.coli* DH5α competent cell/DNA mixture.

8. Spread onto Lysogeny broth (LB) agar plates containing **ampicillin** (100µg/mL) and incubate overnight at 37°C.

9. Isolate plasmid from overnight *E.coli* cultures using Nucleospin Plasmid Quickpure Miniprep kit according to the instructions of the manufacture.

10. Perform Sanger sequencing with primers oWKS-1551 and oWKS-1544 to confirm the sequence of the insert.

2.1.3 Cloning Pveg-sgRNA into pSMB47 using SphI site

The next step would be amplifying the new sgRNA binding site with its related promoter (Pveg), dCas9 handle, and *S.pyogenes* tracer-RNA terminator from the pJMP2 plasmid. To do so, the primers used in this section are designed in such a way that the final PCR product carries SphI site and can be cloned into pSMB47 (Figure 2.4). This vector has its unique SphI site and it is appropriate for conjugal transfers (see Appendix A).

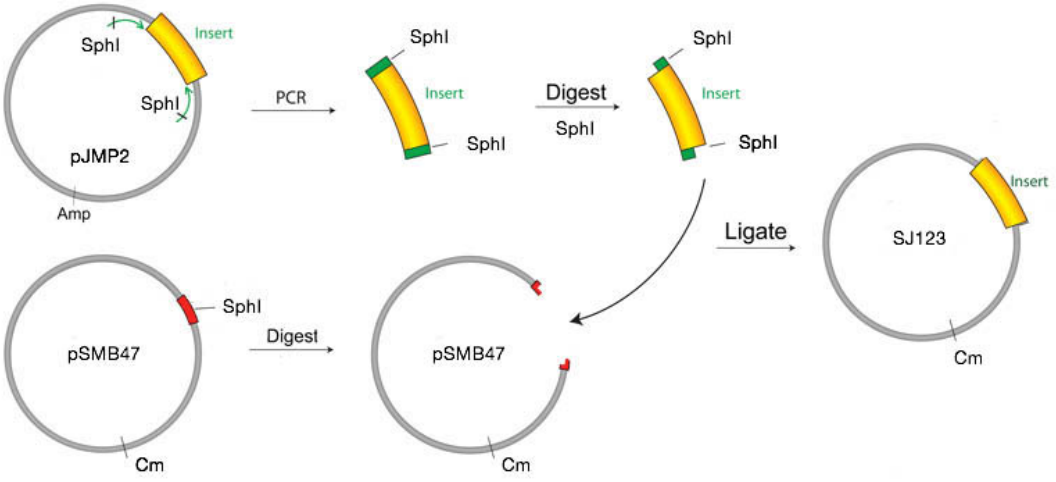

FIGURE 2.4: Schematic of sgRNA cloning by PCR into pSMB47

The direct transfer of pSMB47/sgRNA to *C.difficile* is done via conjugation. Conjugation is the process by which one bacterium transfers DNA in forms of plasmids to another bacterium through direct contact. During conjugation, one bacterium is counted as the donor of the genetic material, and the other serves as the recipient.

In order to have a successful conjugation, only conjugative plasmids are capable of initiating conjugation because they contain a set of transfer or tra genes which promote sexual conjugation between different cells [17]. pSMB47 vector integrates in Tn916 in *Bacillus subtilis* strain BS49 and the integrated element can be transferred into the chromosome of *C.difficile* [18].

The cloning process of sgRNA is done as the following:

1. Use **Accuzyme** DNA Polymerase or other High-Fidelity DNA Polymerases, **oWKS-1544** and **oWKS-1551** primes for this section. Dilute the template 10 times if the

concentration is greater than 100ng/µL.

10mM dNTPs	5µL
10X Accuzyme Buffer	5µL
10µg Reverse Primer	1µL
10µg Forward Primer	1µL
Template DNA	1µL
Accuzyme polymerase	2µL
ddH_2O	35µL
Total	50µL

2. Perform PCR using the thermal cycler program as the following:

Step	**Temperature°C**	**Duration**
Initial denaturation	95	3 minutes
25 cycles	95	15 seconds
	65	15 seconds
	72	1.40 minutes
Final extension	72	2 minutes

3. Place the final product on the electrophoresis gel. The expected fragment is about **660bp** that has Pveg-sgRNA-rrnB terminators in it.

4. If the result of step 3 is as expected, purify the PCR product using **GeneJETGel** Extraction kit according to the instructions of the manufacture.

5. Digest 25µL of the washed product with 0.5µL SphI enzyme, 5µL **CutSmart** buffer and 19.5µL double-distilled water. Incubate at 37°C for 1 hr.

6. Recover the fragment by gel-isolation **GeneJETGel** Extraction kit according to the instructions of the manufacture.

7. Digest 25µL of pSMB47 plasmid with 0.5µL SphI enzyme, 5µL **CutSmart** buffer and 19.5µL double-distilled water. Incubate at 37°C for 1 hour.

8. Dephosphorylate to ensure that the vector does not re-circularize during ligation. Add 1µL **Shrimp Alkaline Phosphatase (SAP)** to the mixture and incubate at 37°C for 1 hour. Heat Inactivate at 65°C for 5 minutes.

9. Set up ligation with a titration of the reaction as the following:

Ratio	1:0	1:1	1:5	1:8
Vector	1µL	1µL	1µL	1µL
Insert	0µL	1µL	5µL	8µL
T4 DNA Ligase Reaction Buffer	5µL	5µL	5µL	5µL
T4 DNA Ligase	0.5µL	0.5µL	0.5µL	0.5µL
ddH_2O	8.5µL	7.5µL	3.5µL	0.5µL
Total	15µL	15µL	15µL	15µL

Leave overnight in room temperature (16°C).

10. Transform 10µL of ligation reactions to 100µL *Ecoli* DH5α competent cells.

11. Spread onto LB agar plates containing **chloramphenicol** (20µg/mL) and incubate overnight at 37°C.

12. If there are transformants on the pSMB47/sgRNA ligation, perform colony PCR to confirm the existence of the insert in the colonies. Use **oWKS-1376** and **oEVE-58** primers for this section. In the case of empty colonies, a 526bp band is expected. Otherwise, the expected band including the insert is 1.2kb.

13. If the result of the previous step is positive, perform Sanger sequencing with the same set of primers to confirm the sequence of Pveg-sgRNA.

2.2 Design and Cloning of dCas9

As explained earlier, sgRNAs recruit dCas9 proteins and detect foreign nucleic acids by complementary base pairing. Thus, we need to clone dCas9 into a vector that is used for conjugal transfers. We cloned the insert into pSigD vector using BamHI and PstI sites (see Appendix A). pSigD is a conjugative plasmid that carries the sigD gene under the control of the nisin-inducible promoter [19]. pJMP1 cannot be digested directly with BamHI because there are 3 BamHI sites in pJMP1 and one happened to be a part of dCas9 protein (Figure 2.5). If we cut the plasmid with BamHI, we end up with an

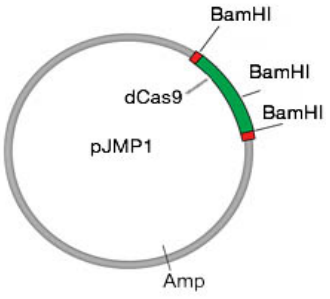

FIGURE 2.5: Schematic of pJMP1 plasmid with 3 BamHI restriction sites

incomplete version of the gene of interest. To avoid excessive digestions, we added BsaI restriction site to the primers used in this section. BsaI recognition site and its cutting site are separated, and the sequence of cutting site can be changed to suit different needs (Figure 2.6). We designed the primers in such a way that the final PCR product contains BsaI recognition site and BamHI cutting site (Figure 2.7).

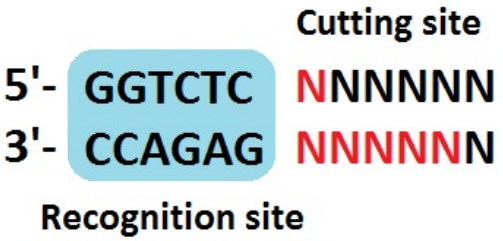

FIGURE 2.6: Schematic of how BsaI recognition site and its cutting site are separated

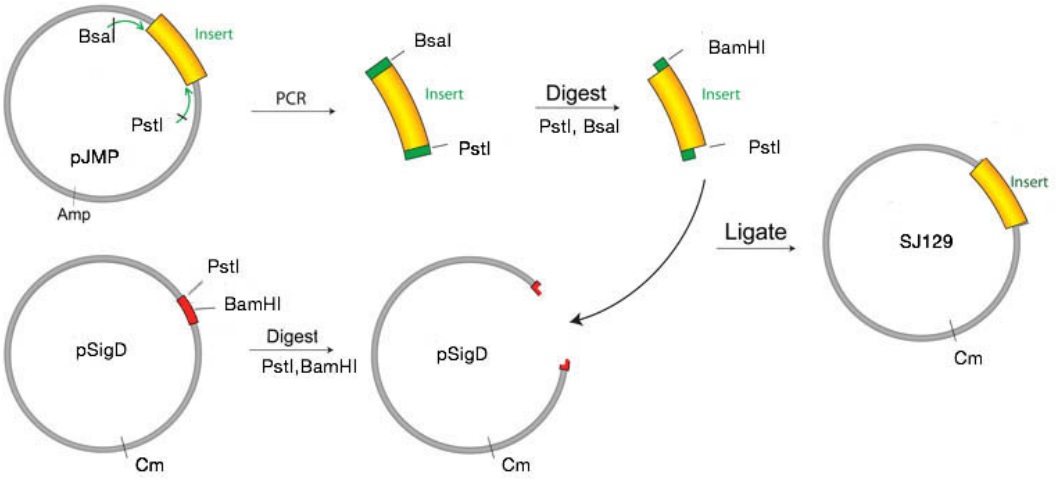

FIGURE 2.7: Schematic of dCas9 cloning by PCR into pSigD

1. Add restriction sites using PCR:
For this section we need pJMP1 template, **oWKS-1541**α and **oWKS-1542**α primes, and Q5 DNA Polymerase to perform PCR reaction. Dilute the template 10 times if the concentration is greater than 100ng/µL. Perform PCR using the thermal cycler program shown below:

Step	**Temperature°C**	**Duration**
Initial denaturation	95	3 minutes
25 cycles	95	15 seconds
	60-70	30 seconds
	72	2.30 minutes
Final extension	72	2 minutes

2. Purify PCR products using proper PCR Purification Kit following the manufacturers instructions.

3. (Optional) Treat the purified PCR products with DpnI to degrade the PCR template plasmid.

DpnI restriction endonuclease	0.5µL
10 CutSmart Buffer	5µL
Purified PCR product	25µL
ddH_2O	19.5µL
Total	50µL

Incubate at 37°C for 1 hour. Heat inactivate at 65°C for 20 minutes. f

2.2.1 Cloning dCas9 into pSigD using BamHI and PstI sites

1. Digest 25µL of the mixture with 0.5µL **BsaHI** enzyme, 0.5µL **PstI** enzyme, 5µL **1X NEBuffer 3.1** buffer, and 21µL double-distilled water. Incubate at 37°C for 1 hr.

2. Digest 25µL of pSigD plasmid with 0.5µL BamHI enzyme 0.5µL PstI enzyme, 5µL **1X NEBuffer 3.1** buffer and 21µL double-distilled water. Incubate at 37°C for 1 hr.

3. Recover the larger fragment corresponding to the DNA back-bone domain by gel-isolation **GeneJETGel** Extraction kit according to the instructions of the manufacture. The expected bands for digested pSigD are 7.5kbp and 750bp. Cut and isolate the larger band.

4. Run on an agarose gel, both your insert and vector. If the concentration of the insert is high compare to the vector, dilute the insert 10 and 5 times.

5. Set up ligation with a titration of the reaction as the following:

Ratio	1:0	1:1	1:5	1:8
Vector	1µL	1µL	1µL	1µL
Insert	0µL	1µL	5µL	8µL
T4 DNA Ligase Reaction Buffer	5µL	5µL	5µL	5µL
T4 DNA Ligase	0.5µL	0.5µL	0.5µL	0.5µL
ddH_2O	8.5µL	7.5µL	3.5µL	0.5µL
Total	15µL	15µL	15µL	15µL

Due to the fact that the insert is large (about 4.2kbp), the chances for successful a ligation are low at room temperature. To slow than the ligation process, place the tubes at **4°C** for **24-48 hours**.

6. Transform 10µL of ligation reactions to 100µL *Ecoli* DH5α competent cells.

7. Spread onto LB agar plates containing **chloramphenicol (20µg/mL)** and incubate overnight at 37°C.

8. If there are colonies on the plates, perform colony PCR to confirm the existence of the insert in the colonies. Use **oWKS-1541**α and **oWKS-1542**α primers for this section as well. The expected PCR product is 4.2kbp.

9. If the results for the previous step is positive, confirm dCas9 in the plasmid by Sanger sequencing. Include the following primers to read the entire 4.2kbp of dCas9: **OWKS-1531, oWKS-1532, oWKS1536, oWKS-1556, oWKS-1541α, oWKS-1542α, oWKS-24, and oWKS-25**.

Olgionucleotides	
oWKS-24	GTGAGCGGATAACAATTTCACACAGG
oWKS-25	GGTTTTCCCAGTCACGACGTTGTAA
oWKS-1243	GATCTGAGCTCCTGCAGTAAAGGAGAAAATT TTATGTTTGCAAAACGATTCAAAACC
oWKS-1376	GGAAGACACCCGAGGCTTGGTTATGCCGGTACTG
oWKS-1531	GCTCCCCTATCAGCTTCA
oWKS-1532	TCGCCCTTTATCCCAGAC
oWKS-1541α	CTAGGGTCTCGGATCCCTAAGGAGGCGTAGTTAATGGATAAG AAATACTCAATAGGCTTAG
oWKS-1542α	CTCGAGCTTAAGGATCCTCCGGAGACGTCGCGGCCGCGGTT
oWKS-1544	GCGCATGCTGCCGATGATAAGCTGTC
oWKS-1551	GCGCATGCGTCTGATCGGATCCTAGAAG
oWKS-1556	GCGCATGCTGCCGATGATAAGCTGTC

TABLE 2.2: List of olgionucleotides used in this study

2.3 Gene transfer into *Clostridium difficile* CD630Δerm

All manipulations of clostridia were undertaken anaerobically in anaerobic workstation and at an incubation temperature of 37°C. *E. coli* CA434 containing **dCas9** gene was transformed to *Clostridium difficile* 630Δerm strain via conjugation as previously described [20]. The selection of *C. difficile* trans-conjugants was done 3 times by subculturing on BHI-YE agar plates containing **thiamphenicol (10 mg/ml)** and *C. difficile* selective supplement **(CDSS)**.

Chapter 3

Results

3.1 Construction of pSJ123 plasmid carrying Pveg-sgRNA

As described earlier in section 2.1, the first step towards the generation of a sgRNA expression construct that targets a specific replication gene in *C. difficile* is performing a mutagenic PCR on the original plasmid carrying Pveg-sgRNA (pJMP2). This will allow us to replace the non-*C. difficile* sgRNA driven from the Pveg promoter with the sequence of the gene of interest.

3.1.1 sgRNA mutagenesis PCR

For this section, the expected PCR product is 7572bp, the same size as pJMP2 plasmid but with the modified version of 20nt sgRNA binding site. We addressed difficulties during the PCR amplification as we did not get any PCR products. The reaction was repeated few times while variables such as the ingredients in the reaction mix were replaced with fresh ones. The PCR thermal cycler program such as the extension time and annealing temperature were modified. The two primers individually with the combination of other primers were checked. We tested oSgr001 with a plasmid located primer oWKS-1243 and oSgr005 with oWKS-1544. All primers showed the expected products. Therefore, we concluded that:

- The target sequence in the template DNA is present.
- The error is not caused by the primers.

We made an assumption that there is more tendency for PCR failure when it comes to long-range PCR amplification. For that reason, we decided to digest pJMP2 with a

proper restriction enzyme and extract the 102-nt desired chimeric non-coding RNA out of the plasmid, clone it into a smaller vector such as pUC19, and repeat the PCR. The expected band sizes for BamHI digestion was around 561 base pairs and 7011 base pairs (Figure 3.1). We introduced a new construct, pSJ121 after cloning sgRNA fragment into pUC19. The PCR reaction with oSgr001/oSgr005 on pSJ121 was successful and we were able to confirm sgRNA005 by Sanger sequencing after circularizing the PCR product. sgRNA005 was cloned back into pJMP plasmid and the construct reported to be correct at the digest level, but needs to be sequence confirmed by Sanger sequencing.

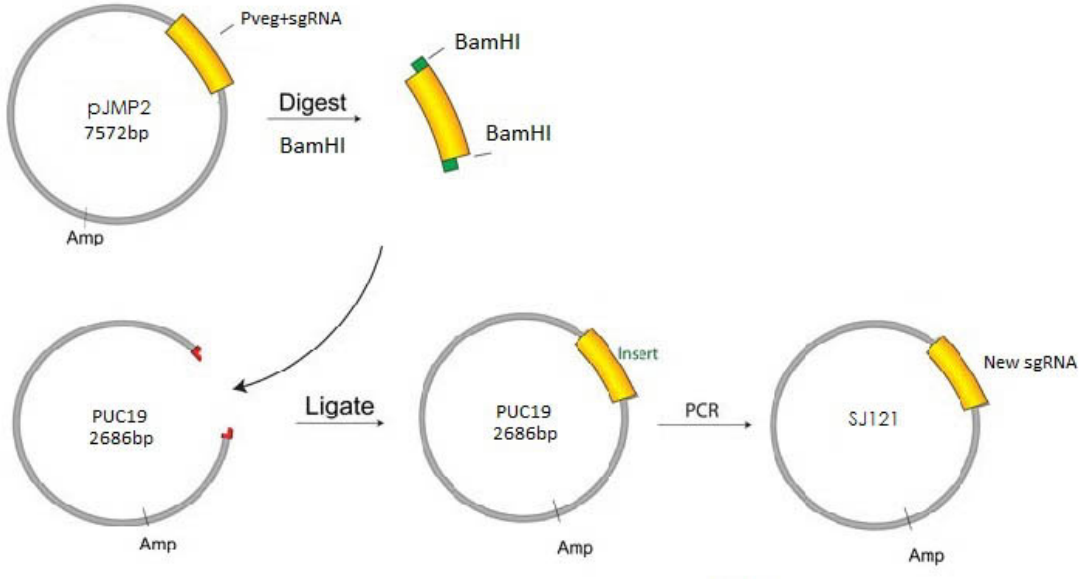

FIGURE 3.1: Schematic of the second strategy for amplifying sgRNA

In parallel to the new approach, we repeated the PCR reaction on pJMP2 plasmid with oSgr001/oSgr004. We reduced the amount of DNA template by 90% and conducted gradient PCR with the annealing temperature of 55°C-72°C. This time the PCR reaction was successful resulting in a new plasmid pSJ122. However, when we repeated the experiment with the exact same settings, the results happened to be different. It is difficult to conclude why the PCR amplification result varies under the same thermal cycler program and the same amount of ingredient. However, we can draw a conclusion that the incorrect template concentration contributed to the failure and the lengths of template DNA is crucial for a successful long PCR. We suggest that if the PCR amplification is positive for annealing temperature of 61.3°C-68.6°C for another couple of times, then perhaps for a successful non-gradient PCR, the annealing temperature can be set between 61°C and 68.5°C (Figure 3.2).

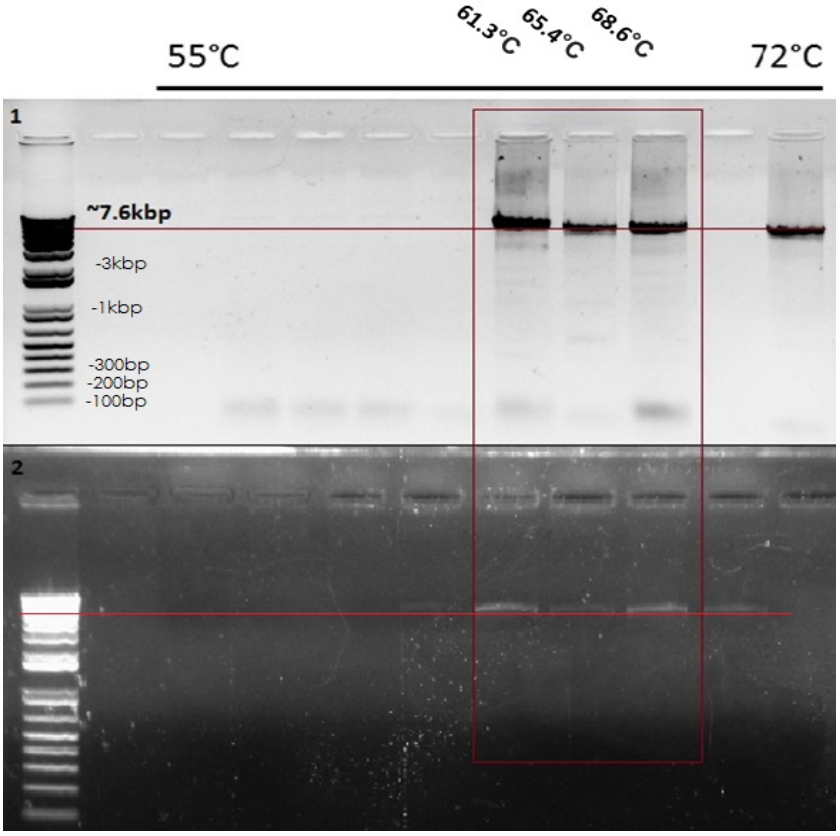

FIGURE 3.2: Electrophoresis results of sgRNA mutagenesis PCR. Different results are shown under the same PCR settings

3.1.2 Cloning Pveg-sgRNA into pSMB47

After completion of sgRNA mutagenesis PCR, the new sgRNA binding site with its related promoter (Pveg), dCas9 handle, and S.pyogenestracer-RNA terminator from the pSJ122 plasmid must be cloned into pSMB47 by means of SphI restriction site. The sequence of genes of interest lacks a SphI site, therefor it was added via PCR amplification with oWKS-1544 and oWKS-1551 on pSJ122. The results showed the expected PCR product of 660bp (Figure 3.3). Our attempts to clone Pveg-sgRNA004 into pSMB47 failed couple of times regardless of the quality and amount of components being used. Since the most common reason for failure of cloning PCR products is an incomplete digestion, we performed sub-cloning strategy in a TOPO vector. However, TOPO cloning did not contribute to our success in having a positive colony. The major problem was the excess of self-ligated colonies. There were so many transformants on the pSMB47-ligations but neither of them carried Pveg-sgRNA. After performing colony PCR on about 35 clones with primers oWKS-1376 and oEVE-58, we were able to report one positive clone with the expected size of 1.2kb (Figure 3.4). We digested the purified plasmid of the potential positive colony, and the digestion showed the correct bands. However, a missing terminal nucleotide was found in the sequencing results

of sgRNA/pSMB47 (Figure 3.5). We sequenced pSJ122 and we noticed that the the deletion occurred during the mutagenesis PCR.

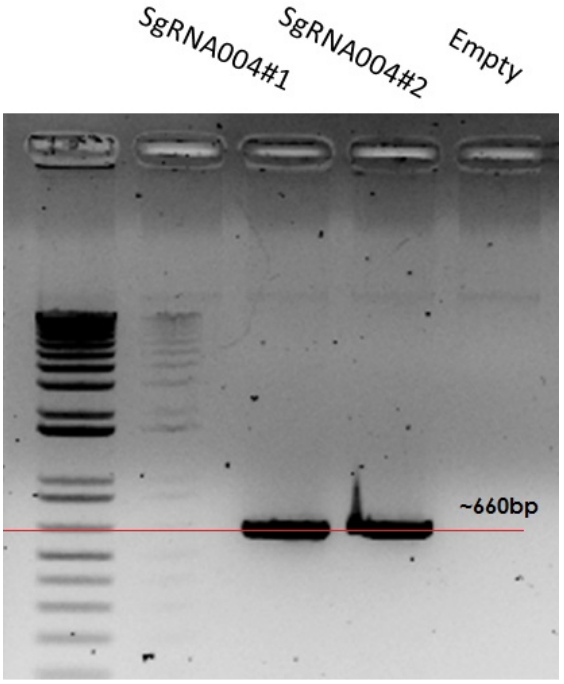

FIGURE 3.3: DNA detection by agarose gel electrophoresis. PCR results of 1% gel electrophoresis of H. pylori genotypes showing PCR results of SgRNA004 gene. lanes 1 and 2 showed PCR products (660 bp) of SgRNA004 gene, lane 3 is the negative control.

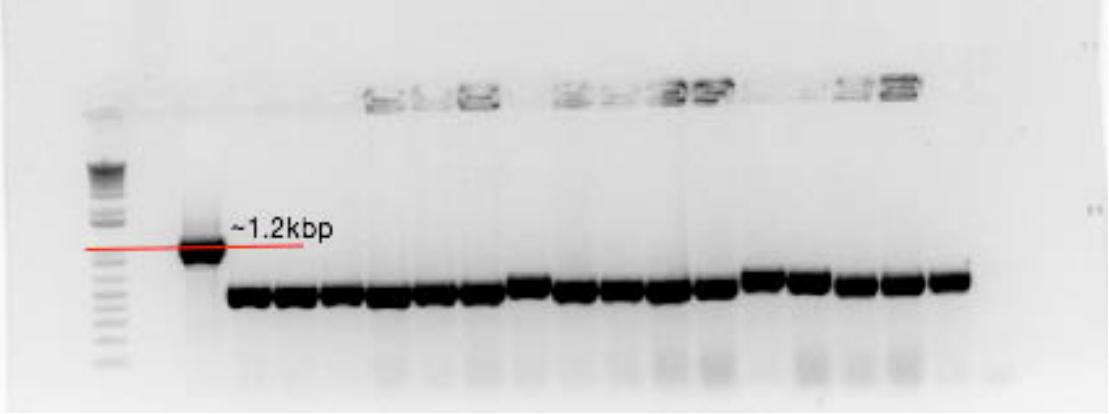

FIGURE 3.4: Electrophoresis results of colony PCR on sgRNA-pSMB47 with the expected band size of 1.2kbp with the insert and 560bp if empty.

We proposed two hypothesis regarding the DNA deletion:

1. **Oligonucleotide synthesis is prone to error.**

   Deletions at a rate of 1 in 100 bases and mismatches and insertions at about 1 in 400 bases can occur as a result of synthesis errors [21].

2. **A spontaneous 1 nucleotide deletion during PCR amplification.**

When one or more nucleotides is removed from the DNA. Deletion can alter the reading frame of the gene.

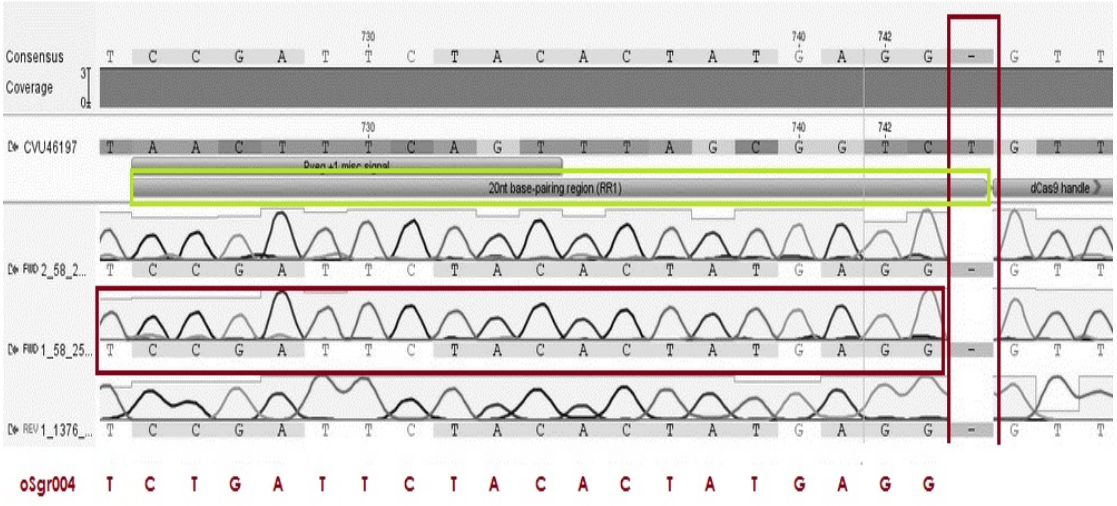

FIGURE 3.5: Schematic of Sanger sequencing result of the only positive sgRNA004-pSMB47 colony. The picture indicates the missing nucleotide in the 20-nt sgRNA binding site

In order to investigate why one nucleotide is missing, we repeated the sgRNA mutagenesis PCR on pJMP2. We sequence two independent clones. Clone#1 has the nucleotide missing and Clone#2 does not (see Appendix B). We can conclude that part of the primers lacks 1 nucleotide. We were either unlucky to continue with the one that had the missing nucleotide or lucky to have one clone without the missing nucleotide. To clarify any ambiguities, more clones should be sequenced.

We repeated the mutagenic PCR with oSgr001/oSgr002 and oSgr001/oSgr007. By using a combination of TOPO cloning and screening tens of colonies, we were able to build pWKS1816 and pWKS1826 plasmids, which are pSMB-Pveg-sgRNA002 (CD0001) and pSMB-Pveg-sgRNA007 (CD1214) respectively. When sgRNA002 and sgRNA004 are expressed individually in the presence of dCas9, we expect to observe reduction in the size of colonies. The size of colonies when targeting CD3657 (DnaC) must be smaller than when targeting CD0001 (DnaA). Once DnaA (the replication initiator) is blocked the replication will not stop immediately but only new rounds on replication are inhibited. By targeting CD3657, the replication will be stopped almost instantaneously. Therefore, the time-frame from targeting a gene to inhibiting the replication process may indicate the size of colonies. CD1214 (Spo0A) is a highly conserved transcriptional regulator that plays a key role in initiating sporulation in *C.difficile* [22]. If CD1214 is nonfunctional, sporulation does not initiate. To observe how the cells behave when SgRNA007 is

expressed in the presence of dCas9, we perhaps need to place the bacteria in a harsh environmental conditions.

3.2 Construction of pSJ129 plasmid carrying dCas9

As explained earlier, in order to transform a new construct into *C.difficile*, we first need to clone the gene of interest in a conjugative plasmid using proper sites.

3.2.1 Adding His-tags to dCas9 by PCR

The dCas9 gene was amplified by gradient PCR (from plasmid pJMP1) using primers oWKS-1541α to add BsaI recognition site with BamHI cutting site to the start codon of the dCas9 coding sequences, and oWKS-1542α to include PstI site. The new dCas9 carrying PstI/BamHI restriction sites was cloned into a plasmid containing the sigD gene under control of a lantibiotic (nisin) inducible promoter as previously described. The expected PCR product was approximately 4.2kbp (Figure 3.6).

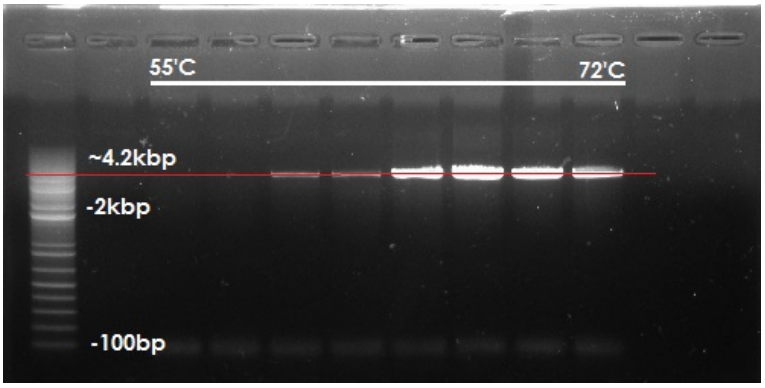

FIGURE 3.6: Electrophoresis results of dCas9 mutagenesis PCR with the expected band size of 4.2kbp

3.2.2 Cloning dCas9 into pSigD using BamHI and PstI sites

Cloning dCas9 into pSigD vector found to be a rather difficult process as we failed to have colonies for few rounds. Since dCas9 is a large fragment (~4.2kbp), we placed the tubes at 4°C to slow down the movement of nucleic acid fragments, and therefore higher the chances of having positive colonies. After incubating the plates for 24 hours, there were few colonies on each plate. We initially performed the colony PCR with oWKS-024 and oWKS-025. We expected the original pSigD with sigD insert to be 1203bp and 4609bp with dCas9 insert. There were multiple non-specific bands on the gel even on

when no template was added to the mix (Figure 3.7). Non-specific bands are usually caused by factors related to cycling times/temperature or PCR components. However, it is difficult to interpret the results due to the fact that non-specific bands also occurred on a template-free mix. We repeated the colony PCR with oWKS-1541α and oWKS-1542α. The PCR amplification showed positive results on clone#5 and #8 (Figure 3.8). We were able to align the entire dCas9 gene in clone#8 with the one in the original plasmid by Sanger sequencing. pSJ129 plasmid was constructed and ready to be used for *C.difficile* conjugation.

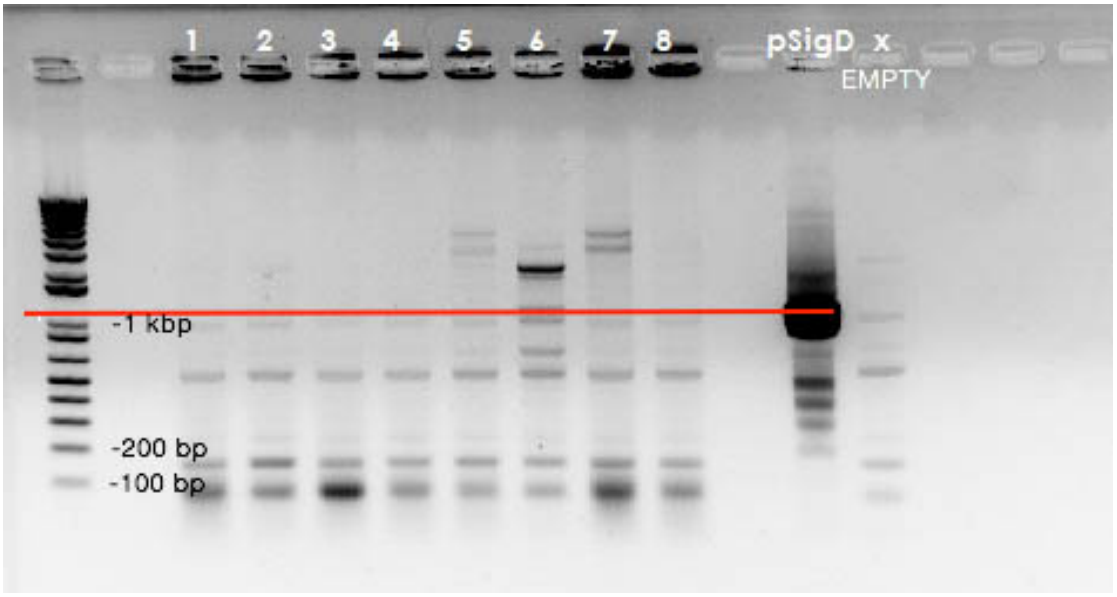

FIGURE 3.7: Electrophoresis results of dCas9 colony PCR. Colony PCR with oWKS-024 and oWKS-025 shows multiple non-specific bands.

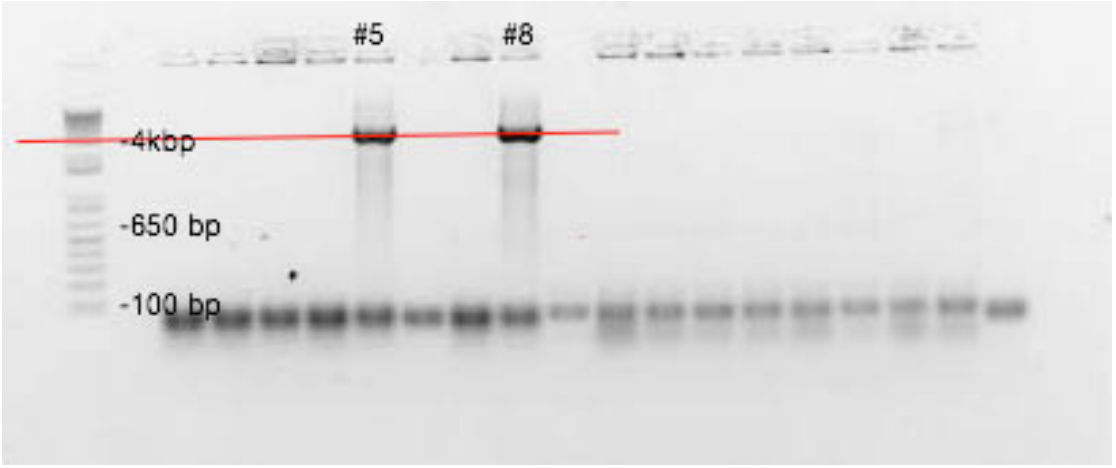

FIGURE 3.8: Electrophoresis results of dCas9 colony PCR. Colony PCR with oWKS-1541α and oWKS-1542α shows positive results on clone#5 and #8.

3.3 Conjugation of pSJ129 to *C.difficile* 630Δerm

Following to the completion of dCas9 cloning into a conjugative plasmid, pSJ129 must be transformed into *E. coli* CA434 cells to generate conjugal donor strains, which can be used to transfer plasmids into *C.difficile* 630Δerm by conjugation.
There were no growth on the patched BHI/CDSS/Thi plates. The conjugation was repeated four times while variables such CA434 transformants, media, plates, and *C. difficile* stock were replaced by fresh ones. However, neither of these changes were effective. We conducted pRPF-based conjugations as a control to figure out whether the problem is caused by the *C. difficile* strain or plasmid. The control conjugations with different plasmids worked efficiently, therefore, we assume that the conjugation of the pSigD plasmid is not as efficient as other conjugative plasmids. As a secondary approach, we designed two primers to modify pRPF185 in such a way that the new engineered plasmid carries SalI/BamHI/NotI polylinker region and the original Ptet. The next step would be designing primes to add SalI-RBS and NotI restriction sites to the sequence of dCas9. Upon successful completion of the conjugation, we will test whether dCas9 can be expressed in *C.difficile* by means of the western blot. Western blotting is an important technique used in cell and molecular biology that allows researchers to identify specific proteins from a complex mixture of proteins extracted from cells [23].

Strains	Description	Source
CA434	*E. coli* HB101 carrying the Inc conjugative plasmid R702	[24]
DH5α	General cloning strain	[25]
pSJ117	*C.difficile* 630Δerm	WKS1241
Plasmids	**Description**	**Source**
pSJ120	pUC-Pveg-sgRNARR1; BamHI fragment of pJMP2 cloned into pUC19	This study
pSJ121	Same as pSJ120 but after mutagenic PCR, carrying oSgr005 instead of RR1	This study
pSJ122	pJMP2-Pveg-sgRNA004	This study
pSJ123	PCR product carrying Pveg-sgRNA004 from pSJ122 cloned into SphI digested pSMB47	This study
pSJ129	PCR product carrying dCas9 from pJMP1 cloned into BamHI/PstI digested pSigD	This study
pSJ132	pSJ129 transformed into CA434	This study
pWKS1819	PCR product carrying Pveg-sgRNA002 cloned into SphI digested pSMB47	This study
pWKS1826	PCR product carrying Pveg-sgRNA007 cloned into SphI digested pSMB47	This study
pJMP1	Pxyl-dcas9; dcas9 cloned into BamHI site	Gross lab, UCSF
pJMP2	Pveg-sgRNARR1; sgRNARR1 cloned into EcoRI site	Gross lab, UCSF
pSigD	Used as a conjucative plasmid: pSigD is a vector that carries the sigD gene under control of a lantibiotic (nisin) inducible promoter	[26]
pSMB47	Used as a conjucative plasmid: integrates in Tn916 in *B.subtilis* strain BS49	[27]
pUC19	Used as a cloning vector	[28]

TABLE 3.1: List of plasmids and strains used in this study

Chapter 4

Conclusion and future perspectives

4.1 Conclusion

The experiments presented in this report demonstrate the steps of designing and creating CRISPR-dCas9 construct to inhibit the bacterial growth by silencing replication genes in *Clostridium difficile*. This construct is composed of two parts: a dCas9 protein and a customized single-guided RNA (sgRNA) with a 20-basepair complementary region to target replication genes in *C.difficile*. Expression of sgRNA in the presence of dCas9 can efficiently silence a target gene and interfere with DNA replication process. Each part was cloned seperately into a significant conjucative plasmid.

Prior to cloning of Pveg-sgRNA into pSMB47 vector, a conjugative plasmid, we had to change the 20nt sgRNA binding site to the sequence of the gene of interest by performing mutagenic PCR. The PCR reaction failed couple of times before we reduced the amount of template added to the PCR mixture by 90%. We were also able to avoid the repetitive PCR failure by extracting Pveg-sgRNA and cloning into a smaller vector. We believe that the large plasmid size was the main contributing factor to the PCR failure. The process of cloning Pveg-sgRNA into pSMB47, also failed regardless of the quality and amount of components being used. We were able to have one positive clone after screening more than 40 colonies. The major problem we encountered was the excess of self-ligated colonies because only one restriction site was involved. When using a single restriction enzyme site for subcloning, we cannot predict or encourage the direction of ligation. It is possible to prevent false colonies by dephosphorylating the plasmid and phosphorylating the insert. Another option is to phosphorylate the primers before carrying out PCR. However, we still had lot of colonies without the insert and

dephosphorylation seemed to be insufficient. On the one hand we can screen as many colonies as possible until we find a positive clone, but on the other hand we can avoid self-ligation and incorrect orientation by adding another restriction site to our design. To do so, we can add a suitable restriction site that is unique in the cloning vector and absent in the Pveg-sgRNA construct (see Appendix C).

Due to internal restriction sites in dCas9, the process of cloning into pSigD (the conjugative plasmid) became a rather challenging task. To avoid cutting the sequence of the gene of interest, we used a special enzyme that has a great feature of cutting outside its recognition site. The cloning process was successful and we were able to confirm the the entire 4.2bp of the gene by Sanger sequencing.

The transformation of dCas9 via conjugation failed four times. We assumed that the current cloning vector is not as efficient as other conjugative vectors such as pRPF185. For that matter, we propose two possible ways to overcome potential conjugation problems: (a) Optimization of conjugation by changing conjugation time, donor strain, and antibiotic selection, and (b) Conduction of a secondary strategy with the use of a more efficient conjugative vector such as pRPF185.

The future research should focus on transforming the plasmids carrying sgRNA(s) and dCas9 into *C.difficile* and analyze what the expected phenotypes are for each of sgRNAs when expressed in the presence of dCas9.

4.2 Future perspective

The CRISPR-dCas9 system offers a powerful approach for silencing genomic sequences, allowing the study of gene function at nucleotide level. This technology is not only limited to bacterial replication genes but it can be used to manipulate the expression of other essential genes. One potential aspect would be developing the CRISPR-dCas9 system to interfere with bacterial cell-to-cell communication (quorum sensing). Quorum sensing is the mechanism by which bacterial populations coordinately control gene expression in response to cell population density [29]. Each bacterium produce and release small, diffusible signal molecules called autoinducers that are sensed by the entire population. This will result in the coordinate production of virulence factors. Currently, there is no direct evidence that *C. difficile* employs quorum sensing in literature. However, the release of two major virulence determinants in *C. difficile*, toxin A and toxin B, is cell population density dependent [30]. Moreover, a homologue of luxS, a gene involved in quorum sensing, has been identified in the genome of *C. difficile* [30]. Therefore, there is a possible coupling between quorum sensing and initiation of DNA

replication or perhaps there is a mutual gene(s) involved in both of the processes. The CRISPR-dCas9 system allows cloning multiple sgRNAs into a single vector for manipulating the expression of multiple genes and thus it is a handy tool for conducting lateral research.

Appendix A

(A) Conjugative plasmids used in this study

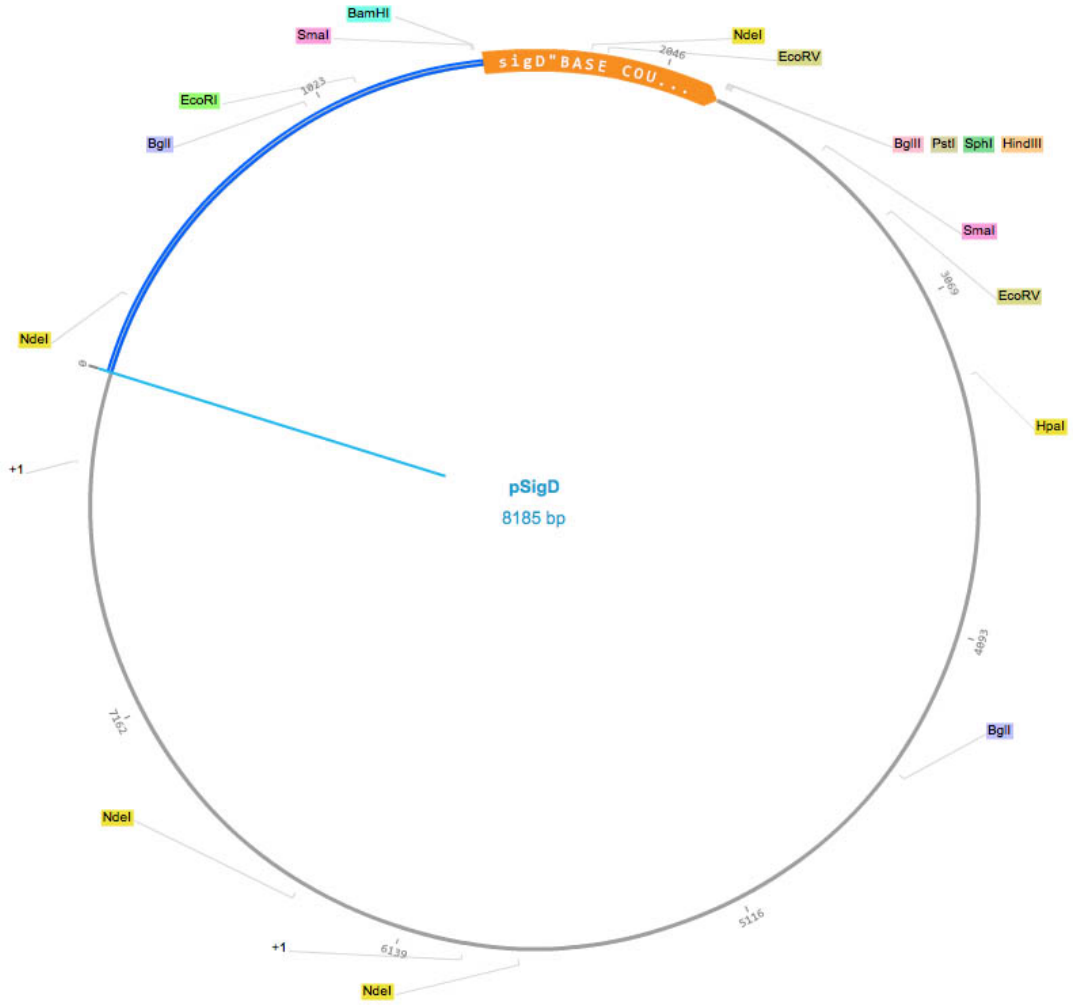

FIGURE A.1: Schematic of pSigD conjugative vector used in this study for dCas9 transformation

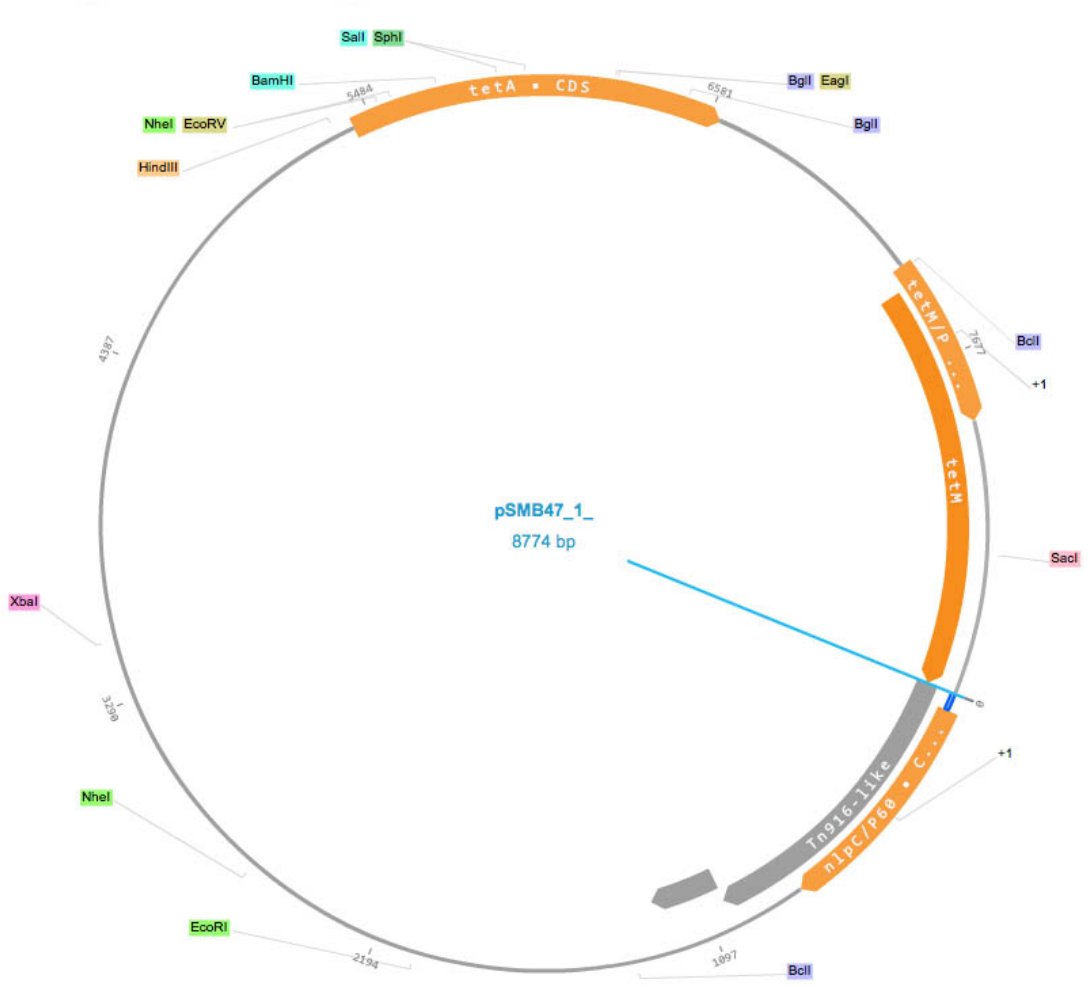

FIGURE A.2: Schematic of pSMB47 conjugative vector used in this study for sgRNA transformation

Appendix B

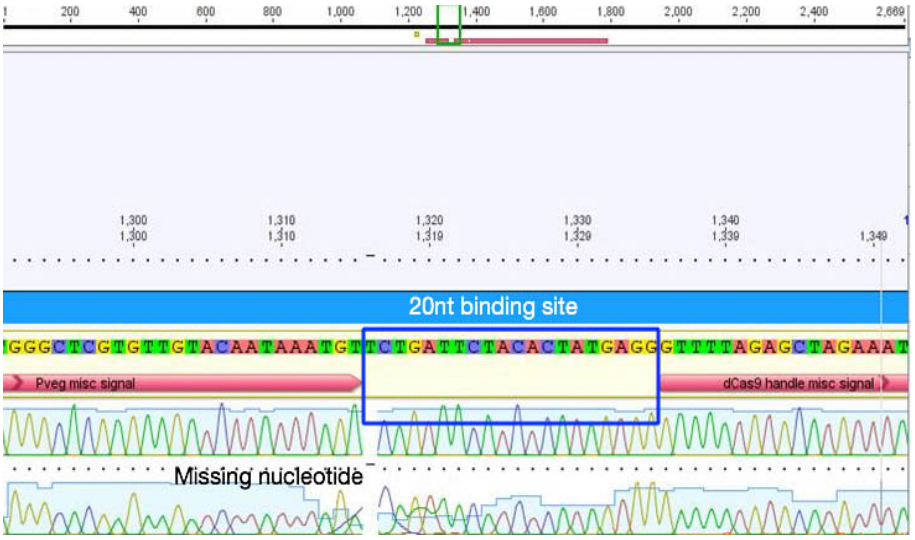

FIGURE B.1: A sanger sequence read of self-ligated sgRNA004 with oWKS-1551 and oWKS-15441. clone#1: is missing nucleotide

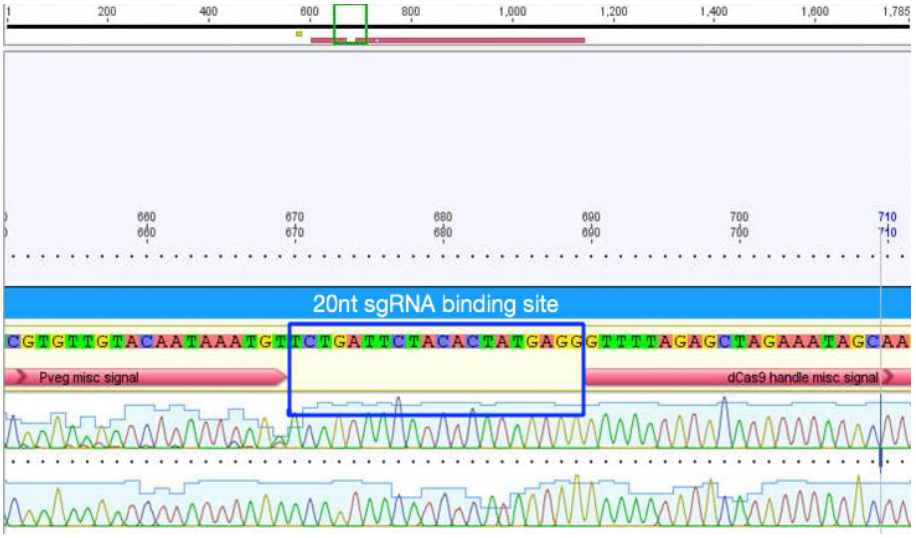

FIGURE B.2: A sanger sequence read of self-ligated sgRNA004 with oWKS-1551 and oWKS-15441. clone#2: has the complete sequence of 20nt-binding site

Appendix C

As it was mentioned in the results section, we digested pJMP2 vector with BamHI, extracted the sequence of the gene of interest, and cloned it into pUC19 to have a shorter mutagenesis PCR. In this case, we can design primers for the second PCR in such away that one BamHI site is included in the PCR product while another BamHI site is excluded. Instead, similar to the initial design, we add a SphI site by means of primer2 as shown in (Figure C.1).

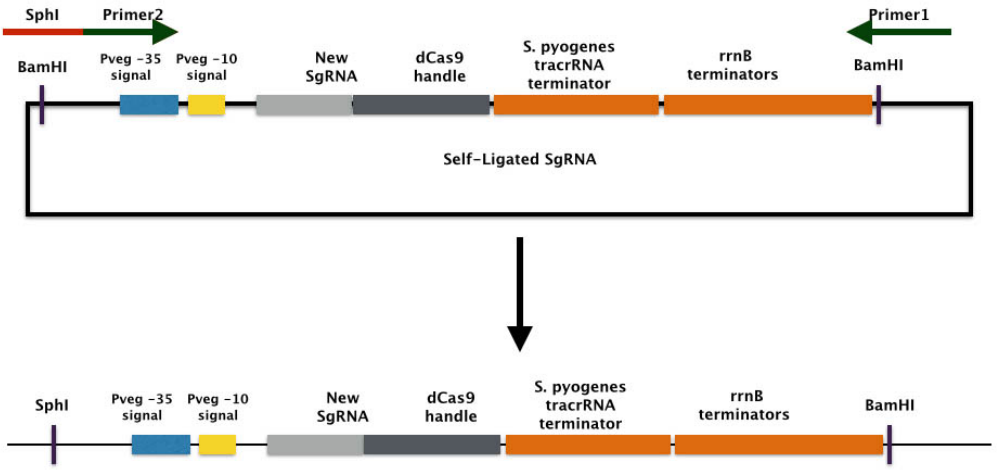

FIGURE C.1: PCR amplification: one BamHI site is included in the PCR product while another BamHI site is excluded. Instead, a SphI site by means of primer2 as shown. The PCR product can be digested by BamHI/SphI and cloned into pSMB47

Bibliography

[1] Fredrik Bäckhed, Ruth E Ley, Justin L Sonnenburg, Daniel A Peterson, and Jeffrey I Gordon. Host-bacterial mutualism in the human intestine. *science*, 307(5717):1915–1920, 2005.

[2] Michael Lynch and Aaron O Richardson. The evolution of spliceosomal introns. *Current opinion in genetics & development*, 12(6):701–710, 2002.

[3] Anthony RM Coates, Gerry Halls, and Yanmin Hu. Novel classes of antibiotics or more of the same? *British journal of pharmacology*, 163(1):184–194, 2011.

[4] Carrie Arnold. To share and share alike: Bacteria swap genes with their neighbors more frequently than researchers have realized. *Scientific American*, 2011.

[5] Stephen T Odonkor and Kennedy K Addo. Bacteria resistance to antibiotics: recent trends and challenges. 2011.

[6] ECCMID Barcelona. P0755:towards a europe-wide c. difficile infection surveillance. 2014.

[7] David A Burns, John T Heap, and Nigel P Minton. Clostridium difficile spore germination: an update. *Research in microbiology*, 161(9):730–734, 2010.

[8] Yoshizumi Ishino, Hideo Shinagawa, Kozo Makino, Mitsuko Amemura, and Atsuo Nakata. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in escherichia coli, and identification of the gene product. *Journal of bacteriology*, 169(12):5429–5433, 1987.

[9] Rodolphe Barrangou, Christophe Fremaux, Hélène Deveau, Melissa Richards, Patrick Boyaval, Sylvain Moineau, Dennis A Romero, and Philippe Horvath. Crispr provides acquired resistance against viruses in prokaryotes. *Science*, 315(5819): 1709–1712, 2007.

[10] Gene W Tyson and Jillian F Banfield. Rapidly evolving crisprs implicated in acquired resistance of microorganisms to viruses. *Environmental microbiology*, 10(1): 200–207, 2008.

[11] Pierre Boudry, Ekaterina Semenova, Marc Monot, Kirill A Datsenko, Anna Lopatina, Ognjen Sekulovic, Maicol Ospina-Bedoya, Louis-Charles Fortier, Konstantin Severinov, Bruno Dupuy, et al. Function of the crispr-cas system of the human pathogen clostridium difficile. *mBio*, 6(5):e01112–15, 2015.

[12] F Ann Ran, Patrick D Hsu, Chie-Yu Lin, Jonathan S Gootenberg, Silvana Konermann, Alexandro E Trevino, David A Scott, Azusa Inoue, Shogo Matoba, Yi Zhang, et al. Double nicking by rna-guided crispr cas9 for enhanced genome editing specificity. *Cell*, 154(6):1380–1389, 2013.

[13] Lei S Qi, Matthew H Larson, Luke A Gilbert, Jennifer A Doudna, Jonathan S Weissman, Adam P Arkin, and Wendell A Lim. Repurposing crispr as an rna-guided platform for sequence-specific control of gene expression. *Cell*, 152(5):1173–1183, 2013.

[14] Wiep Klaas Smits, Houra Merrikh, Carla Yaneth Bonilla, and Alan D Grossman. Primosomal proteins dnad and dnab are recruited to chromosomal regions bound by dnaa in bacillus subtilis. *Journal of bacteriology*, 193(3):640–648, 2011.

[15] Wiep Klaas Smits, Alexi I Goranov, and Alan D Grossman. Ordered association of helicase loader proteins with the bacillus subtilis origin of replication in vivo. *Molecular microbiology*, 75(2):452–461, 2010.

[16] John S Hawkins, Spencer Wong, Jason M Peters, Ricardo Almeida, and Lei S Qi. Targeted transcriptional repression in bacteria using crispr interference (crispri). *CRISPR: Methods and Protocols*, pages 349–362, 2015.

[17] Celeste Yanisch-Perron, Jeffrey Vieira, and Joachim Messing. Improved m13 phage cloning vectors and host strains: nucleotide sequences of the m13mpl8 and puc19 vectors. *Gene*, 33(1):103–119, 1985.

[18] Shonna M McBride and Abraham L Sonenshein. Identification of a genetic locus responsible for antimicrobial peptide resistance in clostridium difficile. *Infection and immunity*, 79(1):167–176, 2011.

[19] Robert W McKee, Mihnea R Mangalea, Erin B Purcell, Erin K Borchardt, and Rita Tamayo. The second messenger cyclic di-gmp regulates clostridium difficile toxin production by controlling expression of sigd. *Journal of bacteriology*, 195(22): 5174–5185, 2013.

[20] Des Purdy, Triona AT O'Keeffe, Michael Elmore, Mike Herbert, Anne McLeod, Monika Bokori-Brown, Anna Ostrowski, and Nigel P Minton. Conjugative transfer of clostridial shuttle vectors from escherichia coli to clostridium difficile through

circumvention of the restriction barrier. *Molecular microbiology*, 46(2):439–452, 2002.

[21] Jingdong Tian, Hui Gong, Nijing Sheng, Xiaochuan Zhou, Erdogan Gulari, Xiaolian Gao, and George Church. Accurate multiplex gene synthesis from programmable dna microchips. *Nature*, 432(7020):1050–1054, 2004.

[22] Laura J Deakin, Simon Clare, Robert P Fagan, Lisa F Dawson, Derek J Pickard, Michael R West, Brendan W Wren, Neil F Fairweather, Gordon Dougan, and Trevor D Lawley. The clostridium difficile spo0a gene is a persistence and transmission factor. *Infection and immunity*, 80(8):2704–2711, 2012.

[23] Tahrin Mahmood, Ping-Chang Yang, et al. Western blot: technique, theory, and trouble shooting. *North American journal of medical sciences*, 4(9):429, 2012.

[24] Des Purdy, Triona AT O'Keeffe, Michael Elmore, Mike Herbert, Anne McLeod, Monika Bokori-Brown, Anna Ostrowski, and Nigel P Minton. Conjugative transfer of clostridial shuttle vectors from escherichia coli to clostridium difficile through circumvention of the restriction barrier. *Molecular microbiology*, 46(2):439–452, 2002.

[25] Seth G Grant, Joel Jessee, Fredric R Bloom, and Douglas Hanahan. Differential plasmid rescue from transgenic mouse dnas into escherichia coli methylation-restriction mutants. *Proceedings of the National Academy of Sciences*, 87(12):4645–4649, 1990.

[26] Robert W McKee, Mihnea R Mangalea, Erin B Purcell, Erin K Borchardt, and Rita Tamayo. The second messenger cyclic di-gmp regulates clostridium difficile toxin production by controlling expression of sigd. *Journal of bacteriology*, 195(22): 5174–5185, 2013.

[27] Shonna M McBride and Abraham L Sonenshein. Identification of a genetic locus responsible for antimicrobial peptide resistance in clostridium difficile. *Infection and immunity*, 79(1):167–176, 2011.

[28] Celeste Yanisch-Perron, Jeffrey Vieira, and Joachim Messing. Improved m13 phage cloning vectors and host strains: nucleotide sequences of the m13mpl8 and puc19 vectors. *Gene*, 33(1):103–119, 1985.

[29] Robert P Fagan and Neil F Fairweather. Clostridium difficile has two parallel and essential sec secretion systems. *Journal of Biological Chemistry*, 286(31):27483–27493, 2011.

[30] Glen P Carter, Des Purdy, Paul Williams, and Nigel P Minton. Quorum sensing in clostridium difficile: analysis of a luxs-type signalling system. *Journal of medical microbiology*, 54(2):119–127, 2005.

Glossary

Lysogeny broth is a nutritionally rich medium for the growth of bacteria. 11

non-template DNA strand DNA molecule has a double-stranded structure. One strand is the coding strand or non-template strand, and the other is the non-coding strand or template strand. 9

PCR The polymerase chain reaction (PCR) is a revolutionary method developed by Kary Mullis in the 1980s. PCR is a technology in molecular biology used to synthesize new strand of DNA and make a single copy or a few copies of a piece of gene of interest. 9

polylinker is a region composed of several restriction enzyme recognition sites engineered into a single cluster. 25

promoter In genetics, a promoter is a part of DNA that initiates transcription of a specific gene. 8

pUC19 is a commonly used cloning vector in *E.coli*. It is 2686 base pairs in length, carries 54 base-pair multiple cloning site, and has a high copy number. 19

Sanger sequencing is a DNA sequencing method in which target DNA is denatured and annealed to an oligonucleotide primer, and extended by DNA polymerase using a mixture of deoxynucleotide triphosphates. 12

TOPO cloning is a molecular biology technique in which PCR products are cloned into specific vectors without the use for DNA ligases. The DNA sequences can be ligated into TOPO with compatible ends. The ligation is complete in only 5 minutes at room temperature. 20