Enhancing the Ministring DNA (msDNA) Purification Using PI-Sce1 Homing Endonuclease/CRISPR-Cas3 Recombinant System

by

Merium Fernando

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Science

in

Pharmacy

Waterloo Ontario, Canada, 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

In the generation of msDNA the recombinant *E. coli* cells are transformed by a msDNA generating precursor plasmid, whereupon expression of the Tel protelomerase enzyme, acting on the *pal* target sequence present in the precursor plasmid, generated linear covalently closed (LCC) msDNA. However, the *in vivo* recombinant platform to produce msDNA results in a mixture of plasmids including unprocessed precursor plasmid, unwanted LCC bacterial backbone, and their topological isoforms, which interferes with the purification of the target species. For larger scale synthesis, the plasmid extract needs to be pretreated with commercially available restriction enzymes before being purified through chromatographic columns. Meanwhile, at the laboratory scale, msDNA is purified from agarose gels based on their size. These purification processes are time-consuming and inefficient and therefore, there is a need to optimize the process.

To address this issue, we developed two *in vivo* recombinant systems for digesting the unwanted prokaryotic backbone and unprocessed precursor plasmid. These systems are the PI-SceI homing endonuclease enzyme system and the clustered regularly interspaced short palindromic repeats-Cas3 (CRISPR-Cas3) system. Homing endonucleases are highly specific DNA cleaving enzymes. The homing endonuclease PI-SceI, encoding gene *vma* from *Saccharomyces cerevisiae* was successfully integrated into the *tel* integrated bacterial chromosome via site-specific recombination using conditional replication and integration (CRIM) plasmid. The double integrants, both *vma* and *tel* integrated recombinant bacteria, were transformed with msDNA synthesizing precursor plasmids and induced the msDNA synthesis and *vma* gene overexpression. Even though the double integrants were able to overexpress the homing endonuclease enzyme and digest the precursor

plasmid, they were not able to synthesize msDNA. Therefore, the Tel protelomerase enzyme was expressed episomally inside the *vma* integrated recombinant bacteria. This *vma* gene is under the control of an inducible P_{BAD} promoter. In the presence of L-arabinose in the media, the Tel protelomerase enzyme was episomally expressed and synthesized msDNA by acting on the precursor plasmid. Subsequently, the overexpressed PI-SceI homing endonuclease enzyme digested the undesired byproducts of msDNA synthesis as expected. Introducing homing endonuclease enzyme recognition sequences into the Tel protelomerase enzyme-expressing plasmid will further improve the purification process. The other recombinant system that was developed is the utilization of the CRISPR-Cas3 system which is naturally present in W3110 *E. coli* K-12 bacteria. A pre-crRNA targeting the origin of replication (*ori*) of the msDNA synthesizing precursor plasmid was successfully designed and cloned into the low copy number plasmid. The pre-crRNA expressing gene cassette was placed under the control of the P_{BAD} promoter. Upon overexpression, crRNA was synthesized inside the W3110 *E. coli* K-12 bacteria. The crRNA bound to the expressed CRISPR-Cas3 protein cascade of the bacteria, guided the effector complex to the target sequence and successfully digested the targeted precursor plasmid. Even though the W3110 tel^+ recombinant bacteria synthesized msDNA in a pre-crRNA expressing background, an efficient degradation of the unwanted by-products of msDNA synthesis was not observed. This could be due to the disruption of the CRISPR locus of W3110 tel^+ recombinant bacteria. Episomal expression of the CRISPR-Cas genes inside W3110 tel^+ recombinant bacteria will enhance the digestion of the non-msDNA species.

Acknowledgements

First of all, I want to express my sincere and utmost gratitude to my supervisor, Dr. Roderick Slavcev, for giving me the opportunity to pursue a master's degree under his guidance. It was truly an honor to work under his supervision, and I am grateful for all the support he has provided throughout the years. I extend my thanks to the committee members, Dr. Emmanuel Ho, Dr. Bernard Duncker, and Dr. Bernard Glick, for serving as committee members and for their long-standing support and guidance.

My sincere gratitude goes to Dr. Shawn Wetting and Mrs. April Wetting for introducing me to the School of Pharmacy for my graduate studies. I am honored to have met such humble individuals, and I appreciate their unconditional support at the beginning of my graduate journey. I take this opportunity to express my sincere gratitude to Dr. Diana Duca for encouraging me to pursue a master's degree. My enthusiasm for molecular biology was ignited while working with her in Glick's lab during her Ph.D. Thank you for encouraging me and providing guidance and support during the initial stages of my graduate studies. I would like to express my sincere gratitude to Dr. Shirley Wong from the Slavcev Lab for the mentorship and knowledgeable insights she provided at the beginning of my graduate studies. I would like to thank Ahsan Ayub for valuable insights into my thesis proposal editing, which I found very helpful.

Special appreciation to Deborah Pushparaja for her exceptional assistance in lab work and for providing valuable insights while editing the thesis. Since I pursued my master's part-time, I needed help with lab work on some of the days, and she never said no when I required assistance. Her humble and honest support in making my part-time graduate studies a success is tremendous. I thank Talia Fiaani for her assistance with my CRISPR-Cas3 project, which was her honors thesis project, and for her support in completing it. I express gratitude to Dr. Steven Hersch from Mediphage for his insights and knowledge contributed to the CRISPR-Cas3 project. My sincere thanks also go to Dr. Nafiseh Nafissi from Mediphage for her knowledgeable insights at the beginning of my graduate studies. I extend my thanks to the Slavcev lab group (Rohini Prakash, Julia Lumini, Heba Alattas, Hayden Huh, Anetta Ille, Jessica Nicastro, Jaya Li, Nicholas Cheng, and Jesse St Jean) for all their support. I would like to thank Melinda Recchia for the support she provided me throughout the years.

My deepest gratitude goes out to my family. I want to express my sincere gratitude to my husband, Menaka Silva, my daughter, Manisha Silva, my father, Flavian Fernando, and my mother, Quinta De Silva, for their continuous support over a long period and for enduring my monotonous student life.

Dedication

To my beloved daughter, Manisha Silva

Table of Contents

List of Tables ..... **xi**
List of Figures ..... **xii**
List of Abbreviations ..... **xiv**
1.Introduction **1**
1.1 Gene therapy **1**
1.2 Viral and non-viral GT vectors **6**
1.3 Non-viral gene delivery vectors used in GT **8**
1.4 Synthesis of msDNA **13**
1.5 Purification of msDNA **22**
1.6 Homing endonucleases **24**
1.7 Conditional-replication, integration and modular (CRIM) plasmids **26**
1.8 CRISPR-Cas3 system of *E. coli* K-12 bacteria **30**
1.8.1 Type 1E CRISPR-Cas system **32**
1.8.2 Strategies employed by phages to evade CRISPR immunity **34**
2. Rationale, hypothesis, and objectives **38**
2.1 Rationale **38**
2.2 Objectives: **39**
3. Materials and methods **41**
3.1 Strains and plasmids **41**
3.1.1 Media and culture conditions **42**
3.2 Plasmid construction **42**
3.2.1 Construction of pSW3 plasmid **42**
3.2.2 Construction of the pACYC184-CI-857-tel plasmid construct **44**
3.2.3 Construction of pDMS-tel recombinant plasmid **46**
3.2.4 Construction of the CRISPR Cas3 pre-crRNA expressing plasmids **47**
3.2.4.1 Phosphorylation and cleanup of PCR amplified product **50**
3.2.4.2 Ligation and transformation of the phosphorylated PCR product **50**
3.2.4.3 Confirmation of the insertion of the pre-crRNA encoding gene sequence **50**
3.2.5 Construction of pcrRNA120-Km^R plasmid (p120K) **51**
3.3 Integration of the pSW3 plasmid into the bacterial chromosome **53**
3.4 Inductions **56**

3.4.1 Regulated expression of the *vma* gene in recombinant bacteria **56**

3.4.2 Induction of ministring DNA (msDNA) synthesis **57**

3.4.3 Induction of ministring DNA generation and endonuclease expression in *tel*$^+$ *vma*$^+$ double integrants **57**

3.4.4 Induction of msDNA synthesis in JM109 *tel*$^+$*vma*$^+$ double integrants **59**

3.4.5 Episomal expression of the Tel protelomerase enzyme **59**

3.4.5.1Induction of msDNA synthesis in pACYC-CI-857-tel (pACYC-tel) carrying W3110/JM109 recombinant bacteria in the presence of pDMS/pNN9 precursor plasmids **60**

3.4.5.2 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-T7 and pDMS/pNN9 precursor plasmids **60**

3.4.5.3 Induction of msDNA synthesis in W3110 bacteria carrying p-pBAD-tel and pDMS/pNN9 plasmids **61**

3.4.5.4 Induction of msDNA synthesis in W3110 *vma*$^+$ bacteria carrying pXG-pBAD-tel and pDMS/pNN9 plasmids **62**

3.4.6 Induction of crRNA over expression **63**

3.4.6.1 Induction of msDNA synthesis in W3110 *tel*$^+$ recombinant bacteria carrying the CRISPR plasmid **63**

3.4.6.2 Induction of msDNA synthesis followed by crRNA overexpression **63**

3.5 Transformation 64

3.6 *In vitro* digestion of the pDMS plasmid 64

3.7 Integration of the *vma* gene into JM109 *tel*$^+$ recombinant bacteria using P1 transduction 65

3.7.1 P1 transduction **65**

3.7.2 Regeneration of the P1rev6 ..... **66**

3.7.3 Spot plate technique to determine the phage titer **67**

3.7.4 Preparation of P1 *vma*$^+$ transducing lysate **68**

3.7.5 Generation of JM109 *tel*$^+$*vma*$^+$ double integrants **69**

3.8 Designing of the CRISPR Cas 3 pre-crRNA encoding gene sequence 70

4. Results 72

4.1 Integration of the PI-SceI homing endonuclease encoding gene (*vma*) into the bacterial chromosome 72

4.2 Expression of the *vma* homing endonuclease gene upon integration into the *E. coli* genome 74

4.2.1 Transformation of the pDMS precursor plasmid into the positive integrant **74**

4.2.2 Induction of the *vma* overexpression in recombinant bacteria **75**

4.2.3 Induction of the *vma* gene over expression in both *tel* and *vma* integrated recombinant bacteria..... **77**

4.2.4 Ministring DNA synthesis in *vma* and *tel* integrated double integrants..... **79**

4.2.5 Ministring DNA synthesis in double integrants with pNN9 precursor plasmid..... **81**

4.3 Generation of JM109 *tel*$^+$*vma*$^+$ double integrants via P1 transduction..... 85

4.3.1 Induction of msDNA synthesis and homing endonuclease enzyme overexpression in JM109 *tel*$^+$*vma*$^+$ double integrants **86**

4.4 Episomal expression of the Tel protelomerase enzyme 90

4.4.1 Induction of msDNA synthesis in W3110/JM109 recombinant bacteria carrying the pACYC-CI-857-tel plasmid in the presence of pDMS/pNN9 precursor plasmids **92**

4.4.2 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-pBAD and pDMS/pNN9 precursor plasmids **97**

4.4.3 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-T7 and pDMS/pNN9 recombinant plasmids **98**

4.4.4 Induction of msDNA synthesis in W3110+pXG-pBAD-tel+pNN9/pDMS carrying recombinant bacteria..... **103**

4.5 Digestion of the msDNA synthesizing precursor plasmid upon over expression of the pre-crRNA targeting the bacterial backbone 110

4.5.1 pre-crRNA over expression in W3110 recombinant bacteria **110**

4.6 Determination of msDNA synthesis in pre-crRNA synthesizing recombinant bacteria 114

4.6.1 Expression of msDNA Synthesis in pre-crRNA expressing W3110 *tel*$^+$ recombinant bacteria..... **114**

4.6.2 Induction of msDNA synthesis..... **115**

4.6.3 Induction of msDNA synthesis followed by CRISPR RNA over expression..... **116**

5. Discussion..... 120

6. Conclusion and future directions 129

7. References..... 133

Appendices..... 143

Appendix A 143

List of Tables

Table 1: Core *attB* sequences (highlighted) of various lambdoid phages. **29**
Table 2: Bacteria and plasmids used in the experiment. **42**
Table 3: Cycling instructions for amplification of the 3.4 kb gene fragment. **45**
Table 4: Forward and reverse primers used to generate the pre-crRNA encoding gene sequence. **48**
Table 5:PCR steps for insertion of the pre-crRNA encoding gene sequences into the pACYC-araC-tel-Duet-1 plasmid..... **49**
Table 6: Primers used in Gibson assembly to replace the Cm^R gene. **52**
Table 7: Gibson assembly PCR steps for replacing the Cm^R gene with the Km^R gene. **53**
Table 8: Primers used in PCR test to confirm the chromosomal integration of the CRIM plasmid.. ..... **54**
Table 9: Summary of the expected test results after msDNA synthesis and PI-SceI enzyme over expression in different bacterial strains.. ..... **58**
Table 10: Sizes of msDNA and bacterial backbone fragments in the presence of pNN9 and pDMS. **88**
Table 11: Plasmid and msDNA fragment sizes after digestion with EcoR1**.** **88**
Table 12: Expected test results upon heat induction in W3110 bacteria carrying p-Tel plasmid and precursor plasmids. **93**
Table 13: Expected test results upon heat induction in JM109 bacteria carrying pDMS and p-Tel plasmids. **94**
Table 14: Summary of the expected fragment sizes versus the observed fragment sizes upon heat-dependent induction of msDNA synthesis in W3110 recombinant bacteria. **95**
Table 15: Summary of the expected fragment sizes versus the observed fragment sizes upon heat-dependent induction of msDNA synthesis in JM109 recombinant bacteria. **97**
Table 16: Summary of the expected fragment sizes versus the observed fragment sizes upon IPTG-dependent induction of msDNA synthesis in BL21(DE3) recombinant bacteria. **101**
Table 17: Summary of the expected fragment sizes versus the observed fragment sizes upon IPTG-dependent induction of msDNA synthesis in W3110 recombinant bacteria. **103**
Table 18:DNA fragment sizes obtained after digesting the plasmid extract with EcoR1. ..... **109**
Table 19: Summary of the expected fragment sizes versus the observed fragment sizes upon L-arabinose-dependent induction of msDNA synthesis in W3110 *vma*$^+$ recombinant bacteria. ... **109**
Table 20: The expected fragment sizes versus observed fragment sizes upon CRISPR RNA overexpression**.**. **114**
Table 21: The expected fragment sizes versus observed fragment sizes upon msDNA synthesis followed by CRISPR RNA overexpression..... **119**

List of Figures

Figure 1: Summary of the number of clinical studies that have been conducted, along with their current phase of study from year 2010 to 2020**.** **1**
Figure 2: Different approaches of using GT in treating genetic diseases. **3**
Figure 3: Vectors used in GT clinical trials from 2010 to 2020. ..... **8**
Figure 4: Vector integration into bacterial chromosome. **11**
Figure 5: Development of MIDGE and MiLV vectors. **12**
Figure 6: Temperature regulated expression of the Tel and TelN enzyme encoding genes in recombinant *E. coli* (R-cells) bacterial cells. **15**
Figure 7:msDNA pNN9 (5.6 kb) precursor parental plasmid. **18**
Figure 8: Synthesis of LCC msDNA vectors using the pNN9 precursor plasmid. **19**
Figure 9: Schematic representation of the msDNA synthesis using the pDMS precursor plasmid. **21**
Figure 10: Isolated plasmid DNA extraction after msDNA generation using a pNN9 precursor plasmid on a 0.8 % agarose gel. **24**
Figure 11:PI-SceI homing endonuclease 30 bp recognition sequence and the cleavage site. **26**
Figure 12: Integration of the bacterial plasmid via site-specific recombination into the bacterial chromosome. **27**
Figure 13: The highlighted 15 bp λ *attB* sequence which is identical to the *attP* site. **28**
Figure 14: The chromosomal location of bacterial *attB* present in wild type *E. coli* K-12 bacteria. **29**
Figure 15: Stages of CRISPR immune responses. **31**
Figure 16: Bacterial CRISPR-cas system. **34**
Figure 17: Transcription and maturation of pre-crRNA **36**
Figure 18:CRISPR interference. **37**
Figure 19:pSW3 plasmid map used in the study. **44**
Figure 20:pACYC-CI-857-tel recombinant plasmid. **46**
Figure 21: Generation of pre-crRNA expressing plasmid. **49**
Figure 22: Schematic representation of the integration of pAH120-vma (pSW3) into the bacterial chromosome via site specific recombination. **55**
Figure 23: Plasmid map of pXG-pBAD-tel plasmid. **62**
Figure 24: Preparation of phage dilution series. **67**
Figure 25: Spot Plate Technique to determine the phage titer. **68**
Figure 26: Designing of pre-crRNA encoding gene sequences. **71**
Figure 27: Colony PCR test results on a 1 % agarose gel confirmed the integration of pSW3 and the presence of the integrated *vma* gene in the bacterial chromosome. **73**
Figure 28:pDMS precursor plasmid digested with the PI-SceI endonuclease enzyme. **75**
Figure 29: Extractions of plasmids after induction of the PI-SceI enzyme in *vma*-integrated W3110 and BW25113 recombinant bacteria. **77**
Figure 30: The induction of PI-SceI enzyme expression in double integrants carrying the pDMS plasmid **78**
Figure 31: Heat Induction of msDNA synthesis in pDMS carrying W3110 *tel* $^{+}$*vma*$^{+}$, BW25113 *tel* $^{+}$*vma*$^{+}$ and W3110 *tel*$^{+}$ recombinant bacteria. **80**

Figure 32: Heat induction of msDNA synthesis in BW25113 *tel*$^+$ recombinant bacteria carrying pDMS precursor plasmid. **81**
Figure 33:msDNA synthesis upon heat induction in double integrants carrying pNN9 plasmid.**82**
Figure 34: Chromosomal location of the *tel* and *vma* genes in double integrants. **84**
Figure 35: Transducing JM109 *tel*$^+$ with *vma* lysate. **85**
Figure 36: Colony PCR for positive transductant colonies. **86**
Figure 37:msDNA Synthesis in JM109 *tel*$^+$*vma*$^+$ double integrants with pNN9 and pDMS msDNA synthesizing precursor plasmids. **89**
Figure 38: Induction of msDNA synthesis and homing endonuclease enzyme over expression.**90**
Figure 39: Induction of msDNA synthesis in recombinant wild type bacteria carrying p-Tel plasmids and pNN9/pDMS precursor plasmids. **91**
Figure 40: Induction of msDNA synthesis in *vma* integrated recombinant bacteria. **92**
Figure 41: Induction of msDNA synthesis in W3110 bacteria carrying pACYC-CI-857-tel and pNN9 plasmids..... **94**
Figure 42: Induction of msDNA Synthesis in JM109+pACYC-CI-857-tel+pDMS recombinant bacteria. **96**
Figure 43: Induction of msDNA synthesis in BL21(DE3) recombinant bacteria **100**
Figure 44: Induction of msDNA synthesis in W3110 recombinant bacteria..... **102**
Figure 45: Induction of msDNA synthesis in W3110+pXG-tel+pNN9/pDMS recombinant bacteria. **105**
Figure 46: Colony PCR to confirm chromosomal integration of *vma* gene in W3110 *vma*$^+$ single integrants. **107**
Figure 47: Induction of msDNA synthesis in W3110 *vma*$^+$+pXG-tel+pNN9/pDMS recombinant bacteria. **108**
Figure 48: Induction of pre-crRNA over expression. **112**
Figure 49: Digestion of the plasmid extract with EcoRV..... **113**
Figure 50: Induction of msDNA synthesis in recombinant bacteria carrying pNN9 and pcrRNA120 plasmids. **116**
Figure 51: Induction of p120k over expression followed by msDNA synthesis. **118**
Figure A1: Confirmation of the cloning of pre-crRNA (targeting the ampicillin resistant gene of the msDNA synthesizing precursor plasmid) encoding crRNA90 gene sequence……..**143**

List of Abbreviations

AAV	adeno-associated virus
ADA-SCID	adenosine deaminase severe combined immunodeficiency
ALL	acute lymphoblastic leukemia
Ap	ampicillin antibiotic
ATP	adenosine triphosphate
bp	base pairs
CART-T cell	chimeric antigen receptor T cell
CCC	circular covalently closed
CCR5	C-C chemokine receptor type 5 CCR5
CGSC	coli genetic stock center
Cm	chloramphenicol antibiotic
CpG	cytosine-guanine dinucleotide
CRIM	conditional replication and integration plasmid
CRISPR	clustered regularly interspaced short palindromic repeats
DNA	deoxyribonucleic acid
DSB	double-stranded break
dsRNA	double-stranded RNA
E. coli	*Escherichia coli*
EMA	european medicines agency
FDA	US Food and Drug Administration
GT	gene therapy

HDR	homology-directed repair
HF	high fidelity
HIV	human immunodeficiency virus
H-NS	histone-like nucleotide structuring repressor protein
HSC	hematopoietic stem cells
IPTG	isopropyl ß-D-1-thiogalactopyranoside
Km	kanamycin antibiotic
LB	luria-bertani
LCC	linear covalently closed
LPLD	Lipoprotein lipase deficiency
MIDGE	minimalistic immunogenically defined gene expression vectors
miRNA	micro RNA
mRNA	messenger RNA
NBRP	national bio resource project
NHEJ	non-homologous end-joining
ori	origin of replication
OTC	ornithine transcarbamylase
PCR	polymerase chain reaction
PFU	Plaque forming units
phage	bacterio phage
PNK	polynucleotide kinase
RNA	ribonucleic acid
RPM	revolutions per minute
rSAP	shrimp alkaline phosphatase

siRNA	small interfering RNA
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
SS	super sequence
T4 PNK	T4 polynucleotide kinase
TALENs	transcription activator-like effector nucleases

1.Introduction

1.1 Gene therapy

Gene therapy (GT), first introduced in 1970 after the discovery of recombinant DNA technology, is best defined as the use of nucleic acids to treat or prevent diseases by either repairing, replacing, or regulating the genes relevant to the disease (Friedmann, 1992;Berling et al., 2023). Even though the concept of GT had existed for several years, the clinical investigations only started in 1990, with the first clinical study for the rare immunodeficiency disease, adenosine deaminase severe combined immunodeficiency (ADA-SCID) taking place at the US National Institute of Health. Since then, numerous clinical studies have been conducted for different diseases, and they can be categorized into three main groups: monogenic and polygenic diseases (genetic diseases), infectious diseases, and cancer. The GT clinical trials with their current phases of study from 2010 to 2020 are as indicated in Figure 1 (Arabi et al., 2022).

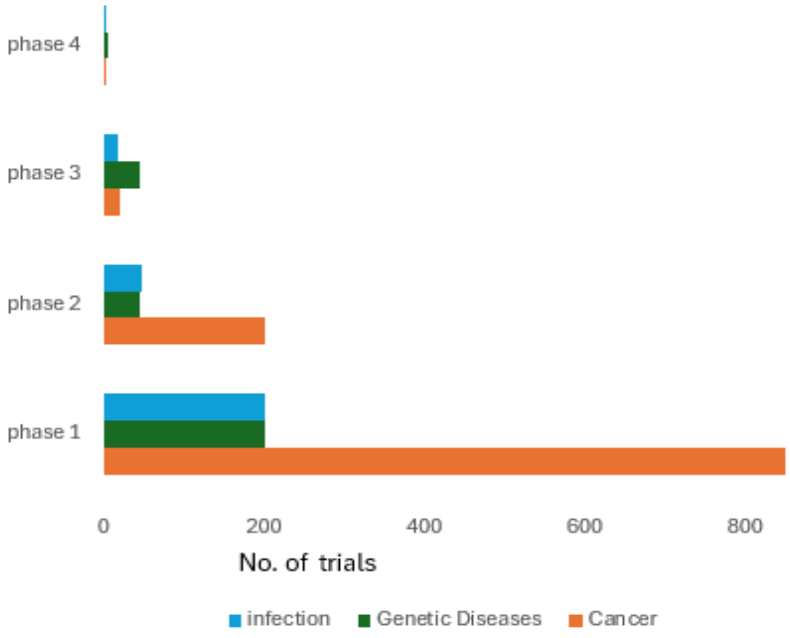


Figure 1:Summary of the number of clinical studies that have been conducted, along with their current phase of study from year 2010 to 2020. Even though most of the clinical trials for

cancer were in phases 1 and 2, they didn't smoothly transition into phases 3 and 4. The majority of phase 3 gene therapy clinical trials are focused on genetic diseases. There are more gene therapy products available for genetic diseases than for cancer. Adapted from Arabi et al., 2022.

GT methods are categorized based on the following criteria: the class of disease (either genetic or complex acquired disease), characteristic of the gene delivery vehicle (either integrating or non-integrating); and based on the way the vector is administered: *in vivo* delivery (directly into the patient cells) or *ex vivo* delivery (transgene is delivered into the cultured cells taken from patients and transplanted back after modifications are done)(Anguela & High, 2018). There are multiple approaches to using GT for therapeutic purposes: to replace a mutated gene with a functional copy to restore the lost cellular function (augmentation GT), knock-down of a mutated gene that encodes defective proteins, which is toxic to cells and introducing new genes to combat the disease. In GT nucleic acids are used to elicit a specific effect in a cell and the therapeutic effect is obtained by the expression of the introduced gene (Komor et al., 2016). The nucleic acids introduced could either be a DNA, messenger RNA (mRNA), small interfering RNA (siRNA), micro-RNA (miRNA) or antisense oligonucleotides (Yin et al., 2014). The goal of augmentation GT of genetic diseases is to confer a controllable expression of the transgene (healthy copy of the mutated gene) at suitable levels and durability to confer therapeutic effect (Figure 2 a) (Anguela & High, 2018). For example, the application of the *ex vivo* GT approach to treating X-linked severe combined immunodeficiency disease replaces the mutated, common gamma chain encoding gene function using retroviral vectors (Salima et al., 2003). Another application of GT is to suppress the expression or knock down the detrimental gene using an RNA interference strategy or gene silencing by degrading the messenger RNA (mRNA) by double-stranded RNA (dsRNA) or genome editing techniques (Figure 2 b) (Anguela & High, 2018; Komor et al., 2016). Genome editing techniques use sequence-specific DNA cleaving nucleases such as zinc-finger nucleases,

transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR-Cas9) (Benjamin et al., 2016;Buchholz, 2009). After the DNA cleaving nucleases introduce the double-stranded DNA breaks at the target site, the DNA breaks can be repaired by two major DNA repair pathways – homology- directed repair (HDR) pathway (via homologous recombination) and non-homologous end-joining (NHEJ) (Lieber, 2010). As shown in Figure 2 (c) in the presence of an exogenously supplied healthy gene copy of the mutated gene, the mutated gene can be corrected via the HDR pathway. The mutated gene can also be knocked-down via a non-homologous end-joining DNA repair pathway if the repair is imperfect. An example of gene editing using CRISPRs, and zinc-finger nucleases is the blocking of the entry of the human immunodeficiency virus (HIV) into cells by disrupting the host endogenous C-C chemokine receptor type 5 (CCR5) (Anguela & High, 2018;Benjamin et al., 2016;Du et al., 2021).

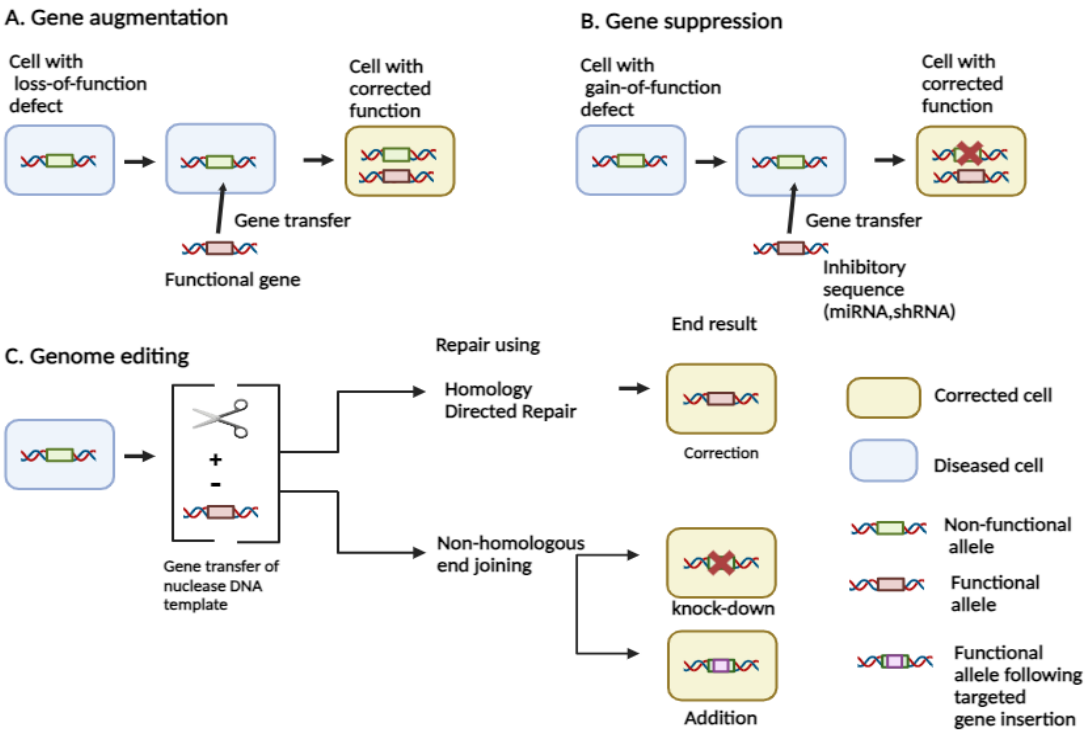


Figure 2:Different approaches of using GT in treating genetic diseases. A. In gene augmentation, the loss of cell function due to the mutated gene will be restored by introducing a healthy functional copy of the defective gene in trans, while the mutated gene is still in the cell. **B.**

Gene suppression is required when a detrimental gene expression results in an accumulation of a toxic product. The detrimental gene expression is suppressed by employing RNA interferences. **C.** The defective gene can be repaired using genome editing techniques. After introducing a double-stranded cut in the defective gene and in the presence of the functional copy of the mutated gene template, via homologous recombination corrects the defective gene. The defective gene is being knock-down by introducing a double-stranded cut and performing non-homologous end-joining results in imperfect repair causes the knock-down of the defective gene. The third approach is the insertion of the corrective gene via non-homologous end-joining. In this approach rather than correcting the defective gene, there is an addition on to the defective gene. Adapted from Anguela & High, 2018.

Yet another approach of GT is the introduction of toxic genes to tumor cells. GT is used to modify cancer cells by introducing tumor-suppressor antigen encoding genes, and suicide genes, where upon expression, causes apoptosis or initiates autonomous cell death in the presence of prodrugs. For example, injection of suicide-gene, thymidine kinase gene (*TK*), expressing adenoviral vectors to tumor cells increases the susceptibility of the tumor cells to prodrugs, which initiates autonomous cell death (Oldfield et al., 1993). Another example is the injection of adenoviral vectors expressing tumor antigen p53 (*p53* tumor suppressor gene) into tumor cells, inducing apoptosis and tumor regression. Another example is the injection of oncolytic viruses that may only replicate and initiate tumor-specific cell lysis (Ginn et al., 2018;Ginn et al., 2013;Parato et al., 2005).

The newly developed “smart drugs” and recombinant vaccines represent further applications of GT. Smart drugs are composed of genetic material and enzyme activities and in the presence of external stimuli change conformation resulting in excising microRNA, which initiates RNA interferences (Bicho D et al., 2015;Lukashev & Zamyatnin, 2016).

Ex vivo GT often uses predominantly integrating vectors to introduce the genetic material into the targeted cells. The vectors are integrated into the genome of stem cells and when the stem cells (undifferentiated cells, replicates, and undergo cell division either to self-renew or to differentiate

into specific cells) replicate, the introduced gene passes onto the daughter cells. In *ex vivo* GT due to the use of integrating vectors, the generation of insertional mutagenesis is a risk. The stability and long-term expression of the episomal, transferred gene is the goal of *in vivo* GT. Additionally, for *in vivo* GT the deleterious immune responses due to the administration of vectors into cells is always a risk (Hardee et al., 2017;Manno et al., 2006;Raper et al., 2003).

The first successful stories of gene therapy applications involved *ex vivo* approaches to treat primary immunodeficiency diseases using autologous hematopoietic stem cells (HSC; self-renewing cells that give rise to all types of blood cells). Since then, several GT treatments have received regulatory approval to date (Anguela & High, 2018;Arabi et al., 2022). In 2012, Glybera, an adeno-associated virus (AAV) based GT drug got conditional marketing approval from the European Medicines Agency (EMA). This drug is used to treat the autosomal recessive genetic disease, familial lipoprotein lipase deficiency (LPLD) (Bryant et al., 2013). The US Food and Drug Administration (FDA) approved another GT-based drug, Imlygic, in 2015. Imlygic utilizes a genetically modified herpes simplex virus type 1 oncolytic virus to treat unresectable lesions in patients with melanoma. Strimvelis, a retrovirus-based *ex vivo* GT drug was approved by EMA in 2016. This is used to treat ADA-SCID (Rehman et al., 2016). Kymriah and Yescartar are chimeric antigen receptor (CAR-T cell) based immunotherapies targeting CD19 antigens present on the surface of B-cell malignancies. These two GT drugs got the FDA approval in 2017 and they are used to treat Non-Hodgkin lymphomas. Kymriah is also used to treat acute lymphoblastic leukemia (ALL) (Brown et al., 2016). Luxturna got FDA approval in 2017 and is being used to treat a rare congenital genetic disease which causes blindness. AAV vectors are used in this GT treatment to replace the mutated biallelic retinal pigment epithelium-specific 65 kilodalton (RPE65) gene function, which results in retinal dystrophy (Anguela & High, 2018;Maguire et al., 2019).

1.2 Viral and non-viral GT vectors

Vectors are used to deliver transgenes or therapeutic genes to the target cells, and they are broadly categorized as viral and non-viral systems. Viral vectors are widely used in transgene delivery because of their natural ability to invade/transfect target cells very efficiently (Boudes, 2014;Rodríguez, 2004). In contrast, non-viral vectors have to be complexed with molecules such as cationic lipids and polymers to transfect specific target cells or must otherwise be forced into target cells using techniques such as electroporation or hydrodynamic injection. However, the transfection methods of non-viral vectors are rapidly being developed (Al-Dosari & Gao, 2009;Lukashev & Zamyatnin, 2016). As indicated in Figure 3, in clinical trials, viral vectors are most commonly used for gene delivery to target cells, because of their higher transfecting ability and efficient expressions in target cells. In comparison, the use of non-viral plasmid DNA vectors accounts for only about 11 % of the transgene delivery methods (Ginn et al., 2018). Even though viral vectors are used more commonly they are not without their own set of shortcomings such as; immunogenicity, cytotoxicity, generation of insertional mutagenesis (ectopic chromosomal integration of viral DNA), difficulty in vector synthesis and low transgenic capacity are few of the drawbacks of their usage in GT (Cooney et al., 2016;Glover et al., 2005).

Almost 70 % of the GT based clinical trials so far have used modified viral vectors such as lentiviral, retroviral, adenoviral and AAV vectors for transgene delivery (Yin et al., 2014). Lentiviral and retroviral vectors are designed to insert genes into the genome of target cells, but using these vectors has a higher risk of gene disruption (Loring et al., 2016). The use of AAV vectors also carries the risk of insertional mutagenesis. Even though the AAV vectors usually exist episomally inside target cells, they may still illegitimately integrate and may further cause issues of toxicity and immunogenicity (Lesch et al., 2009). Certain adenovirus-based viral vectors such

as adenovirus serotype 5 (AdV5) or adeno-associated virus type2 (AAV2) vectors cannot be used in transgene delivery because many people already have a pre-existing immunity against those viruses, leading to their rapid clearance upon administration. Another disadvantage with viral vectors is that the same serotype construct of a viral vector cannot be used repeatedly in transgene delivery in the same patient as they will build immunity against it. The death of Jesse Gelsinger who was in a phase 1 clinical trial to treat ornithine transcarbamylase (OTC) deficiency using adenoviral vectors, was due to a vector-mediated inflammatory response that led to organ failure. In another clinical trial, four patients developed T-cell leukemia due to undirected retroviral vector insertion. These safety concerns have stimulated heavy research in the development of safer vehicles for gene delivery (Ponder, 2003;Samson et al., 2008).

Compared to viral vectors, non-viral vectors are far less immunogenic. Plasmids have been recognized as non-viral gene delivery vectors since 1990. Plasmids are extrachromosomal, self-replicating, double-stranded, mostly circular DNA molecules (occasionally linear plasmids and also plasmids made from RNA exit) that are transferable from one bacterial cell to another. The plasmids are found in bacteria, algae and in some eukaryotes. These plasmid vectors are easier to produce, ship and store and also have a longer shelf life than viral vectors (Liu & Huang, 2002). Moreover, plasmid vectors are easy to manipulate, easy to design for use in therapeutic applications and can be used repeatedly into the same patient in transgene delivery (Ramamoorth & Narvekar, 2015;Rao & Zacks, 2014). Although non-viral vectors have drawn attention owing to properties such as biocompatibility, lower immunogenicity, and less cytotoxicity over viral vectors, the use of non-viral vectors in GT has been limited. This is primarily due to their poor transfection efficiency, which results in poor introduction and low transient expression in target cells. Recently, there has been an increase in the use of non-viral vectors, with noticeable

improvements in their transfecting efficiency, specificity, duration of the gene expression, and safety (Nafissi & Slavcev, 2012). Gene therapy via non-viral vectors offers immense therapeutic promise, however, both the safety as well as the efficiency of the delivery vectors must be improved in order to exploit their full therapeutic potential (Liu & Huang, 2002;Nafissi, Alqawlaq, et al., 2014).

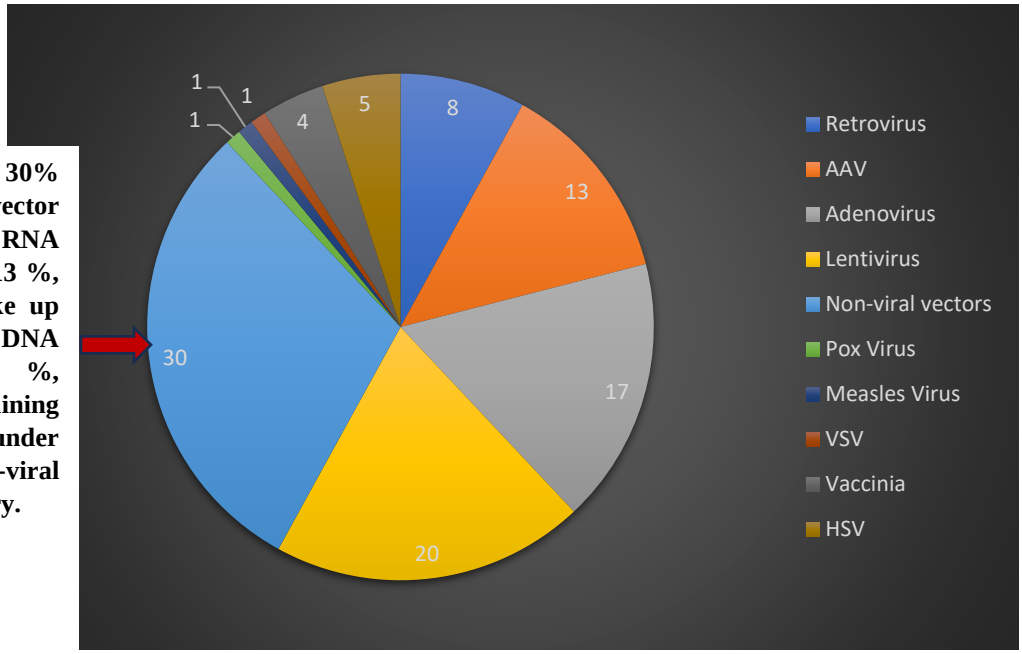


Figure 3:Vectors used in GT clinical trials from 2010 to 2020. During this period, viral vectors were used 70 % of the time, while non-viral vectors were employed in GT only 30 % of the time. Adapted from. Arabi et al., 2022.

1.3 Non-viral gene delivery vectors used in GT

In order for GT to be successful, the non-viral DNA vector carrying the gene of interest should traverse the cell membrane barrier, overcome cytoplasmic barriers such as endosomes, and enter the host nucleus through the nuclear membrane, and express the gene (Miller & Dean, 2009). Conventional, non-viral plasmid DNA vectors often result in the delivery of unwanted prokaryotic sequences, such as CpG motifs, antibiotic resistance genes, and the bacterial origins of replication

to the target, which may lead to unwanted immunological responses. Unmethylated CpG motifs present in bacterial plasmid DNA (exist abundantly in prokaryotic DNA) activate host Toll-like receptors and polyclonal B-cells which results in the activation of both innate and adaptive immune responses inside the host (Klinman et al., 1996). These plasmid vectors may also impart the potential for chromosomal integration, thus potentiating oncogenesis (Spies et al., 2003). Generation of mutations in the host proto-oncogenes, DNA repair genes, or tumor suppressor genes due to the non-target vector integration could cause cancer. The plasmid DNA vectors must also be immuno- and bio-compatible to be able to traverse cellular membranes. Modified gene delivery techniques such as synthetic carriers or physical methods are used to enhance transfection efficiency. This can be also achieved by changing the conformation and the composition of the DNA vectors (Darquet et al., 1997).

DNA minivectors: Smaller DNA minivectors devoid of unwanted prokaryotic sequences have shown great promise as non-viral GT vectors. Mini circular covalently closed (CCC) DNA vectors (DNA minicircles) are synthesized using bacteriophage λ integrase (Int)-*attP* or phage P1-derived Cre-*loxP* site-specific recombination systems (Darquet et al., 1997;Chen et al., 2003). DNA minicircles are devoid of unnecessary prokaryotic elements and are comprised of only the eukaryotic expression cassette. Therefore, they have been shown to enhance the immune and biocompatibility profiles of conventional plasmid vectors. These attributes coupled with their small size enhances the vectors' bioavailability and biocompatibility. Even though these mini vectors have higher nuclear translocation and extended gene expression capabilities compared to the conventional plasmid vectors, their circular profile could still potentially lead to vector insertional mutagenic events (Nafissi & Slavcev, 2012). In contrast, the use of linear vectors avoids

the risk of insertional mutagenesis since integration of a linear plasmid into a chromosome results in a cytotoxic double-stranded break (DSB) that is lethal to the cell (Figure 4).

Another modified set of plasmid vectors known as linear covalently closed (LCC) DNA vectors have been investigated. These torsion-free, double-stranded DNA molecules are not subject to degradation by ExoV exonuclease activity as the ends of the double-stranded DNA molecules are covalently closed (Hertwig et al., 2003). LCC DNA vectors impart a higher level of safety compared to their circular counterparts. As shown in Figure 4A, LCC DNA vector integration into the bacterial chromosome results in the disruption of the bacterial chromosome at the site of integration, which ultimately results in cellular death. Therefore, non-target vector integration of LCC vectors reduces the risk of insertional mutagenesis. In contrast, integration of the minicircle vector or plasmid (Figure 4B), into the non-target sites in the chromosome results in insertional mutagenesis as the integration does not break the chromosomal DNA. As a result, the cells continue to divide with the insert, perpetuating the mutation (Nafissi & Slavcev, 2012).

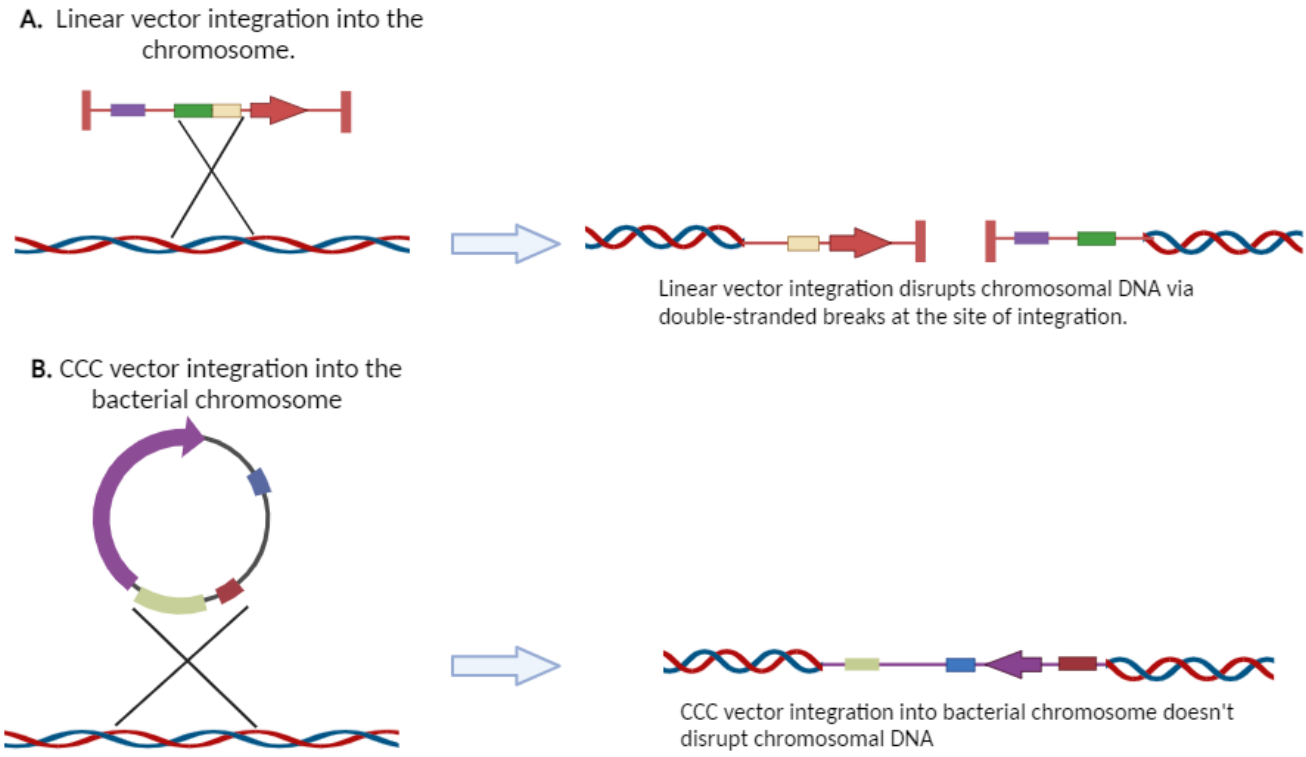


Figure 4:Vector integration into bacterial chromosome. A. Linear vector integration into the chromosome results in the disruption of the chromosomal DNA by making double-stranded breaks at the site of integration. **B.** Undirected, mini CCC DNA vector integration into the chromosomal DNA does not disrupt the chromosome, instead imparts molecular continuity. Adapted from Nafissi & Slavcev, 2012.

These LCC DNA vectors are prepared by using different *in vitro* systems. LCC minimalistic immunogenically defined gene expression (MIDGE) vectors (Figure 5A) are synthesized by endonuclease excision of the prokaryotic backbone and *in vitro* capping of the open ends with hairpin loops. These minimal DNA vectors are comprised of a eukaryotic expression unit (promoter, coding gene) and RNA stabilizing sequences (Schakowski et al., 2001). According to Darquet et al.,1997 reduction in the vector size increases the nuclear translocation of the vector, which results in higher gene expression rates. Compared to conventional plasmid vectors, MIDGE

vectors impart higher transfection efficiency and higher gene expression rates due to the efficient nuclear translocation of the smaller size MIDGE vectors. Another LCC DNA vector named micro-linear vector (MiLV) is synthesized similarly to MIDGE vectors with an additional step of further amplification of the LCC vector by polymerase-chain-reaction (PCR) (Figure 5B) (Darquet et al., 1997;Schakowski et al., 2001;Wang et al., 2012).

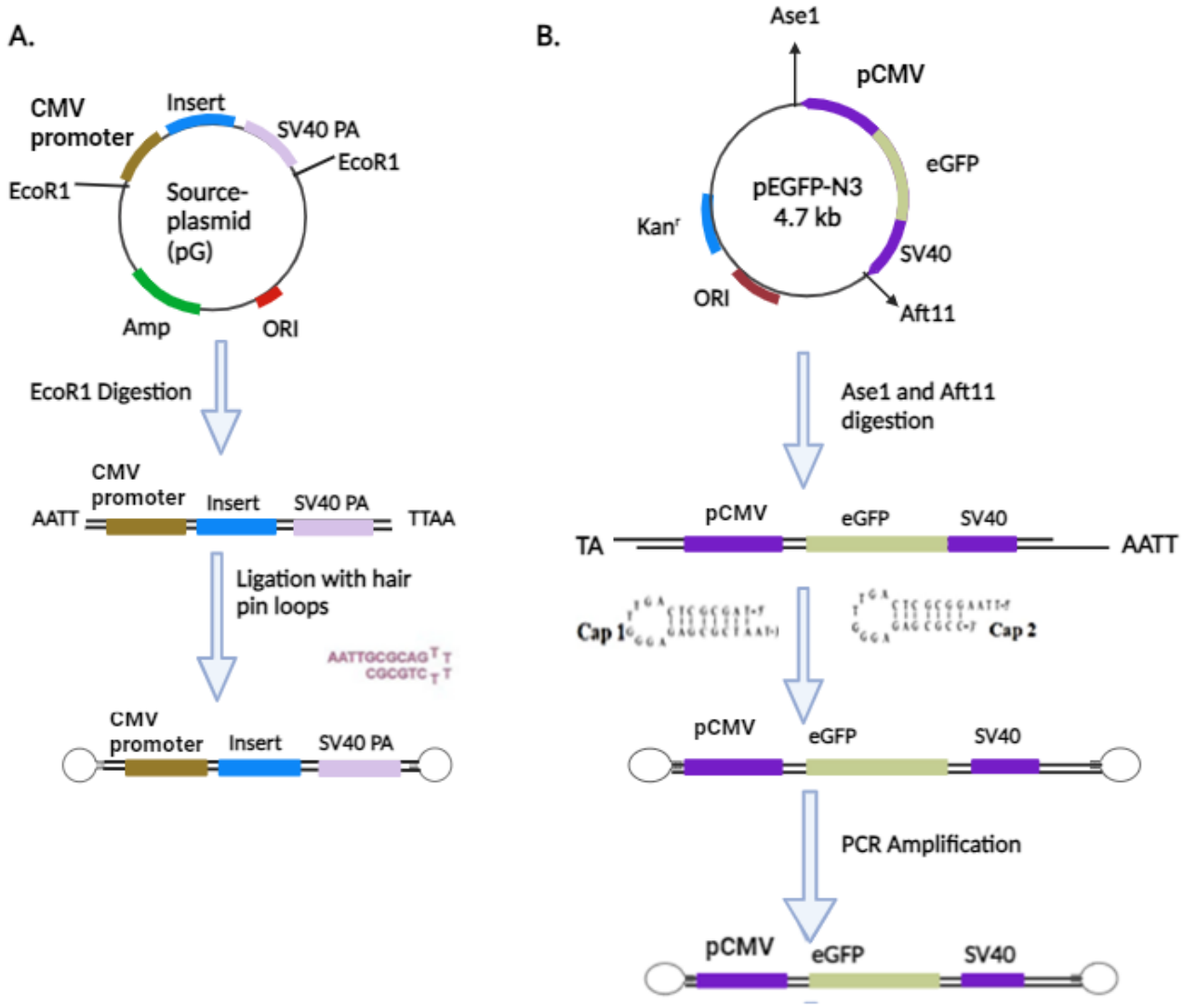


Figure 5:Development of MIDGE and MiLV vectors. A. MIDGE vectors are synthesized by the excision of the bacterial backbone of the parent plasmid by endonucleases; followed by ligation of the open linear ends with hairpin loops. **B.** MiLV vectors are synthesized similar to MIDGE

vector synthesis which is followed by PCR amplification of the synthesized MiLV. Adapted from Schakowski et al., 2001;Wang et al., 2012.

Large-scale *in vitro* production of MIDGE and MiLV vectors are costly and time-consuming due to the process requiring multiple steps to be completed. As such, there is a need for an efficient and inexpensive way to synthesize LCC vectors. Nafissi & Slavcev (2012) developed an inexpensive, *in vivo* recombinant platform for the synthesis of LCC DNA vectors, named DNA-ministrings (msDNA), exploiting the Tel/*Pal* recombination system (Nafissi & Slavcev, 2012). DNA ministrings are LCC DNA mini vectors used for high-efficiency transgene delivery. Like MiLV and MIDGE, these vectors contain only the gene of interest and regulatory sequences. Compared to conventional plasmid vectors, msDNA vectors are much safer as they lack unwanted immunostimulatory prokaryotic sequences and minimize the potential for malignant vector integration events. The smaller size of msDNA enhances transfection efficiency and also increases the copy number per unit mass. The LCC msDNA is redosable and possesses the highest level of sequence fidelity compared to in vitro synthesized versions of LCC vectors. Previous studies have shown that compared to conventional pDNA vectors, msDNA conferred enhanced *in vivo* transgene expression in target cells (Talebnia et al., 2023). Furthermore, modifying the LCC DNA vectors with DNA nuclear targeting sequences (DTS) and covalently linking nuclear localization signals (NLS) enhances nuclear translocation (Miller & Dean, 2009).

1.4 Synthesis of msDNA

The *E. coli* N15 bacteriophage is the first bacteriophage to exist as a LCC plasmid in its lysogenic (prophage) state which is achieved by the protelomerase enzyme TelN acting on the 56 bp *telRL* recognition sequence (Rybchin & Svarchevsky, 1999). The cleaving and joining activity

initiated by TelN protelomerase enzyme results in the generation of LCC double-stranded plasmid DNA (Grigoriev & Łobocka, 2001;Heinrich et al., 2002). Similarly, the bacteriophage PY54, derived from *Yersinia enterocolitica*, possesses a paralog of the N15 TelN protelomerase enzyme, named Tel. This enzyme acts on the 42 bp palindromic target sequence, leading to the formation of LCC plasmid DNA (Hertwig et al., 2003). DNA ministrings are produced in a recombinant platform *in vivo* by utilizing Tel-*pal* and TelN-*telRL* recombination systems. Tel-mediated recognition and binding to the *pal* target sequence result in a cleaving-joining activity of the precursor plasmid, producing LCC DNA vectors (Huang et al., 1997;Nafissi & Slavcev, 2012).

Recombinant *E. coli* cells (R-cells) are constructed by integrating the Tel or TelN protelomerase enzyme encoding gene cassette into the bacterial chromosome *lacZ* gene via homologous recombination. These Tel or TelN enzyme encoding genes are placed under the control of the strong λ *pL* and *pR* promoters. This promoter activity is regulated by the temperature-sensitive λ-repressor, CI[Ts]857. At 30 ^{0}C the promoter is at an inactive state as the λ-repressor abrogates the promoter activity and when shifting the temperature to 37 ^{0}C and above, dissociation of the λ-repressor, CI[Ts]857 results in the activation of the strong promoter activity (Figure 6).

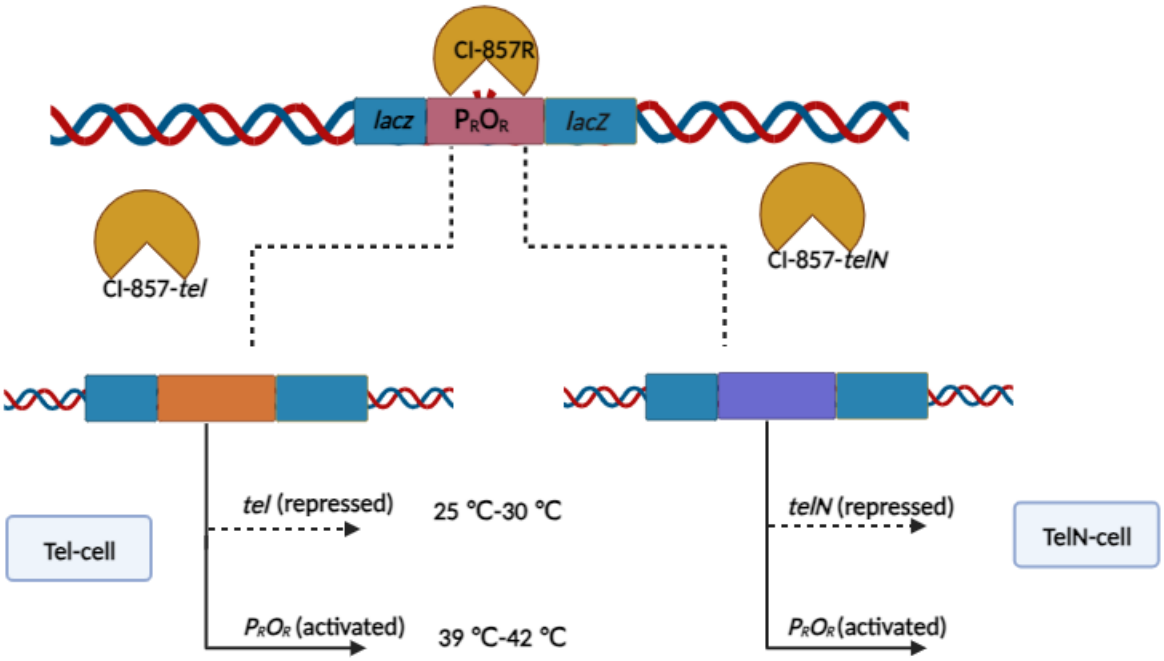
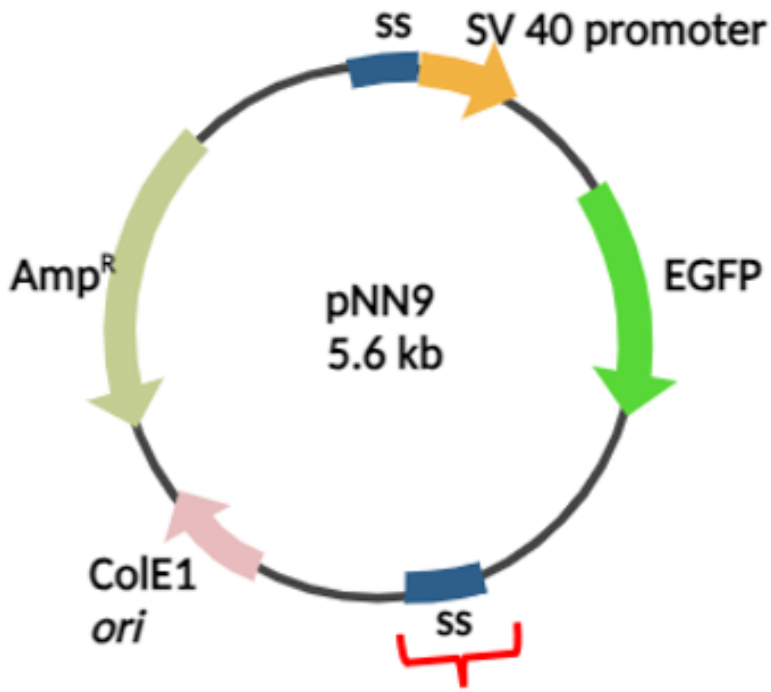


Figure 6:Temperature regulated expression of the Tel and TelN enzyme encoding genes in recombinant *E. coli* (R-cells) bacterial cells. Tel or TelN enzyme-encoding genes are under the control of the strong λ promoters *pL* and *pR* in recombinant cells. This expression cassette is integrated into the bacterial chromosomal *lacZ* gene via homologous recombination. The promoter activity is regulated by the temperature-sensitive λ- repressor CI[Ts]857. At 30 ^{0}C the promoter is in the repressed state as the λ- repressor binds to the operators and abrogates the promoter activity. When the temperature shifted to 42 ^{0}C the promoter gets activated as the repressor dissociates from the operators. Adapted from Nafissi & Slavcev, 2012.

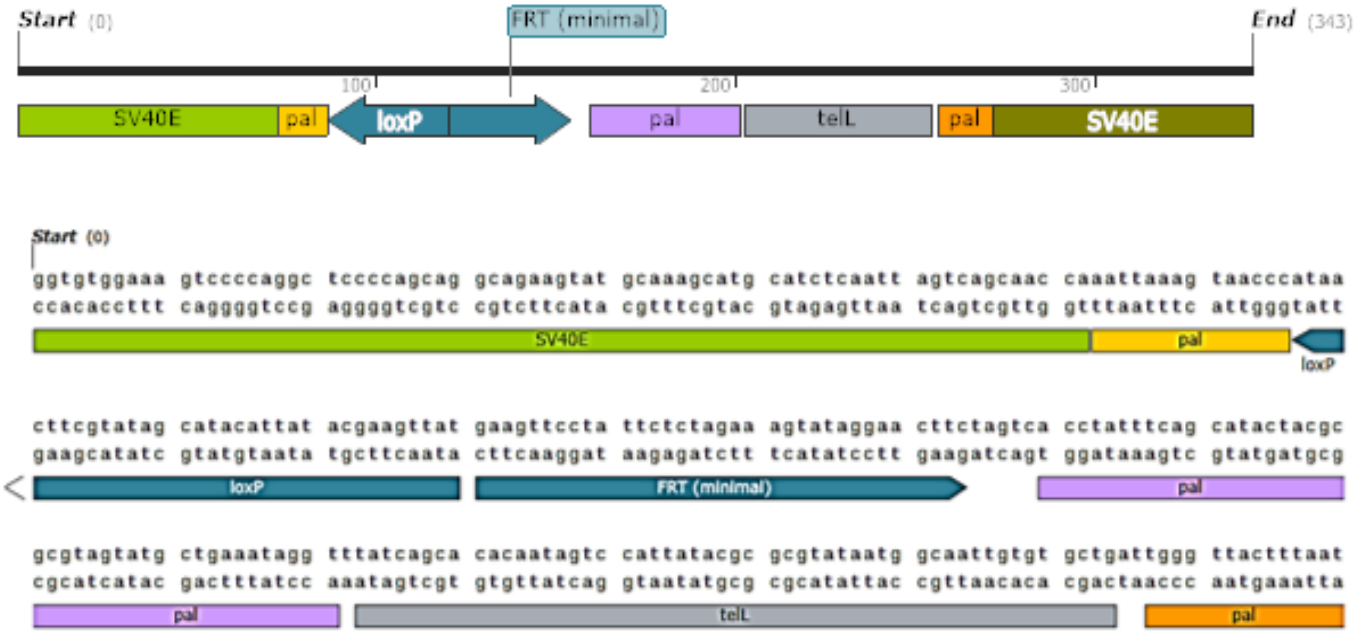
The plasmid carrying the enhanced green fluorescent protein (*egfp)* expression cassette as indicated in Figure 7 (referred to as pNN9) was constructed and transformed into the R-cell systems. This plasmid contains two 343 bp, multi-target sequences, named super sequence (SS), that are placed upstream of the SV40 promoter and downstream of the polyadenylation signal-sequence of the *egfp* expression cassette (Figure 7). This multi-target SS consists of a modified *pal* target site of Tel, to include *telRL* (TelN), and *loxP* (Cre) target sites between Tel-binding sites of *pal*. Both *pal* target sites in the SS are flanked by SV40 enhancer sequences, which enhances nuclear translocation. The purpose of having a parental

plasmid with SS, with multiple target sites for Tel, TelN, Cre enzymes is to introduce the same parental minimal expression plasmid construct into different R-cell systems for the generation of either mini LCC or mini CCC conformations of the minimal expression cassette. In the pNN9 plasmid, the gene of interest (GOI) is flanked by SS, and its bacterial backbone consists of an ampicillin resistance cassette (Ap^R) and a bacterial origin of replication. As shown in Figure 8, when the pNN9 precursor plasmid is introduced into *Tel /TelN* integrated R-cells, shifting the temperature to 42 ^{0}C activates the expression of *tel/telN* Telomerases. Tel/TelN acting on the *pal/telRL* target sites present in the pNN9 precursor plasmid results in the formation of LCC, msDNA (~2.6 kb in size) and also the LCC bacterial backbone (~3.0 kb) (Nafissi & Slavcev, 2012;Nafissi, Sum, et al., 2014).

A.


ss
SV 40 promoter
Amp^R
pNN9
5.6 kb
EGFP
ColE1
ori
ss

B.


Start (0)
FRT (minimal)
End (343)
100
200
300
SV40E
pal
loxP
pal
telL
pal
SV40E
Start (0)
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt agtcagcaac caaattaaag taacccataa
ccacaccttt caggggtccg aggggtcgtc cgtcttcata cgtttcgtac gtagagttaa tcagtcgttg gtttaatttc attgggtatt
SV40E
pal
loxP
cttcgtatag catacattat acgaagttat gaagttccta ttctctagaa agtataggaa cttctagtca cctatttcag catactacgc
gaagcatatc gtatgtaata tgcttcaata cttcaaggat aagagatctt tcatatcctt gaagatcagt ggataaagtc gtatgatgcg
loxP
FRT (minimal)
pal
gcgtagtatg ctgaaatagg tttatcagca cacaatagtc cattatacgc gcgtataatg gcaattgtgt gctgattggg ttactttaat
cgcatcatac gactttatcc aaatagtcgt gtgttatcag gtaatatgcg cgcatattac cgttaacaca cgactaaccc aatgaaatta
pal
telL
pal

Target Sequence	Start	End
SV40E	**1**	**72**
pal	**73**	**87**
loxP	**88**	**120**
FRT	**121**	**154**
pal	**160**	**201**
telRL	**203**	**256**
pal	**257**	**271**
SV40E	**272**	**343**

Figure 7:msDNA pNN9 (5.6 kb) precursor parental plasmid. A. The map of the PNN9 precursor plasmid, consisting of the *egfp* expression cassette flanked by SS, the origin of replication and the ampicillin resistance gene in the bacterial backbone of the plasmid. **B.** Multi target SS consisting of the target sites for Tel, TelN, and Cre recombinases. The modified *pal* target site is flanked by two SV40 enhancer sequences, and the rest of the target sites are positioned in the *pal* non-binding sites in the plasmid. Adapted from Nafissi & Slavcev, 2012.

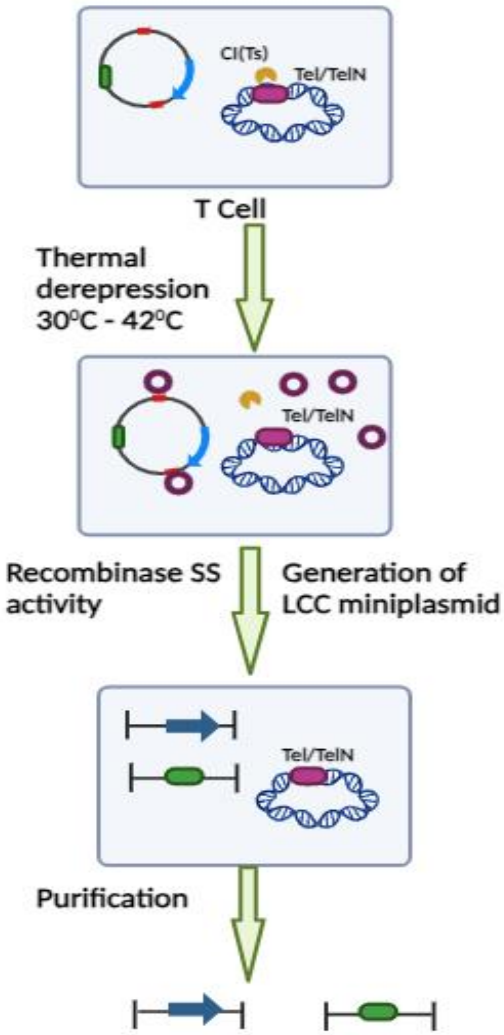


Figure 8:Synthesis of LCC msDNA vectors using the pNN9 precursor plasmid. The pNN9 precursor plasmid is transformed into the *Tel* /TelN integrated recombinant cells. Shifting the temperature in the growth media to 42 ^{0}C activates the expression of the protelomerase. Expressed Tel acting on the *pal* target sites present in the pNN9 precursor plasmid results in the synthesis of LCC, msDNA and LCC bacterial backbone. Adapted from Nafissi, Sum, et al., 2014.

Another msDNA synthesizing precursor plasmid, pDMS, used in this study, was constructed based on pNN9 precursor plasmid (Figure 9). The pNN9 and pDMS plasmids can be distinguished by the presence of specific components. Notably, the pDMS plasmid features a 128 bp PI-Sce1 homing endonuclease enzyme recognition sequence (containing 30 bp three tandem repeats of the homing endonuclease enzyme recognition sequence) and a 585 bp internal ribosome entry site

(IRES) sequence, as illustrated in Figure 9A. The eukaryotic minimal expression cassette in the pDMS plasmid includes a polylinker, which allows for the insertion of any desired gene in transcriptional fusion with the green, fluorescent reporter (GFP) gene (Wong et al., 2016). The bacterial backbone consists of beta-lactamase encoding gene for ampicillin-resistance and PI-SceI homing endonuclease recognition sequences, as indicated in Figure 9 A.

The *E. coli* K-12, W3110, and BW25113 bacterial strains are being used in this study. The *araB-araD* genes have been deleted from the genome of BW25113 bacteria, and a genetic alteration has been implemented in the lacZ gene on the chromosome of the BW25113 bacteria (Grenier et al., 2014). The *tel* gene has been integrated into the chromosome of BW25113 bacterial strain using bacterial phage P1 transduction. *E. coli* K-12 W3110 and BW25113 *E. coli* bacterial strains are well studied and have been widely used in many applications. As shown in Figure 9B, msDNA is synthesized in a simple one-step, heat-inducible process. The *tel*-integrated, genetically engineered W3110 and BW25113 R-cells are transformed by pDMS precursor plasmid. They are grown at 30 ^{0}C and their protelomerase enzyme expression is activated by shifting the temperature to 42 ^{0}C to inactivate the repressor. Protelomerase acting on the *pal* sites present in the SS synthesizes msDNA (~3.3 kb) and LCC bacterial backbone (3.1 kb). Finally, the plasmids are isolated using commercially available kits and the msDNA is agarose gel-purified at laboratory scale. The same procedure is employed to synthesize msDNA in the presence of the pNN9 precursor plasmid, resulting in the production of msDNA with a size of 2.6 kb, along with unwanted bacterial backbone with a size of 3.0 kb.

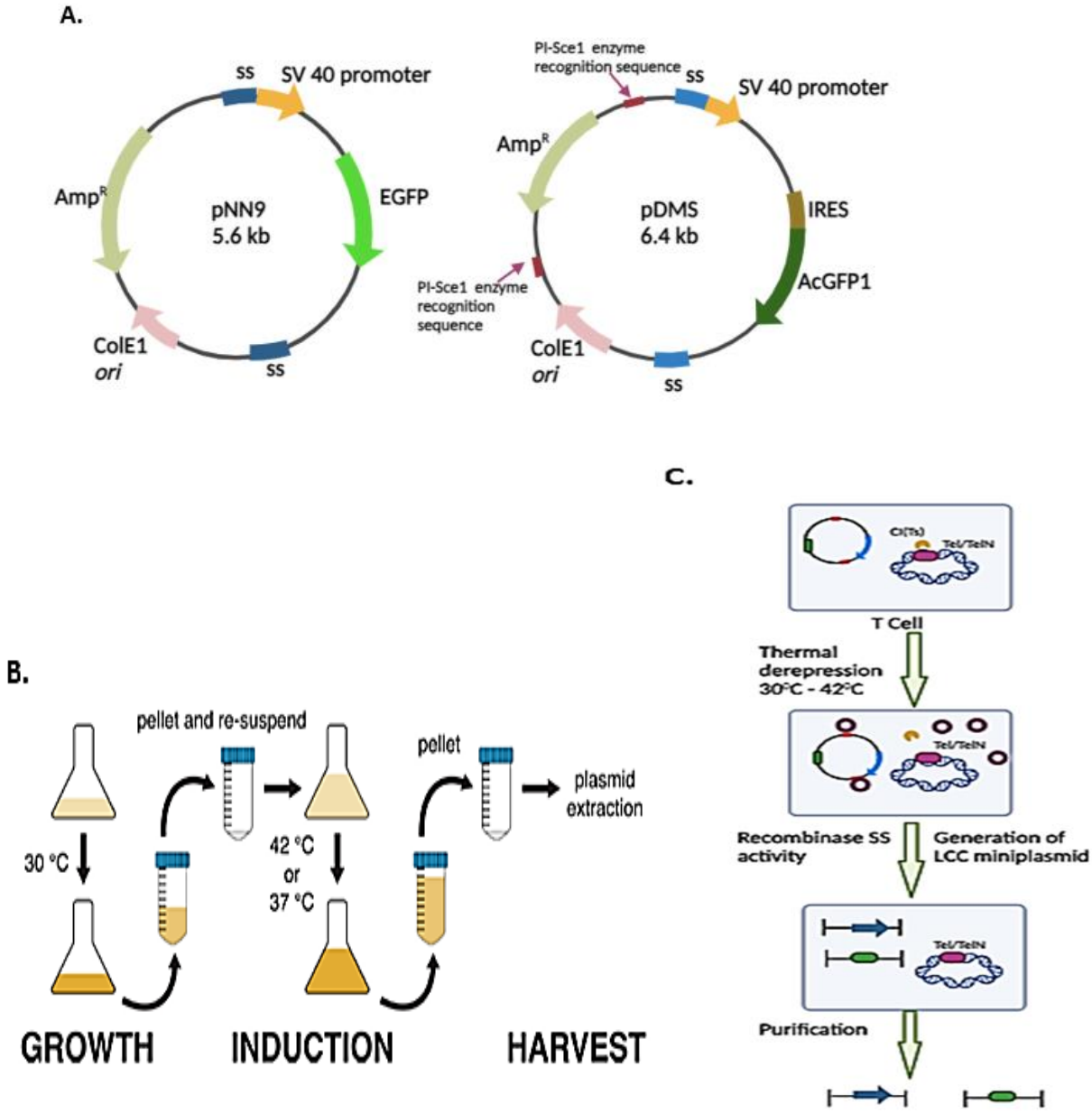


Figure 9:Schematic representation of the msDNA synthesis using the pDMS precursor plasmid. A. The circular maps illustrate the differences between the pNN9 and pDMS msDNA synthesizing precursor plasmids. Notably, pDMS contains the 585 bp IRES sequence and the 128 bp PI-Sce1 enzyme recognition sequence, resulting in the production of msDNA and unwanted bacterial backbones of different sizes. **B.** Schematic representation of the synthesis of msDNA, using a simple one step-heat inducible procedure. The pDMS transformed R-cells are grown in LB media at 30 ^{0}C and shifting the temperature in the media to 42 ^{0}C activates the expression of the *Tel* gene. Telomerase activity on the *pal* sites in the SS generates msDNA and the LCC bacterial backbone. Finally, the plasmids are extracted. **C.** When the temperature is

shifted to 42 ^{0}C, Telomerase is over expressed and acts on *pal* target site. As a result, it synthesizes 3.1 kb msDNA and 3.3 kb LCC bacterial backbone. Adapted from Wong et al., 2016.

1.5 Purification of msDNA

The LCC msDNA is synthesized from CCC precursor pDNA, using a heat-inducible, scalable, *in vivo* recombinant platform. Generally, pDNA extractions consist of different monomeric, topological isoforms such as: supercoiled CCC pDNA, open circular (OC) and linear open forms (LO) (Sum et al., 2014). In pDNA extraction, generally, the most predominant form is the supercoiled CCC pDNA. Figure 10 illustrates the isolated plasmid isoforms when run on an agarose gel, after msDNA synthesis, using pNN9 as the precursor plasmid. It is a mixture of different DNA isoforms such as open linear, supercoiled, LCC bacterial backbone, and the LCC msDNA. The LCC msDNA must be purified from the mixture. At the laboratory scale, the synthesized msDNA is purified via agarose gel extraction, while at larger scale synthesis, it is purified using anion exchange chromatography. In the anion exchange chromatography procedure, the negatively charged DNA binds to the positively charged ligands on the stationary phase. When a salt gradient is applied, this displaces the ionic interactions and as a result, elutes the DNA in the order of increasing charge density. The interactions between the DNA and the ligands on the stationary phase depend on the net charge and the charge density of the DNA molecule. Supercoiled CCC pDNA has a higher charge density, which binds strongly with the stationary phase and elutes last with the salt gradient. Anion exchange chromatography relies upon molecule separation based on charge density as a function of length, which is not always attainable when there are two or more species similar in size. DNA ministring purification based on size from agarose gels also has its own limitations; for instance, the plasmid has to be pre-treated with

enzyme to cut and degrade the un-wanted LCC bacterial backbone if the size of the plasmid backbone is similar to the msDNA. Moreover, the laboratory-scale purification of msDNA by agarose gel extraction provides low yields and bottlenecks the scalability of the manufacturing system (Chen et al., 2005;Sum et al., 2014).

In the process of minicircle production, unprocessed parental precursor plasmids and unwanted bacterial backbones are digested by restriction enzymes after minicircle synthesis. Next, the minicircles are purified via cesium chloride ultra-centrifugation. This purification process is time-consuming, expensive and labour-intensive. Therefore, to overcome these issues the investigators introduced a homing endonuclease I-Sce1 encoding gene (discussed below) along with the endonuclease recognition site into the minicircle synthesizing parental precursor plasmid. After the parental plasmid is converted into the minicircle DNA and the bacterial backbone DNA, endonuclease activity is induced to digest the backbone plasmid leaving only the episomal minicircle vector DNA in the bacteria. The minicircle is then isolated using commercially available one-step affinity columns (Chen et al., 2005). Similarly, for the current research one-step, homing endonuclease recombinant purification technology is being introduced in order to enhance the msDNA purification. The underlying principle of the one-step msDNA purification system is to destroy the unwanted bacterial plasmid DNA after protelomerase mediated msDNA synthesis, leaving msDNA as the only episomal DNA species in the cell after processing.

Another approach utilized in the study involves the utilization of the CRISPR-Cas 3 system, which is naturally present in *E. coli* K-12 bacteria. In this method, pre-crRNAs targeting the bacterial backbone portion of the precursor plasmids are expressed, and the cr-RNA guides the CRISPR-Cas3 cascade protein complex to bind with the targeted sequence, digesting the unwanted bacterial

backbone portion of the precursor plasmid. This approach results in the digestion of the unwanted byproducts of msDNA synthesis.

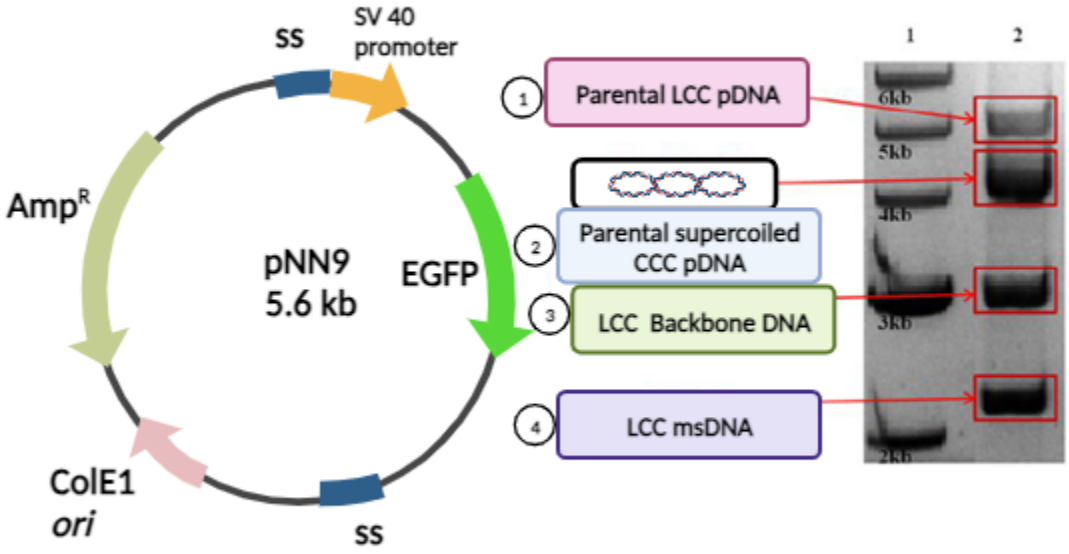


Figure 10:Isolated plasmid DNA extraction after msDNA generation using a pNN9 precursor plasmid on a 0.8 % agarose gel. There is a mixture of different DNA isoforms generated in the process of msDNA synthesis. 1. OL form of the parental CCC plasmid (~5.6 kbp) 2. Super coiled parental CCC plasmid (~4 kbp). 3.LCC bacterial back bone (3.0 kb) 4. LCC msDNA (2.6 kb). Adapted from Sum et al., 2014.

1.6 Homing endonucleases

Homing endonucleases are highly specific DNA cleaving, rare cutting enzymes (Grindl et al., 1998). According to Wende et al.,1996, homing endonucleases are found in eubacteria, archaea, and eukaryotes (Wolfgang Wende, 1996). Homing endonucleases are encoded by open reading frames (ORFs — DNA sequences that provide genetic instructions to synthesize proteins) embedded within introns and inteins. They are classified into four groups based on the following sequence motifs. They are; LAGLIDADG, GIY-YIG, H-N-H, and the His-Cys Box (Perler et al.,

1994). The introns in the gene are spliced out during RNA processing, and the remaining exons are joined together to form mature mRNA. The open reading frames inside these intervening intron sequences encode homing endonucleases, which recognize specific DNA sequences and cleave the DNA. Similarly, within the proteins, the intervening inteins are self-spliced out, and homing endonucleases are synthesized (Chevalier, 2001). Homing endonucleases transfer their host genetic materials to an allele that lacks introns or inteins through a process called homing. The nucleolytic enzyme PI-SceI belongs to the LAGLIDADG family and is naturally synthesized in *Saccharomyces cerevisiae* (Grindl et al., 1998). PI-SceI, previously known as VDE (*VMA1* derived endonuclease), results from a protein-splicing reaction. *VMA1* gene encodes a 119 kDa precursor protein which undergoes protein-splicing, leading to the formation of a 69 kDa vacuolar membrane-bound H+-ATPase and a 50 kDa endonuclease (Gimbles et al., 1993).

PI-SceI recognizes larger DNA sequences (depending on the substrate, the recognition sequence varies and may exceed 40 bps) and nucleotide site recognition is not very stringent like in type II restriction enzymes. In the presence of divalent cations such as Mn^{2+} and Mg^{2+}, PI-SceI cleaves the double-stranded DNA molecule into two DNA fragments. This endonuclease is not only a nucleolytic enzyme but also acts as a splicing enzyme (Gimble et al., 2003). The nucleolytic activity of the enzyme is due to the presence of two LAGLIDADG motifs. According to Thorner and Gimble. (1992), PI-SceI recognizes a 30 bp sequence and the recognition site is not stringent as in other restriction enzymes. Therefore, single base-pair changes do not affect the cleavage of the target by endonucleases. The recognition site indicated in Figure 11 is one of the known recognition sites of this Homing endonuclease (Gimble & Thorner, 1992).

Figure 11:PI-SceI homing endonuclease 30 bp recognition sequence and the cleavage site. The 30 bp PI-SceI recognition sequence is as highlighted. PI-SceI, unlike other restriction enzymes, does not have a very stringent recognition and they do not have a defined cleavage site. The one shown above is a common cleavage site known so far, but it could have several other cleavage sites, which are still unknown. A single base-pair mutation of the recognition site does not affect the binding of the enzyme to the target site. Adapted from Gimble & Thorner, 1992.

1.7 Conditional-replication, integration and modular (CRIM) plasmids

Recombinant enzymes are overexpressed by cloning the enzyme encoding genes into expression plasmids, but the high-copy number artifacts generation due to this could adversely affect physiological studies (Hasan et al., 1994;Metcalf et al., 1994,Glick, 1995). Therefore, integrating genes into the bacterial chromosome helps the studying of the corresponding gene function in single copies and this leads to the creation of novel *E. coli*, bacterial strains stably expressing multiple foreign genes in the same cell (Landy & Ross, 1977). In this study CRIM plasmids are used to integrate the PI-SceI homing endonuclease encoding gene into the bacterial genome. These plasmids consist of an R6K origin of replication, a lambda phage attachment (*attP*) site, and an antibiotic-resistance marker. The R6K origin of replication requires the transacting protein π, which is synthesized in pir^+ bacterial cells, to replicate (Haldimann & Wanner, 2001;Izsvák et al., 2009). CRIM plasmids integrate into the bacterial chromosome through site-specific recombination which takes place between the phage attachment site present in the CRIM plasmid and the corresponding bacterial attachment (*attB*) site present in the bacterial chromosome of pir^- bacteria. Figure 12 illustrates the schematic representation of the integration of the integrative plasmid via site-specific recombination into the bacterial chromosome. In order to

integrate the plasmid via site-specific recombination, the phage derived integrase (Int) and the integration host factor (IHF) encoded by *E. coli* is required (Haldimann & Wanner, 2001;Martinez-Morales et al., 1999;Zucca et al., 2013).

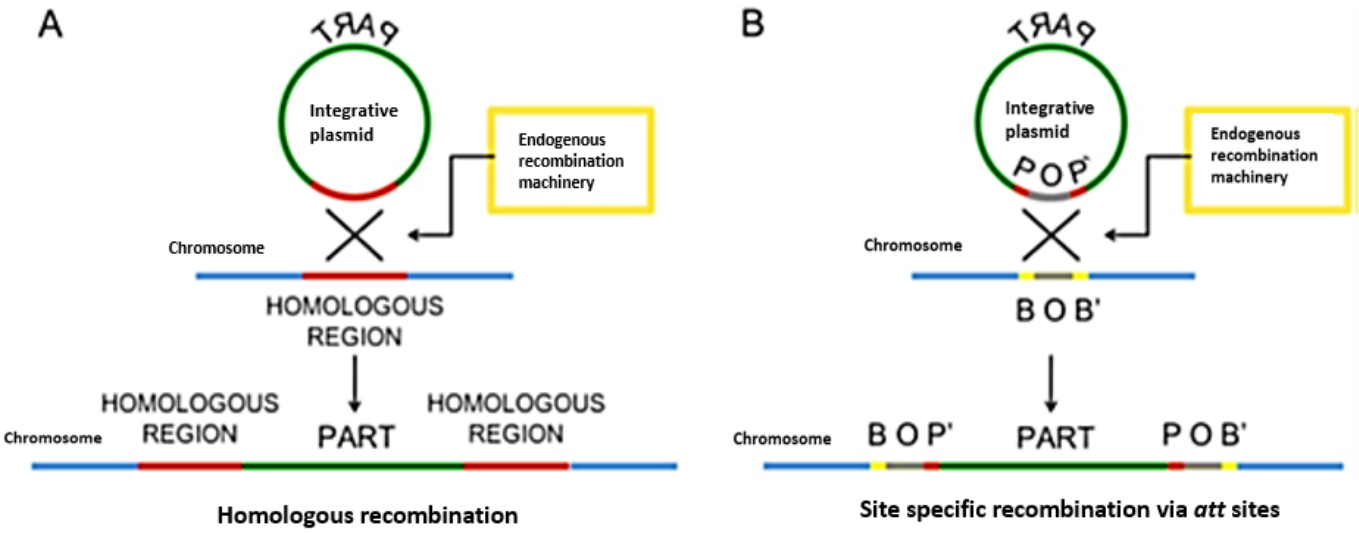


Figure 12:Integration of the bacterial plasmid via site-specific recombination into the bacterial chromosome. A. In the integrative plasmid, the region highlighted in red is homologous with the region highlighted in red in the bacterial chromosome. After a single crossover event generated by the bacterial recombination system, the whole plasmid is integrated into the bacterial chromosome. In the RecA$^-$ *E. coli* cells the homologous recombinases are provided by a helper plasmid. **B.** The integrative plasmid has the phage attachment site (*attP*) which is homologous with the bacterial attachment site present in the bacterial chromosome (*attB*), where site-specific recombination takes place. The homologous region (common core) "O" is flanked by two different arms of *attP* (P and P') and *attB* (B and B') sequences. The single crossover event generated by the phage integrases encoded by helper plasmids integrate the plasmid into the bacterial chromosome. Adapted from Zucca et al., 2013.

Generally, bacteriophages insert their DNA into the bacterial chromosome either through transposition like phage Mu or through site-specific recombination as for phage λ. The phage λ inserts their DNA at a specific location in the chromosome, through site-specific recombination. Also, certain plasmids integrate into the bacterial chromosome by site-specific recombination. This concept was utilized when generating CRIM plasmids; this is by inserting different *attP* sites into the CRIM plasmid to facilitate integration at various phage

attachment sites in the bacterial chromosome. The site-specific recombinases, belonging to the integrase family, perform site-specific recombination activity between the *attP* site in the CRIM plasmid and the corresponding *attB* site present in the bacterial chromosome (Argos et al., 1986;Xu et al., 2001). In order to insert the λDNA to the bacterial chromosome, there has to be a minimum 21 bp long *attB* and a 234 bp *attP*. As shown in Figure 13, the highlighted 15-bp λ *attB* core gene sequence is identical to the 15-bp gene sequence present in the 234-bp *attP* site (Campbell, 1992).

C	T	**G**	**C**	**T**	**T**	**t**	**t**	**T**	**t**	**a**	**t**	**A**	**c**	**t**	**A**	**A**	C	T	T	G
G	A	**C**	**G**	**A**	**A**	**a**	**a**	**A**	**a**	**t**	**a**	**T**	**g**	**a**	**T**	**T**	G	A	A	C

Figure 13:The highlighted 15 bp λ *attB* sequence which is identical to the *attP* site. Adapted from Campbell, 1992.

There are five different phage attachment sites that are used by phages λ, HK022, φ80, P21, and P22 for integration. The location of the corresponding bacterial attachment sites present in the chromosome of the wild type *E. coli* K-12 is as indicated in Figure 14. The corresponding core gene sequences of *attB* of different phages are shown in Table 1. CRIM plasmids are integrated into the bacterial chromosome by supplying integrase (Int) enzymes in trans. This is achieved by transforming helper plasmids carrying the Int enzyme-encoding genes. The integrated CRIM plasmids can be retrieved from the bacterial chromosomes by supplying excisionase in trans by transforming helper plasmids carrying the Xis-encoding genes (Barzel et al., 2011;Boyd et al., 2000;Haldimann & Wanner, 2001).

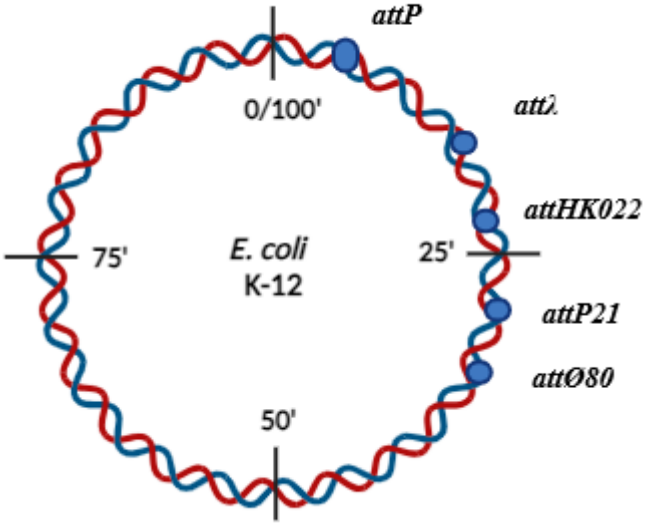


Figure 14:The chromosomal location of bacterial *attB* present in wild type *E. coli* K-12 bacteria. Adapted from Haldimann & Wanner, 2001.

Phage	Core gene sequence of the corresponding *attB* present in the bacterial chromosome.
λ (36)	Ct**gCTTttTtatActAA**cTTG
HK022 (58,73)	gtga**CTTTaggtgaa**tAAGGtTgA
P21 (65)	AgATGa**tGCtGCgcCATAT**
P22 (37)	CagCGC**ATTcgtAATGcGAAG**
Φ80 (37)	**AACAcTTTcttAAAt**gGTT

Table 1:Core *attB* sequences (highlighted) of various lambdoid phages. Adapted from Haldimann & Wanner, 2001.

1.8 CRISPR-Cas3 system of *E. coli* K-12 bacteria

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) – CRISPR-Associated Systems (Cas) are RNA-dependent immune systems used by bacteria and archaea to defend against nucleic acids from invading viruses and mobile genetic elements (MGE) (Sorek et al., 2013). CRISPR-Cas systems are found in both bacteria and archaea and consist of one or more CRISPR arrays and Cas genes (Xue & Sashital, 2019). The CRISPR array is comprised of target-specific spacer sequences alternating with repeat sequences (Bolotin et al., 2005). The spacer sequences are acquired from nucleic acids present in invading viruses. The Cas genes encode Cas proteins which form the effector complex. CRISPR immune responses in bacteria can be broken down into three stages. Adaptation, expression and maturation and interference. In the adaptation process, foreign DNA sequences are acquired (protospacers) and inserted between the repeat sequences as "spacers" in the CRISPR array (Barrangou et al., 2007). The memory within the realm of adaptive immunity is established through the acquisition of new spacers. In the expression and maturation stage, the long transcripts of the CRISPR array (pre-crRNA) go through processing to form a mature functional truncated crRNA (Carte et al., 2008). In the CRISPR interference, the crRNA acts as a guide to recognize incoming nucleic acids by binding to complementary protospacer sequences present in the target sequence. The binding of crRNA with the effector complex guides the effector complex to the target sequence (Marraffini & Sontheimer, 2008). The nucleases present in the effector complex digest the target sequence (Figure 15) (Xue & Sashital, 2019).

CRISPR-Cas systems are divided into two main classes (Class 1 and Class 2), six types, and 30 subtypes. The primary distinction between the two classes lies in the composition of the effector complex. Types I, II, and IV belong to Class 1 of the CRISPR-Cas system, where the effector

complex is composed of a multiprotein cascade consisting of several Cas proteins. Class 2 effector complexes are comprised of a single multidomain Cas protein (Murugan et al., 2017).

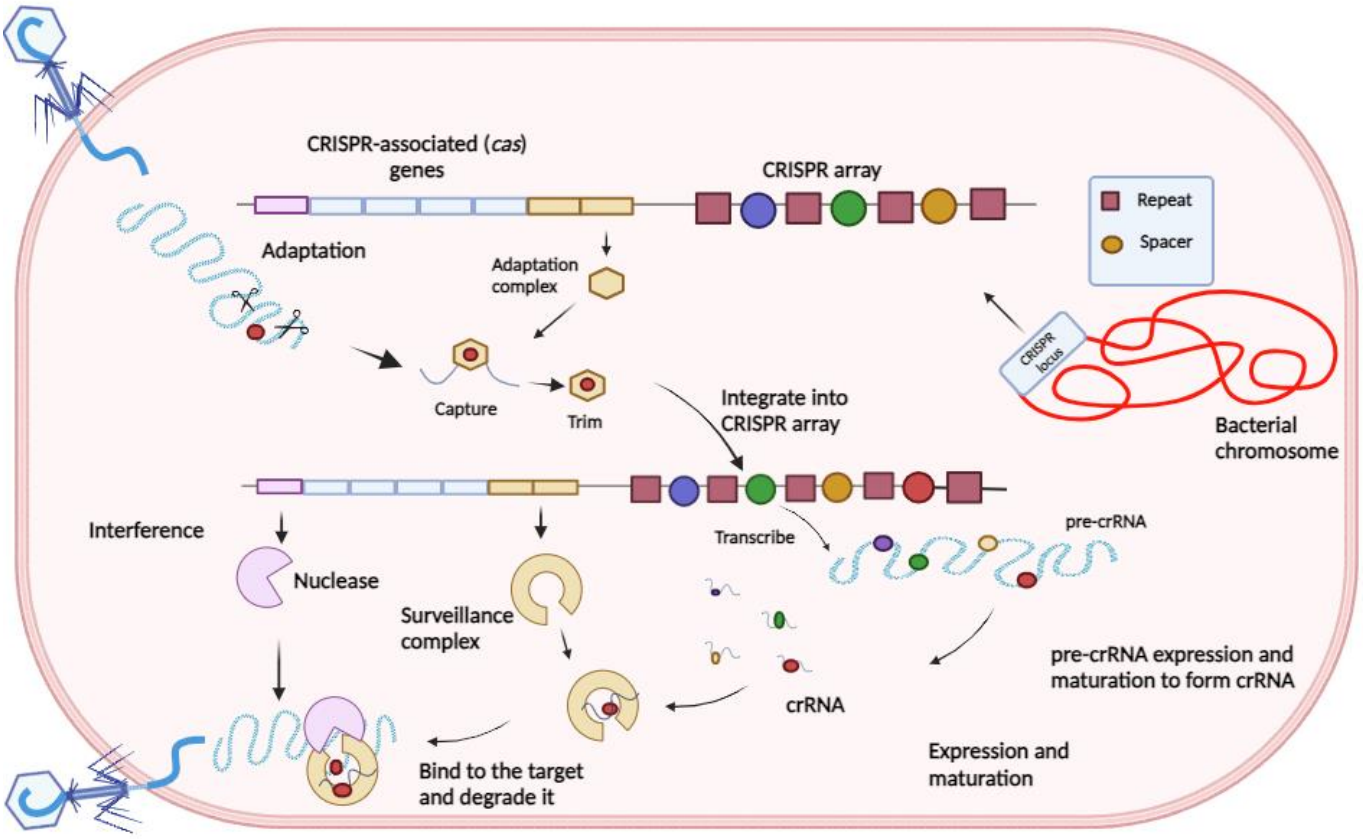


Figure 15:Stages of CRISPR immune responses. CRISPR immune response can be broken down into three stages: Adaptation, expression and maturation and interference. Adaptation: when bacterial cells invade by a phage, the phage DNA pieces are recruited by adaptation complex and integrate into the CRISPR array. Expression and maturation; CRISPR array transcribed the long pre-crRNA transcript. Upon processing it forms a matured crRNA consisting of spacer sequences flanked with partial repeat sequences. Interference: crRNA binds with Cas proteins and form an effector surveillance complex, where it binds with the invading phage DNA. Nucleases present in the effector complex degrade the DNA. Adapted from Xue & Sashital, 2019.

Due to the simplicity of the composition of the effector protein complex of the Class 2 system, they are mostly used in biotechnology for customizable applications for the purpose of cleavage of specific DNA/RNA nucleic acids. Even though the Class 2 CRISPR systems are simpler, 90 %

of bacteria and archaea have a Class 1 system (Makarova et al., 2015). The classification of different types of CRISPR systems is based on their mechanisms of action and the type of target nucleic acid. For instance, Type I systems cleave and degrade DNA; Type II systems cleave DNA and type III systems cleave DNA and RNA. Type I and II systems require the following factors for effective targeting: complementarity between the crRNA spacer sequence and the protospacer sequence present in the target nucleic acid, and the presence of a protospacer adjacent motif (PAM) upstream to the protospacer sequence (Makarova et al., 2011) (Figure 16 B).

1.8.1 Type 1E CRISPR-Cas system

E. coli K-12 bacteria has a Type 1 E CRISPR-Cas system, and its CRIPSR-Cas locus consists of two CRISPR arrays and a *cas* operon (Figure 16A). The *cas* operon expresses eight CRISPR-Cas genes (*cas 1,2,3,5,6,7,8e* and *11*). Cas1 and Cas2 proteins are involved in CRISPR adaptation. Upon expression of the *cas* genes, a cascade complex is formed. The multi subunit cascade complex is required for crRNA binding with the target sequence, cleavage, and degradation of the target nucleic acids (Brouns et al., 2008).The pre-CRISPR RNA is transcribed unidirectionally by the promoter sequences present in the leader sequence, upstream of the CRISPR array. The pre-CRISPR RNA, upon processing, results in the generation of mature CRISPR RNAs (crRNA) consisting of an individual spacer sequence flanked by partial repeat sequences. The pre-crRNA forms stem-loop structures in the repeat sequences. The Cas6 endoribonuclease recognizes the stem-loop structures and cleaves it at the base of the stem (Figure 17). Following the cleaving of the stem-loop structure at the base, Cas6 continues to bind to the 3' end of the crRNA. Subsequently, the rest of the Cas proteins assemble and form the cascade complex. The cascade protein in the effector complex binds with the AAG PAM sequence in the target nucleic acid. The Type 1 E CRISPR-Cas system in *E. coli* recognizes 5’ AAG 3’/ 3’ TTC 5’ PAM sequences.

Subsequently, the spacer of the crRNA binds with the complementary protospacer sequence and forms a DNA-RNA hybrid and initiates the cleavage or degradation of the target nucleic acids within minutes of invasion. This is called CRISPR interference (Gasiunas et al., 2012). As shown in Figure 18, the spacer sequence binds with the target strand (T-strand) of the protospacer sequence and unwinds the target DNA. Cas 3-helicase-nuclease is recruited by the cascade protein through protein-protein interactions. Cas3 first nicks the non-target (NT-strand) strand of the protospacer sequence and subsequently digests the T strand at multiple locations and ultimately degrades the target sequence. There are two types of adaptation that take place in *E. coli,* known as naïve and primed adaptation. In naïve adaptation, the new spacers are acquired from existing or foreign DNA sequences that are present inside the cell, which have the 5' AAG 3'/3' TTC 5' PAM motifs. In primed adaptation, the new spacers are acquired from sequences surrounding already existing spacers targeting protospacers sequences. This way, bacteria strengthen and update its memory in order to fight against specific invaders in the future (Shiriaeva et al., 2022).

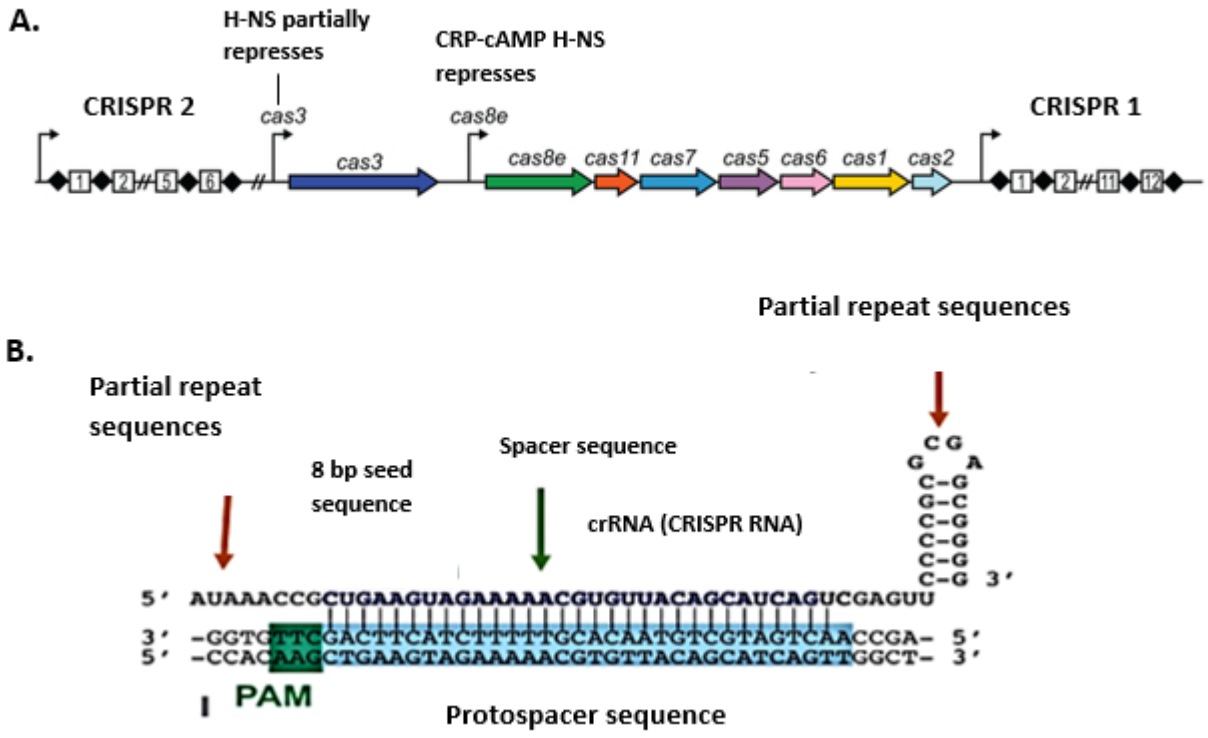


Figure 16:Bacterial CRISPR-cas system. A. Type 1-E CRISPR-Cas operon of *E. coli* K-12 bacteria. The promoters are indicated from arrows. The repressors of *cas3* and *cas8e* are represented from arrows. The CRISPR array consists of spacer sequences flanked by repeat sequences followed by leader sequence where promoter sequences are present **B.** Upon processing of the pre-crRNA forms shorter crRNAs consisting of single spacer sequence flanked by partial repeat sequences. The spacer sequence is complementary to the target protospacer sequence. The cascade protein in the effector complex binds with the AAG PAM sequence of the *E. coli* target sequence. The 8 bp seed sequence in 5' end of the crRNA is very critical in guiding the effector complex bind with the target sequence. Adapted from Gomaa et al., 2014;Xue & Sashital, 2019.

1.8.2 Strategies employed by phages to evade CRISPR immunity

Even though CRISPR-Cas immunity is a robust defensive mechanism used by bacteria to fight against invading bacterial phages, phages have escaped CRISPR immunity (van Houte et al., 2016). Research is being conducted using the model organism Enterobacteria to test the CRISPR immune response through phage-challenge and plasmid-based assays. The escape mutations generated in the protospacer sequences, especially within the PAM sequence and the seed region,

inhibit cascade complex binding to the target sequence. Even the presence of a certain number of mismatches between the spacer sequence and the target protospacer sequence doesn't disrupt effective targeting. The mismatches in the "seed region" (8 bases in the 5' end of the spacer sequence) of the spacer sequence, disrupts effective targeting (Jiang, Bikard, et al., 2013).

Anti CRIPSR proteins encoded by phages inactivate the CRISPR immune response. Anti-CRISPR genes have been found in the genome of prophages and they encode small proteins which interfere with CRISPR interference. This is by blocking the CRISPR-Cas cascade complex binding to the target sequence (Borges et al., 2017). Research indicates that initially, when a bacterium becomes infected with a phage, the CRISPR immune system effectively eliminates the phage. However, upon repeated exposure of the phage to bacteria, the phage begins to produce enough anti-CRISPR proteins that thwart the CRISPR defenses, preventing elimination (Dong et al., 2017).

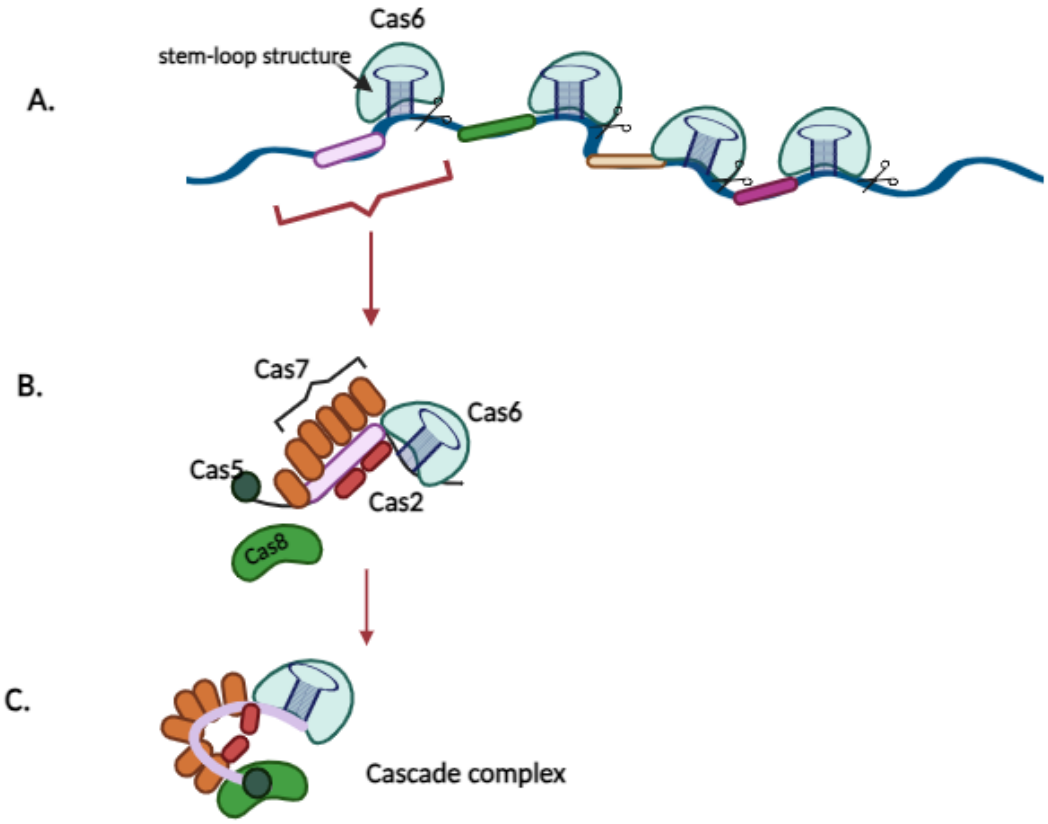


Figure 17:Transcription and maturation of pre-crRNA. A. Transcription of the CRISPR array results in the synthesis of long pre-crRNA transcript. The stem-loop structures are formed in repeat sequences. Cas6 endoribonulcease cut the at the base of the stem-loop structure. Following the cleaving of the stem-loop structure at the base, the Cas6 continued to bind to the 3' end of the crRNA**. B-C.** Subsequently the rest of the cas proteins assembled in sequence to form the cascade complex. Adapted from Xue & Sashital, 2019.

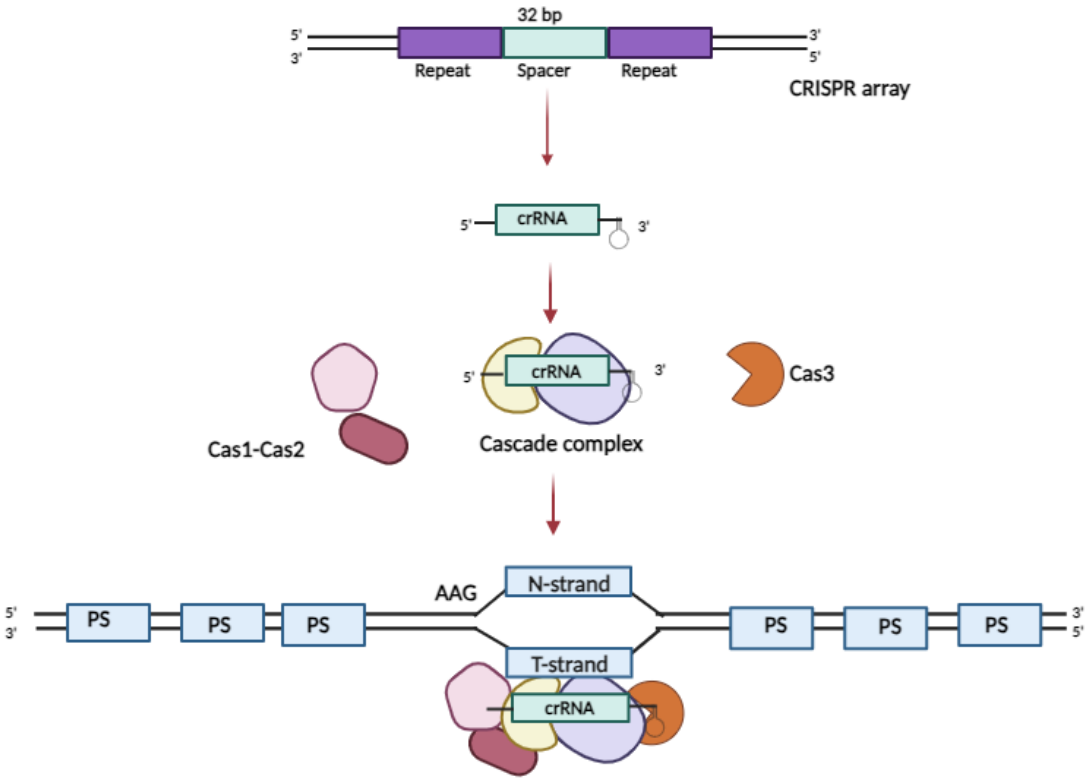


Figure 18:CRISPR interference. CRISPR array consists of spacer sequences flanked by repeat sequences. The crRNA transcribed from CRISPR array binds with the cascade protein to form the effector complex. This effector complex binds with the target protospacer sequence which has complementary sequence with the spacer sequence. The effector complex first binds with AAG PAM sequence and then the spacer sequence forms the DNA-RNA hybrid with the target T strand. Cas3- helicase protein is recruited by the cascade protein via protein-protein interactions. Subsequently Cas3 nicks first the non-target N-strand and the T-strand at multiple locations. Thereby degrading the target DNA. Adapted from Shiriaeva et al., 2022.

2. Rationale, hypothesis, and objectives

2.1 Rationale

Considering the advancement of msDNA is expanding, larger scale synthesis and cost-effective purification is a necessity to eventually use this in clinical applications. The *in vivo* recombinant platform to produce msDNA results in a mixture of unprocessed precursor plasmids, unwanted prokaryotic LCC plasmids, and their topological isoforms (such as supercoiled CCC, circular open DNA and linear open DNA). Therefore, purifying msDNA vectors from other DNA species and isoforms is a crucial step in the purification process. Currently, at the laboratory scale, msDNA is purified using electrophoretic approaches, which results in a significant loss of DNA yield. Additionally, use of standard plasmid isolation, followed by treating the bacterial lysate with commercially available restriction enzymes to purify msDNA, is not scalable in the mid-to-large scale setting. This is primarily due to the fact that conducting these downstream steps is both time-consuming and expensive. Therefore, the need to optimize this process is necessary.

For this project, we focus on the development of an *in vivo* recombinant system to digest unwanted prokaryotic precursor plasmid and byproduct DNA species upon processing of msDNA from precursor parental plasmid in an attempt to improve purification efficiency, using the P1-Sce1 homing endonuclease to isolate msDNA. This project focuses on integrating the homing endonuclease gene (*vma*), encoding the PI-Sce1 homing endonuclease, under the control of an inducible promoter (a system that can be efficiently and rapidly turned on and off) into the bacterial host chromosome as a single copy cassette.

An additional approach that will be investigated to purify msDNA from unwanted DNA species is the use of the RNA guided CRISPR-Cas 3 system; an endonucleolytic system that is naturally

present in *E. coli* bacteria, to target and degrade the unprocessed precursor plasmid and unwanted bacterial backbone.

Hypothesis: Recombinant bacteria engineered to encode and express *vma* homing endonuclease under the control of an inducible promoter will improve the purification of msDNA by digesting any unprocessed, non msDNA species following msDNA processing *in vivo*. Furthermore, recombinant bacteria carrying a pre-CRISPR RNA encoding genetic element targeting endogenous Cas3 to the precursor plasmid backbone will improve the purification of msDNA by digesting any unprocessed, non msDNA species following msDNA processing *in vivo*.

2.2 Objectives:

The main objective of the research project is to develop a recombinant system which degrades all the unwanted DNA by products after msDNA synthesis leaving only msDNA intact, thereby simplifying purification.

1. Integration of the PI-SceI homing endonuclease encoding gene (*vma*) into the bacterial chromosome.
2. Expression of the *vma* homing endonuclease gene upon integration into the *E. coli* genome to determine whether the introduced homing endonuclease imparts any adverse effects on the synthesis of msDNA.
3. Integration of the *vma* gene into JM109-*tel* recombinant bacteria via P1-transduction and induction of msDNA synthesis, followed by over expression of the *vma* gene.

4. Episomal expression of the Tel protelomerase enzyme to induce msDNA synthesis in recombinant bacteria.

5. Digestion of the msDNA synthesizing precursor plasmid upon over expression of the pre-crRNA targeting the bacterial backbone.

6. Determination of msDNA synthesis in pre-crRNA synthesizing recombinant bacteria.

3. Materials and methods

3.1 Strains and plasmids

A complete list of plasmids, bacterial strains and their genotypes used for this study can be found in Table 2.

Strain / Plasmid	Genotype	Source
	Bacteria	
DH5α	F-, *Δ (argF-lac)169, ö80dlacZ58(M15), ÄphoA8, glnV44(AS), λ- , deoR481, rfbC1, gyrA96(NalR), recA1, endA1, thi-1, hsdR17*	Coli genetic stock center (CGSC) # 12384
W3110	F-, λ-, *IN (rrnD-rrnE)1, rph-1*	CGSC # 4474
W3110 *vma*$^+$	F-, λ-, *IN (rrnD-rrnE)1, rph-1; Km-vma*	This study
W3110 *tel*$^+$	F-, λ-, *IN (rrnD-rrnE)1, rph-1 lacZ: Cm-CI857-tel*	This study
W3110 *tel*$^+$*vma*$^+$	F-, λ-, *IN (rrnD-rrnE)1, rph-1 lacZ: Cm-CI857-tel ;araC-araBAD-vma*	This study
BW25113	F-, *Δ(araD-araB)567, ΔlacZ4787(: rrnB-3), λ-, rph-1, Δ(rhaDrhaB)568, hsdR514*	CGSC #7636
BW25113 *vma*$^+$	F-, *Δ(araD-araB)567, ΔlacZ4787(: rrnB-3), λ-, rph-1, Δ(rhaDrhaB)568, hsdR514; Km-vma*	This study
BW25113 *tel*$^+$	F-, *Δ(araD-araB)567, ΔlacZ4787(: rrnB-3), λ-, rph-1, Δ(rhaDrhaB)568, hsdR514, lacZ: Cm-cI857-tel*	This study
JM109	F' *traD36 proAB+ lacIq lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14− recA1 gyrA96 relA1 hsdR17*	New England Bio labs (NEB)
JM109 *tel*$^+$	F' *traD36 proAB+ lacIq lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14− recA1 gyrA96 relA1 hsdR17; Cm-CI857-tel*	Mediphage Bioceuticals
JM109 *tel*$^+$*vma*$^+$	F' *traD36 proAB+ lacIq lacZΔM15/ Δ(lac-proAB) endA1 glnV44 thi-1 e14− recA1 gyrA96 relA1 hsdR17; Cm-CI857-tel-vma*	This study
BL21(DE3)	F$^-$*ompT hsdSB* (r_B^-, m_B^-) *gal dcm (DE3)*	Mediphage Bioceuticals
	Plasmids	

pAH120	*attP* λ integration plasmid (Kn^R)	National Bio Resource Project (NBRP)
pInt(ts)	λ *ts-int* λ (Ap^R)	Addgene
pACYC184	Cloning vector*;tet*^{R}*,Cm*^{R}	Accession# X06403.1 Addgene
pACYC184-CI-857-tel	*CI-857-PL-tel-PR -Cm*^{R}	This study
pACYC-tel-T7	pACYC-Duet-T7-*tel*	Mediphage Bioceuticals
pSW3	pAH120; *vma*	This study
pDMS	pGL2-*egfp*;(2XSS) (second SS downstream of *SV40* polyA sequence); 3 tandem PI-SceI target sequences integrated	This study

Table 2:Bacteria and plasmids used in experiments

3.1.1 Media and culture conditions

Luria-Bertani (LB) broth (pH 7.0) complex media was used to grow bacterial strains. Super optimal broth (SOB) and super optimal broth with catabolite repression (SOC) were prepared, as outlined in Cold Spring Harbor Protocols (2018). Antibiotics used to select for plasmid maintenance and their concentration were as follows: ampicillin (Ap) (100 µg/mL; aqueous), chloramphenicol (Cm) (25 µg/mL; ethanol), and kanamycin (Km) (50 µg/mL; aqueous). Single copy integrants (chromosomal integration of the *vma* gene) were selected using Cm (6 µg/mL) and Km (10 µg/mL) (Haldimann & Wanner, 2001).

3.2 Plasmid construction

3.2.1 Construction of pSW3 plasmid

The *vma* (PI-SceI homing endonuclease encoding) gene (NCBI accession no# NC_001136.10), was previously cloned into the pAH120 integrating plasmid, referred to as, “pSW3”, by Dr. Shirley

Wong (Figure 19). Using site-specific recombination between the *λ attP* site of the plasmid and the corresponding bacterial *attB* attachment site of the bacterial chromosome, the entire integrating plasmid (including the *vma* expression cassette) was integrated into the bacterial chromosome. The glycerol stocks of DH5α bacteria carrying the pSW3 plasmid were streaked on agar plates supplemented with Km and incubated at 37 ℃ overnight. LB broth (5 mL) supplemented with 2.5 µL of Km was inoculated with a single colony and grown at 37 ℃ overnight at 250 RPM. The culture was then centrifuged for 5 min at 10,000 RPM, and the cell pellet was collected. The plasmids were extracted from the pellet using the Monarch Plasmid Miniprep Kit (New England Biolabs (NEB), Ipswich, USA). The presence of the *vma* gene in the pSW3 plasmid was confirmed through gene-specific polymerase chain reaction (PCR) using forward (5' AACCGTTGGCCTCAATCG 3') and reverse (5' GATGGGCCCTTAGCAATTATGG 3') primers. Amplification was performed using One Taq 2X PCR Master Mix NEB for 30 s at 94 ^{0}C for initial denaturation, followed by 30 cycles each of 15 s at 94 °C, 15 s at 52 °C, 1 min 48 s at 68 °C, and 5 min at 68 °C for the final extension to generate a 1832 bp fragment. The presence of the *vma* gene in pSW3 was further confirmed using Sanger Sequencing (Centre for Applied Genomics (TCAG), Toronto, CA).

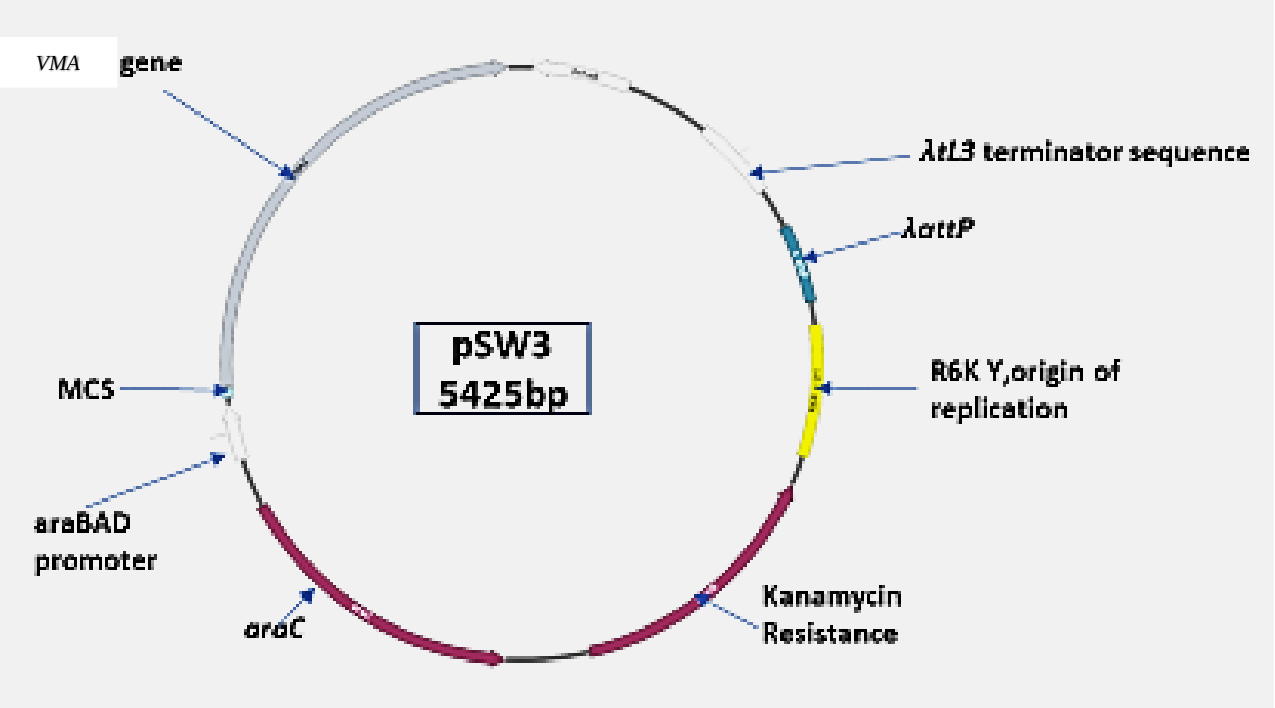


Figure 19:pSW3 plasmid map used in the study. The PI-SceI endonuclease encoding gene**,** ***vma*****,** is under the control of the P_{BAD} promoter. This *vma-araC-*P_{BAD} gene cassette was cloned into the pAH120 integrating plasmid. The construct is named pSW3, which is 5,425 bp in size.

3.2.2 Construction of the pACYC184-CI-857-tel plasmid construct

The Tel protelomerase enzyme encoding *tel* gene was cloned into the prokaryotic inducible plasmid, pPL451, by Dr. Nafiseh Nafissi in 2012 (Nafissi & Slavcev, 2012). The cloned gene was placed under the control of strong promoters (λ *pL* and *pR)*, under which the expression of the cloned gene was regulated by the heat-labile λ repressor, CI857. The 3.4 kb *cI857-pL-tel-tL gene* cassette was amplified using forward (5'- *CCCGGG*TATGTTATGTTCTGAGGGGAGTG -3') and reverse (5' - *GAGCTC*AAGTTGGGTAACGCCAGG - 3') primers. The italicized regions denote Xma1 and Sac1 restriction enzyme sites, respectively. The 3.4 kb insert was PCR amplified using Phusion Flash High-Fidelity PCR master mix from Thermo Fisher Scientific (Waltham, USA). The PCR conditions are outlined in Table 3. The pACYC184 (4.2 kb)

(Accession: X06403.1) low copy number plasmid was subjected to restriction enzyme digestion using NEB (Ipswich, USA) EcoRV restriction enzyme, followed by blunt end ligation of the PCR amplified *cI857-pL-tel-tL* (3.4 kb) insert. The PCR amplified *cI857-pL-tel-tL* insert was phosphorylated using T4 Polynucleotide Kinase from NEB. The EcoRV digested pACYC184 plasmid (vector) was dephosphorylated using NEB Shrimp Alkaline Phosphatase. The phosphorylated *cI857-pL-tel-tL* insert and the dephosphorylated pACYC184 vector were ligated using NEB T4 DNA ligase. The ligation was performed at a 5:1 insert:vector ratio and was incubated overnight at 16 ℃. Strain DH5α was transformed with 10 µL of the ligation mixture via electroporation. Tetracycline sensitive and chloramphenicol resistant (Cm^R) colonies were picked and grown at 30 ℃ overnight to isolate plasmids. The presence of the insert in the pACYC-CI-857-tel recombinant plasmid (Figure 20) was confirmed by digesting the plasmid with Xma1 and Sac1 restriction enzymes from NEB. The cloning was further confirmed by sending the constructed recombinant plasmid for sanger sequencing (Centre for Applied Genomics (TCAG), Toronto, CA).

Cycle Step	**Temperature (℃)**	**Time (Second)**	**Cycles**
Initial denaturation	98	10	1
Denaturation	98	1	30
Annealing	62.5	5	
Extension	72	51	
Final Extension	72	60	1
	4	Hold	

Table 3:Cycling instructions for amplification of the 3.4 kb gene fragment.

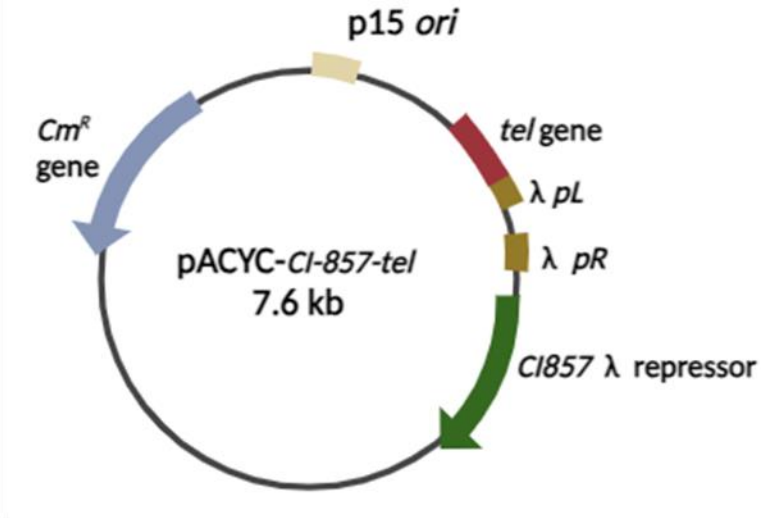


Figure 20:pACYC-CI-857-tel recombinant plasmid. The *tel* gene was placed under the control of strong *pL-pR* promoters, and the overexpression of the *tel* gene is regulated by the temperature-sensitive CI-857 λ repressor. A chloramphenicol-resistant recombinant pACYC-CI-857-tel plasmid was constructed by cloning a 3.4 kb insert into the 4.2 kb pACYC184 vector plasmid using the traditional restriction enzyme digest-ligation method. The resulting recombinant plasmid is 7.6 kb in size.

3.2.3 Construction of pDMS-tel recombinant plasmid

The Tel protelomerase enzyme encoding *tel* gene expression cassette was PCR amplified from the pPL451 recombinant plasmid, as outlined in section 3.2.2. The cloning of the *cI857-pL-tel-tL* cassette into the pDMS msDNA synthesizing precursor plasmid was performed in two steps. First, the 3.4 kb *cI857-pL-tel-tL* cassette was PCR amplified from the pPL451 plasmid and blunt-end ligated into the pUC18 plasmid (2.7 kb) (NCBI accession# L08752) (Norrander et al., 1983). The pUC18 plasmid was digested with Sma1(NEB, Ipswich, USA), which disrupted the *lacZ* gene segment in the plasmid. The digested pUC18 plasmid was subjected to dephosphorylation by rSAP from NEB. The 3.4 kb insert was phosphorylated using NEB T4 Polynucleotide Kinase. The dephosphorylated vector and the phosphorylated insert were ligated at a 5:1 insert:vector ratio using NEB T4 DNA ligase. The ligation reaction mixture consisted of 50 ng of vector and 315 ng

of insert, mixed in a final volume of 20 µL. Electrocompetent DH5α cells (50 µL) were transformed with 10 µL of the ligation reaction. The blue/white screening on IPTG + X-gal plates was performed to isolate positive transformants. Recombinant bacteria carrying pUC18-tel appeared white on IPTG + X-gal plates. The insert was isolated from the pUC18-tel recombinant plasmid by digesting the recombinant plasmid with Sma1 and Sac1 restriction enzymes. The digestion mixture was run on a 0.8 % agarose gel and the insert was gel extracted using the NEB Monarch DNA gel extraction kit. The double-digested insert was ligated with the double-digested (both Sma1 and Sac1) pDMS vector using NEB T4 DNA ligase enzyme. Electrocompetent DH5α cells were transformed with 10 µL of ligation reaction and spread onto selective LB + agar plates supplemented with Ap. The plates were incubated overnight at 30 ^{0}C.

3.2.4 Construction of the CRISPR Cas3 pre-crRNA expressing plasmids

The 90 bp pre-crRNA encoding gene sequence was cloned into the pACYC-pBAD-tel-Duet-1 plasmid via PCR. The forward and reverse primers were designed by including parts of the 90 bp sequence in their 5' ends. The pACYC-araC-tel-Duet-1 plasmid is 6,795 bp in size. Primers were designed to amplify the entire plasmid, avoiding the *tel* gene. The forward and reverse primer sequences indicated in Table 4 were used to insert the pre-crRNA encoding 90 bp gene sequence into the pACYC-araC-tel-Duet 1 plasmid. Each primer contains a part of the crRNA sequence and a sequence that binds with the pACYC-tel-pBAD-Duet1 plasmid upstream of the P_{BAD} promoter. For PCR amplification, Bio-rad (Hercules, USA) C1000 Thermal Cycler was used. The PCR reaction contained 12.5 µL of the Thermo Fischer Scientific (Waltham, USA) Phusion High Fidelity PCR Master Mix, 10 ng of plasmid DNA, 0.1 µM of forward and reverse primers, and sterile water within a final volume of 25 µL. Table 5 describes the settings used for the amplification of the targeted gene product. The PCR products were run on a 0.7 % agarose gel at

90 V for 1 h and 30 min. The successful insertion of the pre-crRNA encoding gene sequence into the pACYC-Duet 1 plasmid, should produce a ~ 5 kb PCR fragment upon amplification (Figure 21). The PCR product was gel purified using the Qiagen (Hilden, Germany) QIAquick Gel Extraction Kit. The concentration and the purity of gel-extracted DNA was measured using the Thermo Fisher Scientific (Waltham, USA) NanoDrop 2000.

PRIMER	SEQUENCE
Target- 1 FP	5’ ATCCAGTTCGATGTGAGTTCCCCGCGCCAGCGGGGATAAACCGCTGAAACCTCAGGCATTTGAG 3’
Target- 1 RP	5’ CTCAACAGCGGTAAGATCCGGTTTATCCCCGCTGGCGCGGGGAACTCTGGATTTTCATTATATCTCCTTC 3’
Target- 2 FP	5’ TCCATAGGCTCCGGAGTTCCCCGCGCCAGCGGGGATAAACCGCTGAAACCTCAGGCATTTGAG 3’
Target-2 RP	5’ AAAACGCCAGCAACGCGGCCGGTTTATCCCCGCTGGCGCGGGGAACTCTGGATTTTCATTATATCTCCTTC 3’

Table 4:Forward and reverse primers used to generate the pre-crRNA encoding gene sequence. The designed two spacer sequences are referred to as target 1 and 2. Target 1 protospacer sequence is in the Ap resistant gene of the precursor plasmid. Target 2 protospacer sequence is in the *ori* of the precursor plasmid. The text highlighted in green, and blue indicates the sequences of the forward and reverse primers which bind to the pACYC-araC-tel-Duet-1 plasmid. The text highlighted in pink and grey are the 90 bp target 1 and 2 sequences attached to the 5’ ends of the forward and reverse primers. Each forward and reverse primer contains approximately half of the crRNA sequence.

Cycle step	Temperature	Time	Number of cycles
Initial Denaturation	98 ^{0}C	10 s	1
Denaturation	98 ^{0}C	1 s	
Annealing	56 ^{0}C (using target 1 primers) 67.7 ^{0}C (using target 2 primers)	5 s	
Extension	72 ^{0}C	1 min 20 s (Target 1) 1 min 23 s (Target 2)	30
Final extension	72 ^{0}C	60 s	1
Incubation	4 ^{0}C	Incubate	N/A

Table 5:PCR steps for insertion of the pre-crRNA encoding gene sequences into the pACYC-araC-tel-Duet-1 plasmid.

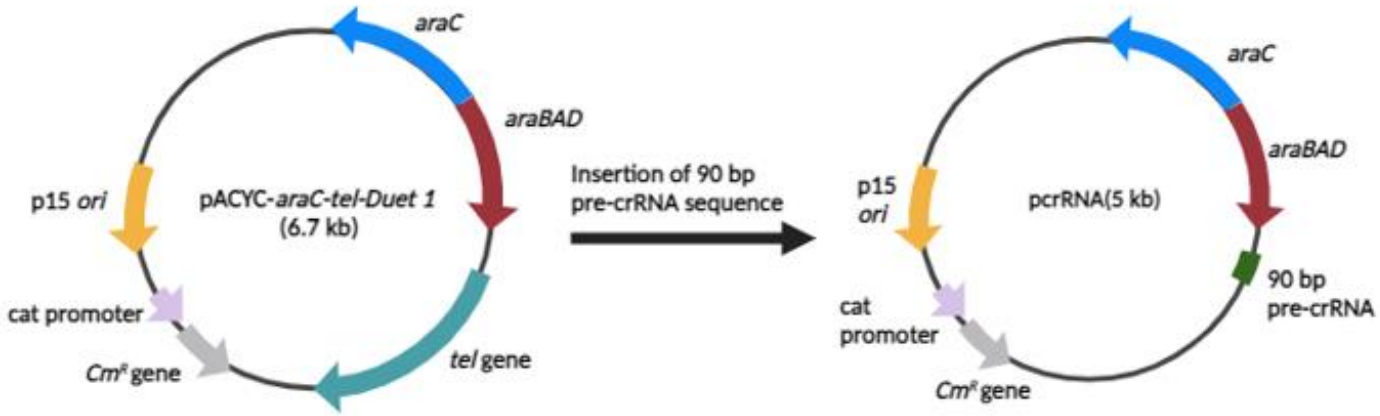


Figure 21:Generation of pre-crRNA expressing plasmid. The successful insertion of the pre-crRNA results in generating a 5 kb pcrRNA plasmid.

3.2.4.1 Phosphorylation and cleanup of PCR amplified product

The gel extracted DNA fragment of the amplified pACYC-Duet 1 plasmid, flanked with crRNA sequences, was phosphorylated using the NEB (Ipswich, USA) T4 Polynucleotide Kinase (PNK). Each reaction contained 300 ng of DNA, 5 µL of 10X T4 PNK buffer, 1 µL of T4 PNK enzyme, 5 µL of 10 µM ATP and sterile water in a final reaction volume of 50 µL. The phosphorylation was conducted to phosphorylate the 3' ends of the amplified DNA fragment to enable successful ligation. The reaction was carried out on ice and incubated at 37 ^{0}C for 30 min. The reaction mixture was heat inactivated by incubating at 60 ^{0}C for 20 min. The phosphorylated PCR product was purified using the NEB Monarch PCR & DNA Cleanup kit.

3.2.4.2 Ligation and transformation of the phosphorylated PCR product

The phosphorylated PCR amplified DNA fragment containing the pre-crRNA encoding gene sequences, was ligated using the NEB (Ipswich, USA) T4 DNA ligase in a 20 µL final reaction volume. The reaction contained 300 ng of DNA, 2 µL of 10X T4 DNA Ligase Buffer and sterile water. The reaction was carried out on ice and incubated at 16 ^{0}C overnight. The reaction was then inactivated by incubating at 65 ^{0}C for 10 min. Electrocompetent DH5α cells were then transformed with 10 µL of the ligation mixture.

3.2.4.3 Confirmation of the insertion of the pre-crRNA encoding gene sequence

Two plasmids were constructed: one referred to as, "pcrRNA90", which expresses the Target 1 pre-crRNA sequence, and the other referred to as, "pcrRNA120", which expresses the Target 2 pre-crRNA sequence. To confirm the insertion of the 90 bp pre-crRNA encoding gene sequence in the pACYC-araC-Duet 1 plasmid, positive colonies on LB + Cm agar plates were picked and cultured in 5 mL of LB broth + Cm overnight at 37 ℃. The plasmids were extracted using the NEB (Ipswich, USA) Monarch Plasmid Miniprep Kit. The extracted plasmid was digested with the High

Fidelity (HF) EcoRV restriction enzyme obtained from NEB. The 25 µL final reaction volume contained 500 ng of DNA, 2.5 µL of 10X rCutSmart Buffer, 0.5 µL of HF-EcoRV enzyme and sterile water. The digestion reaction mixture was incubated at 37 ℃ for 15 min and run on a 0.8 % agarose gel to visualize the fragment size difference between the pACYC-araC-tel-Duet 1 plasmid and the generated pcrRNA plasmids. The insertion of the pre-crRNA encoding gene sequence was further confirmed by sequencing the generated pcrRNA plasmids. The primers were designed flanking the inserted pre-crRNA gene sequence and sent along with the plasmids to the Sick Kids Centre for Applied Genomics TCAG Facility (Toronto, CA) for Sanger sequencing. The sequence of the primers flanking the pre-crRNA gene sequence in the constructed pcrRNA plasmids were 5' TCCTACCTGACGCTTTTTAT 3' (forward primer) and 5' CTACCGGAAGCAGTGTGAC 3' (reverse primer).

3.2.5 Construction of pcrRNA120-Km^R plasmid (p120K)

Gibson assembly was used to replace the Cm^R gene in the CRISPR plasmid with the Km^R gene. The primers to perform Gibson assembly were designed using the Benchling software. The primers used in the experiment are shown in Table 6. The primers were designed to amplify the entire CRISPR plasmid excluding the Cm^R gene (vector). The insert (Km^R gene) was amplified from the pSW3 plasmid and ligated with the PCR amplified vector fragment via Gibson assembly. The PCR conditions to perform the amplifications are outlined in Table 7. For PCR amplification, the Bio-rad (Hercules, USA) C1000 Thermal Cycler was used. The PCR reaction was conducted using the pcrRNA120 and pSW3 plasmids. Each PCR reaction contained 25 uL of the Thermo Fisher Scientific (Waltham, USA) Phusion High Fidelity PCR Master Mix, 10 ng of plasmid DNA, 0.5 µM forward and reverse primers, and sterile water within a final volume of 50 µL. The PCR amplified DNA fragments were run on a 0.8 % agarose gel for 1 h and 30 min and gel extracted

using the Qiagen (Hilden, Germany) QIAquick Gel Extraction Kit. The concentration and purity of gel-extracted DNA was measured using Thermo Fisher Scientific (Waltham, USA) NanoDrop 2000. Gibson assembly was performed using the NEB (Ipswich, USA) NEBuilder HiFi DNA assembly master mix. Each reaction contained 10 µL of the NEBuilder HiFi master mix, a combined total of less than 0.2 pmol of vector and insert DNA, and sterile water in a final volume of 20 µL. The mixture was incubated at 50 ^{0}C for 1 h in the Bio-rad (Hercules, USA) C1000 Thermal Cycler. Electrocompetent DH5α bacteria was transformed with 10 µL of the assembly mixture and grew on Km (50 µg/mL) supplemented LB agar plates. Positive colonies were picked and grown overnight in 5 mL of LB broth supplemented with Km (50 µg/mL). The concentration of the Km added in the experiment is as outlined in section 3.1.1. The next day plasmids were extracted from the cultures.

Primers	Sequence
1. FP-pcrRNA120	5’ gaaacgatcctcatcctgtcttttagcttccttagctcctg 3’
2. RP-pcrRNA120	5’ gataccgtcaaaacttcattagggcagggtcgttaaatag 3’
3. FP-pSW3	5’ ctatttaacgaccctgccctaatgaagttttgacggtatcg 3’
4. RP-pSW3	5’ aggagctaaggaagctaaaagacaggatgaggatcgtttc 3’

Table 6:Primers used in Gibson assembly to replace the Cm^{R} gene. The primers specified in the table were used in Gibson assembly to replace the Cm^{R} gene with the Km^{R} gene in the pcrRNA120 plasmid. Primers 1 and 2 were used to amplify the CRISPR plasmid excluding the Cm^{R} gene and primers 3 and 4 were used to amplify the Km^{R} gene in the pSW3 plasmid.

Cycle step	Temperature	Time	Number of Cycles
Initial Denaturation	98 ^{0}C	10 s	1
Denaturation	98 ^{0}C	1 s	
Annealing	60.6 ^{0}C (amplification of pcrRNA120 plasmid) 59.7 ^{0}C (amplification of pSW3 plasmid)	5 s	30
Extension	72^{0}C	1 min (pcrRNA120 plasmid amplification) 13 sec (amplification of pSW3 plasmid)	
Final extension	72 ^{0}C	60 s	1
Incubation	4 ^{0}C	Incubate	N/A

Table 7:Gibson assembly PCR Steps for replacing the Cm^R gene with the Km^R Gene. PCR conditions for the two amplifications are included in the table. The green highlighted text shows the annealing temperature and the extension time used to amplify the pcrRNA120 plasmid. The blue highlighted text shows the annealing temperature and the extension time used in the amplification of Km^R gene from the pSW3 plasmid.

3.3 Integration of the pSW3 plasmid into the bacterial chromosome

Initially, W3110, BW25113, W3110 *tel*$^+$ and BW25113 *tel*$^+$ strains were transformed using a low-copy number pINT helper plasmid via electroporation. Because the pINT helper plasmid is temperature-sensitive, the recombinant bacteria carrying the helper plasmid were grown at 30 ℃ and made electrocompetent. The electro-competent recombinant bacterial cells carrying the pINT helper plasmid were transformed with the pSW3 CRIM plasmid via electroporation. The transformed cells were grown in 1 mL of SOC without antibiotics at 37 ℃ for 1 h, and then at 42 ℃ for 30 min. The transformed cells were spread onto selective media and incubated at 37 ℃ overnight. Positive integrants were then selected for Km-resistance conferred by the CRIM

plasmid (carrying *vma*) and were further screened by colony PCR using four primers (P1–P4), as highlighted in Table 8. The integration of the pSW3 CRIM plasmid was confirmed using colony PCR. A single positive colony was chosen and resuspended in 20 µL of nuclease-free water. Then, 5 µL of the cell suspension was mixed with 0.5 U of Taq DNA polymerase (NEB, Ipswich, USA), 1 µL of primers P1–P4, 2.5 µL of 1X PCR buffer (NEB), and 2.5 mM of $MgCl_2$ with deoxynucleoside triphosphates (dNTPs), in a final volume of 25 µL. PCR amplification was performed for 25 cycles for 1 min each, starting with the denaturing step at 94 ℃, followed by the annealing step at 63 ℃, and then the final extension step for 1 min at 72 ℃.

Phage	Primer P1 Sequence	Primer P4 Sequence	Temp (^{0}C)	No integrant with P1 and P4 (expected band size bp)	Single integrant with P1 and P2, P3 and P4 (expected band size bp)	Multiple integrant with P1 and P2, P3 and P2, P3 and P4 (expected band size bp)
λ	GGCATCACGGCAATATAC	TCTGGTCTGGTAGCAATG	63	741	577, 666	577, 502, 666
HK022	GGAATCAATGCCTGAGTG	GGCATCAACAGCACATTC	59	740	289, 824	289, 373, 824
Φ80	CTGCTTGTGGTGGTGAAT	TAAGGCAAGACGATCAGG	63	546	409, 732	409, 595, 732
P21	ATCGCCTGTATGAACCTG	TAGAACTACCACCTGACC	57	506	568, 620	568, 682, 620
P22	AAGTGGATCGGCATTGGT	TTCGTATCGACAGGAAGG	63	735	376, 926	376, 568, 926
P22(EcoB)	AAGTGGATCGGCATTGGT	CGATTGAACCGCAGATTACG	63	609	376, 801	376, 568, 801

Table 8:Primers used in PCR test to confirm the chromosomal integration of the CRIM plasmid. Primers P2 (ACTTAACGGCTGACATGG) and P3 (ACGAGTATCGAGATGGCA) are the same for all *attP* sites. The four highlighted primers are used to perform colony PCR to confirm the integration of the pSW3 plasmid into the bacterial chromosome. Adapted from Haldimann & Wanner, 2001.

The integrative plasmid, pSW3, bearing the λ *attP* homologous region, integrates into the bacterial chromosome via site-specific recombination (Figure 22). To screen the positive integrants, four primers—P1, P2, P3, and P4–were used (Table 8). Single copy integrants showed

two bands when PCR was performed, owing to the loss of the λ *attB* sequence (P1-P4 fragment) where the pSW3 plasmid was inserted. As shown in Figure 22 A, two bands (577 bp and 666 bp in size), should be generated upon amplification of the BOP' region using the P1-P2 primer set, and the POB' region using the P3-P4 primer set. Multi-copy integrants were expected to show an additional gene fragment due to the P3-P2 amplification, as shown in Figure (22 B).

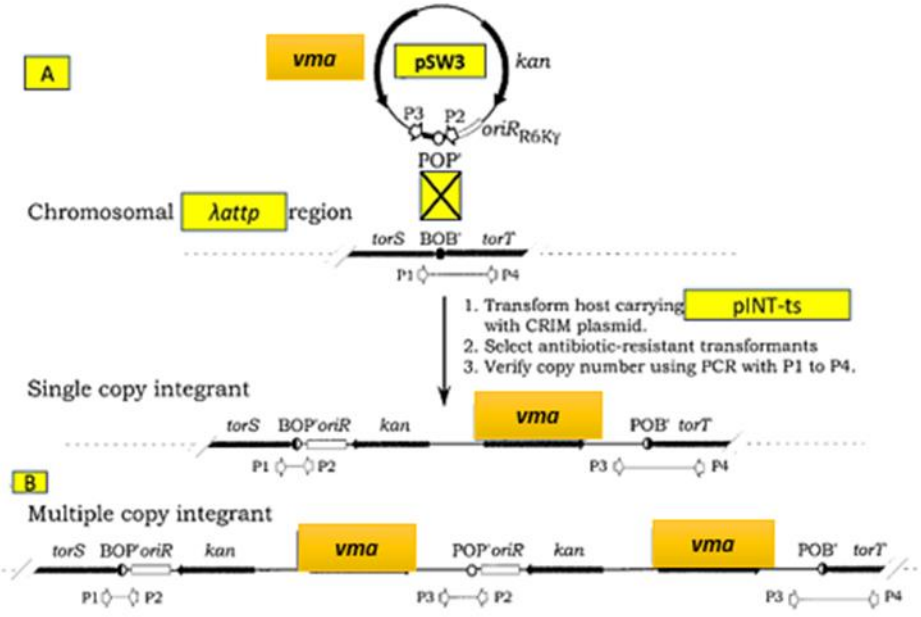


Figure 22:Schematic representation of the integration of pAH120-vma (pSW3) into the bacterial chromosome via site specific recombination. A. The λ *attP* region, also known as POP', has the "O" common core sequence flanked by different arms of λ *att*p sequence P and P'. After a single crossover event, generated by the integrase enzyme encoded by the pINT helper plasmid, the pAH120 plasmid was integrated into the bacterial chromosome via site specific recombination. P1, P2, P3 and P4 primers were used for PCR to identify positive integrants. **B.** In the event of multiple copy integration of the integrative plasmid, three bands are generated from PCR instead of two, as a single copy integration event. Adapted from Haldimann & Wanner, 2001.

3.4 Inductions

3.4.1 Regulated expression of the *vma* gene in recombinant bacteria

The P_{BAD} promoter is an inducible promoter. It is induced by the presence of L-arabinose which can be quickly turned off by the addition of glucose or its competitive inhibitor, fructose. Mg^{2+}, a cofactor, is essential for PI-SceI enzyme activity. A final concentration of 2.5 mM of Mg^{2+} is required *in vitro* to obtain the maximum enzyme activity, where a higher concentration has an inhibitory effect on endonuclease activity (Gimble et al., 2003;Gimbles et al., 1993).

Bacterial strains W3110, W3110 *vma*$^+$, BW25113 and BW25113 *vma*$^+$ were transformed by the msDNA precursor plasmid, pDMS, via electroporation. Colonies that grew on LB agar + Ap plates were cultured in 5 mL of LB broth supplemented with Ap and 1 % glucose and grown over night at 30 ℃. The next day, 10 mL of LB broth was inoculated with overnight cultures at a dilution of 1:100 and grown at 30 ℃ until the A_{600}=0.4. The cultures were centrifuged at 10 000 RPM for 10 min using the Beckman Coulter (Mississauga, CA) Avanti J-E Refrigerated Centrifuge and resuspended in 20 mL of LB broth supplemented with 1% L-arabinose and antibiotics. The L-arabinose induced cultures were grown at 37 ℃. The plasmids were extracted after 3, 6, and 16 h of incubation, and were run on a 0.8 % agarose gel at 90 V for 1 h and 30 min to visualize. Based on the observed results, ideal conditions such as, optimum temperature and the duration of incubation with the inducer, were determined for optimal endonuclease activity. Densitometry analysis was conducted using ImageJ software to determine the intensities of bands on the agarose gel.

A similar approach was followed to induce *vma* gene over expression in double integrants carrying the pDMS plasmid.

3.4.2 Induction of ministring DNA (msDNA) synthesis

Bacterial strains W3110, W3110 *tel*$^+$, BW25113 and BW25113 *tel*$^+$ were transformed by the msDNA generating precursor plasmid, pDMS/pNN9, via electroporation. W3110 *tel*$^+$ integrants were identified using blue/white screening on IPTG + X-gal plates. Cm-resistant (Cm^R) bacterial colonies, integrated with the *tel* gene expression cassette appeared white on IPTG + X-gal plates. Positive chromosomal integrants were transformed by the msDNA generating precursor plasmid via electroporation. Single-copy integrants carrying pDMS are Cm and Ap resistant. Considering the BW25113 strain does not have the complete *lacZ* gene, *tel* gene-integrated positive integrants were identified based on Cm resistance and were further confirmed via gene-specific PCR using *tel* gene-specific primers. The *tel* gene was amplified using forward (5' GGCGTTATGAAAATCCATTTTCGC 3') and reverse (5' TGTTTCATCGCATCGATAACAC 3') primers. Amplification was performed using the One Taq 2X PCR Master Mix (NEB, Ipswich, USA) for 120 s at 94 ℃ for initial denaturation, followed by 30 cycles each of 15 s at 94 ℃, 60 s at 48 ℃, 1 min 51 s at 68 ℃, and 5 min at 68 ℃ for the final extension, to generate a 1,840 bp fragment. The msDNA was generated following the protocol published by Wong et al., (2016), in pDMS/pNN9-bearing W3110, W3110 *tel*$^+$, BW25113 and BW25113 *tel*$^+$ bacterial strains.

3.4.3 Induction of ministring DNA generation and endonuclease expression in *tel*$^+$ *vma*$^+$ double integrants

The protocol previously described by Wong et al. (2016), was followed with slight modifications for the synthesis of msDNA in both BW25113 *tel*$^+$*vma*$^+$ and W3110 *tel*$^+$*vma*$^+$ double-integrated recombinant bacteria. An overnight culture of the double integrants was grown in LB media supplemented with 1 % glucose, Km, Cm and Ap antibiotics. The following day, the culture was subcultured at a ratio of 1:100 in 10 mL of LB broth supplemented with 1 % glucose and the

aforementioned antibiotics and grown at 30 ^{0}C until A_{600} = 0.8. The culture was centrifuged at 10,000 × g for 10 min and the cell pellet was resuspended in 20 mL of LB broth supplemented with 1 % glucose and antibiotics. The heat induction procedure was then followed to induce msDNA synthesis by incubating the culture at 42 ℃ on an orbital shaker for 1 h and then at 37 ℃ for 1.5 h. Finally, the cultures were kept at 30 ℃ for 2 h. Briefly following msDNA synthesis, 20 mL of culture was centrifuged at 10,000 × g for 10 min and was resuspended with LB containing 1 % L-arabinose (inducer) and 2.5 mM $MgCl_2$ before incubating at 37 ℃ for 16 h. The episomal DNA species were then isolated and separated by 0.8 % gel electrophoresis. Inductions were performed as indicated in Table 9.

Recombinant bacterial transformed with pDMS precursor plasmid	**Heat induction: *tel* expression**	**L-arabinose induction: endonuclease expression**	**Heat induction followed by L-arabinose induction: *tel* and endonuclease**
W3110	pDMS	pDMS	pDMS
W3110 *tel*$^+$	pDMS+LCC bacterial backbone + msDNA	pDMS	pDMS+LCC bacterial backbone +msDNA
W3110 *tel*$^+$*vma*$^+$	pDMS+LCC bacterial backbone +msDNA	*Smear (digested pDMS)	**msDNA
W3110 *vma*$^+$	pDMS	*Smear (digested pDMS)	*Smear (digested pDMS)
BW25113	pDMS	pDMS	pDMS
BW25113 *tel*$^+$	pDMS+LCC bacterial backbone +msDNA	pDMS	pDMS+LCC bacterial backbone +msDNA
BW25113 *tel*$^+$*vma*$^+$	pDMS+LCC bacterial backbone +msDNA	*Smear (digested pDMS)	**msDNA
BW25113 *vma*$^+$	pDMS	*Smear (digested pDMS)	*Smear (digested pDMS)

Table 9:Summary of the expected test results after msDNA synthesis and PI-SceI enzyme over expression in different bacterial strains.* The PI-SceI homing endonuclease enzyme is over expressed by inducing the *vma* (PI-SceI encoding) gene which is under the control of the L-

arabinose promoter. This PI-SceI enzyme digests its corresponding recognition sequence bearing pDMS plasmid. The exonucleases present inside the bacteria further digest the partially digested plasmid. As a result, the plasmid extraction when run on an agarose gel will appear as a smear. ** After msDNA synthesis, PI-SceI enzyme over expression is induced by adding L-arabinose to the media. As a result, the unprocessed pDMS precursor plasmids and unwanted LCC bacterial plasmids after msDNA generation, are digested by the PI-SceI homing endonuclease enzyme. The partially digested plasmids are further digested by exonucleases present in the bacteria, leaving only the msDNA.

3.4.4 Induction of msDNA synthesis in JM109 *tel*$^+$*vma*$^+$ double integrants

The msDNA was synthesized in JM109 *tel*$^+$*vma*$^+$ recombinant bacteria carrying pDMS/pNN9 plasmids, as described in section 3.4.3. To further confirm msDNA synthesis, both induced and uninduced plasmid extracts were digested with the NEB EcoR1 restriction enzyme. The 25 µL reaction mixture consisted of 500 ng of plasmid DNA, 2.5 µL of NEB 10X Cutsmart Buffer, 0.5 µL of NEB High Fidelity (HF) EcoR1 restriction enzyme, and nuclease-free water. The reaction mixture was incubated at 37 ℃ for 15 min and run on a 0.8 % agarose gel for 1 h and 30 min at 90 V.

3.4.5 Episomal expression of the Tel protelomerase enzyme

Instead of integrating the Tel protelomerase enzyme-encoding gene cassette into the bacterial chromosome, Tel protelomerase enzyme expression was achieved episomally. To improve msDNA synthesis, a low copy number plasmid carrying the Tel protelomerase enzyme encoding gene cassette was introduced into recombinant bacterial cells along with the msDNA synthesizing precursor plasmid. The episomal expression of the Tel protelomerase enzyme acts on the precursor plasmid and synthesizes msDNA upon induction.

3.4.5.1 Induction of msDNA synthesis in pACYC-CI-857-tel (pACYC-tel) carrying W3110/JM109 recombinant bacteria in the presence of pDMS/pNN9 precursor plasmids

W3110 and JM109 bacterial cells were made electrocompetent by following the protocol outlined in Gonzales et al., (2013). Electrocompetent W3110 and JM109 bacteria were transformed with pACYC-tel and pDMS/pNN9 msDNA synthesizing precursor plasmids via electroporation. At first, electrocompetent W3110 and JM109 bacteria were transformed with pACYC-tel recombinant plasmid via electroporation. Chloramphenicol-resistant, W3110/JM109 bacterial colonies carrying the pACYC-tel plasmid were grown overnight at 30 ℃ and made electrocompetent. Electrocompetent recombinant bacteria carrying pACYC-tel were transformed with pNN9/pDMS msDNA synthesizing precursor plasmids via electroporation. Positive colonies, both Cm and Ap resistant, were grown overnight at 30 ℃. The next day, the overnight culture was subcultured at a 1:100 dilution in 10 mL LB broth and grown until A_{600}=0.8 was reached. Heat induction was performed to induce msDNA synthesis, as outlined in Wong et al., (2016). Plasmids were isolated after induction and run on a 0.8 % agarose gel to visualize. The presence of both plasmids and msDNA was further confirmed by digesting the plasmid extract with BamH1, obtained from NEB (Ipswich, USA).

3.4.5.2 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-T7 and pDMS/pNN9 precursor plasmids

In the pACYC-tel-T7 recombinant plasmid, the *tel* gene was placed under the regulation of the T7 inducible promoter. The *tel* gene expression gene cassette was cloned into a low copy number pACYC-Duet plasmid (*ori*: p15). This pACYC-tel-T7 recombinant plasmid construct was a gift from Mediphage Bioceuticals, Toronto, CA. BL21(DE3) and W3110 cells were transformed with the pACYC-tel-T7 and pNN9/pDMS msDNA synthesizing precursor plasmids via sequential transformation, as outlined in section 3.4.5.1. Both Cm and Ap resistant positive colonies were

picked and grown at 37 ℃ overnight. On the next day, the overnight culture was subcultured at a 1:100 in LB broth supplemented with Cm and Ap. When the A600=0.8, isopropyl ß-D-1-thiogalactopyranoside (IPTG) was added at a final concentration of 0.4 mM or 1 mM and induced for 6 h. After induction, the plasmids were isolated and run on a 0.8 % agarose gel to visualize.

3.4.5.3 Induction of msDNA synthesis in W3110 bacteria carrying p-pBAD-tel and pDMS/pNN9 plasmids

In this study, two plasmids were employed to express Tel protelomerase episomally, where the *tel* gene was placed under the control of the P_{BAD} promoter. One plasmid contains the *ori* p15 (pACYC-tel-pBAD), and the other contains pSC101. W3110 recombinant bacteria were sequentially transformed with pACYC-tel-pBAD/ pXG-pBAD-tel and pNN9/pDMS precursor plasmids as outlined in section 3.4.5.1. Both Cm^R and Ap^R positive colonies were selected and grown overnight at 30 ℃ in LB broth supplemented with Ap, Cm and 1 % glucose. The next day, the cells were subcultured at a 1:100 dilution in 10 mL of LB media supplemented with antibiotics and glucose. When the A_{600}=0.8, the culture was centrifuged at 10,000 ×g for 10 min. The supernatant was discarded, and the cell pellet was resuspended in 10 mL of LB broth supplemented with Cm. L-arabinose was added at a final concentration of 1 % and induced for 6 h at 37 ℃. After 6 h of induction, cells were collected through centrifugation and plasmids were extracted. The extracted plasmids were run on a 0.8 % agarose gel at 90 V for 1 h and 30 min to visualize. The plasmid map of the pXG-pBAD-tel plasmid is shown in Figure 23.

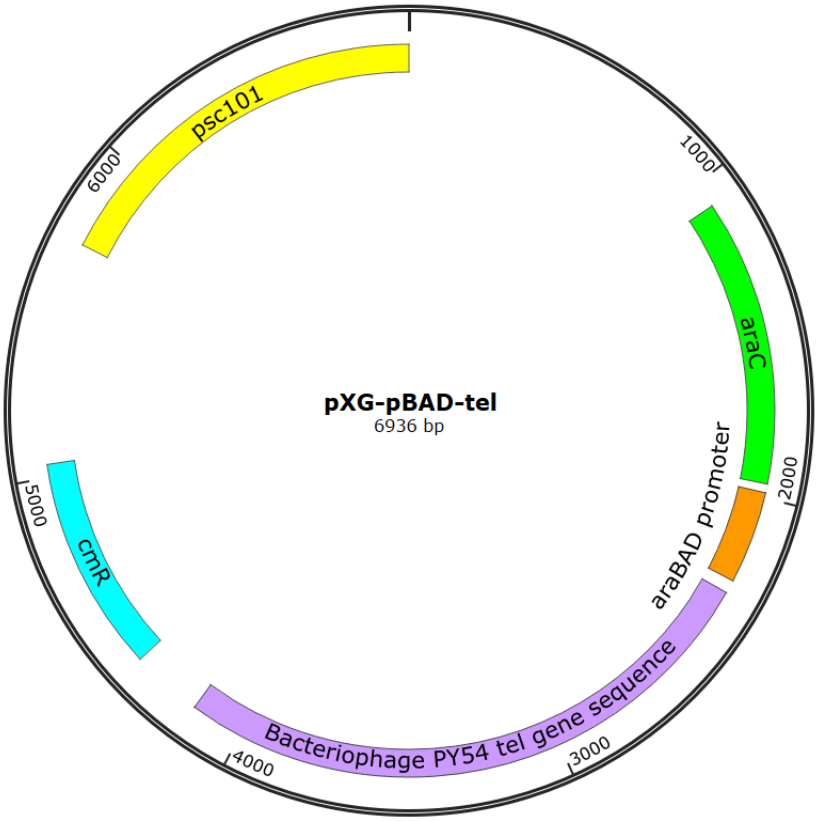


Figure 23:Plasmid map of pXG-pBAD-tel plasmid. The *tel* gene is placed under the control of P_{BAD} inducible promoter. The plasmid is 6.9 kb in size and has pSC101 *ori.* This is a gift from Mediphage Bioceuticals, Toronto.

3.4.5.4 Induction of msDNA synthesis in W3110 vma$^+$ bacteria carrying pXG-pBAD-tel and pDMS/pNN9 plasmids

The presence of the integrated *vma* gene in W3110 *vma*$^+$ recombinant bacteria was confirmed by performing a colony PCR, as outlined in section 3.3. W3110 *vma*$^+$ recombinant bacteria were made electrocompetent and sequential transformation of the pXG-pBAD-tel plasmid and pNN9/pDMS precursor plasmids was performed, as described in section 3.4.5.1. The msDNA synthesis was induced in recombinant bacteria carrying the Tel plasmids and the precursor plasmids as described in section 3.4.5.3.

3.4.6 Induction of crRNA over expression

The pre-crRNA encoding gene cassette is placed under the control of the P_{BAD} inducible promoter. W3110 recombinant bacteria were sequentially transformed with CRISPR plasmids and the pDMS precursor plasmid via electroporation. The pre-crRNA was expressed as described in section 3.4.1. The induction was performed for 6 h at 30 °C, and the plasmids were extracted using the Monarch Plasmid Miniprep Kit from NEB (Ipswich, USA). The concentration and the purity of extracted plasmids were measured using the Thermo Fisher Scientific (Waltham, USA) NanoDrop 2000. The extracted plasmids (500 ng) were digested with the HF EcoRV restriction enzyme obtained from NEB (Ipswich, USA). The 500 ng of undigested extracted plasmid DNA and EcoRV-digested plasmids were run on a 0.8 % agarose gel at 90 V for 1 h and 30 min to visualize.

3.4.6.1 Induction of msDNA synthesis in W3110 *tel*$^+$ recombinant bacteria carrying the CRISPR plasmid

The msDNA was synthesized in W3110 tel$^+$ recombinant bacteria carrying the CRISPR plasmid and pNN9 precursor plasmid, as outlined in section 3.4.3. After msDNA synthesis, the plasmids were isolated using the NEB (Ipswich, USA) Monarch Plasmid Miniprep Kit, and the concentrations of the extracted plasmid DNA were determined using the Thermo Fisher Scientific (Waltham, USA) NanoDrop 2000. The 500 ng of extracted plasmid DNA were run on a 0.8 % agarose gel to visualize the data.

3.4.6.2 Induction of msDNA synthesis followed by crRNA overexpression

msDNA synthesis was induced in W3110 *tel*$^+$ recombinant bacteria carrying pNN9 and pcrRNA120 plasmids, as described in section 3.4.3. After msDNA synthesis, the cultures were centrifuged at 10,000 ×g for 10 min and resuspended in LB broth supplemented with antibiotics

and L-arabinose at a final concentration of 1 %. The induction was performed at 30 ℃ for 4 h or overnight (approximately 16 h). The plasmids were extracted and run on a 0.8 % agarose gel to visualize (300 ng of DNA was loaded to each lane).

3.5 Transformation

The bacterial cells were transformed with recombinant plasmids via electroporation. First, the cells were made electrocompetent by following the protocol described by Calvin & Hanawalt, 1988. The 50 µl of electrocompetent cells were mixed with approximately 1000 ng of the recombinant plasmid of interest and incubated on ice. Subsequently, 100 µl of the cell suspension was added to an electroporation cuvette, and a voltage of 1000 V was applied. Pre-warmed 1000 µl of LB broth without antibiotics was added, and the mixture was incubated on a shaker at 37 / 30 ℃ for 1 h. Following incubation, 100 µl of the cell suspension was streaked onto selective media containing antibiotics and incubated overnight at the desired temperature.

3.6 *In vitro* digestion of the pDMS plasmid

In order to confirm the presence of P1-Sce1 enzyme recognition sequences in the pDMS precursor plasmid, pDMS was subjected to digestion with the commercially available NEB (NEB, Ipswich, USA) P1-Sce1 homing endonuclease enzyme. In each reaction, 500 ng of plasmid DNA was digested with 0.5 U of P1-Sce1 enzyme in a total reaction volume of 25 µL. The 25 µL reaction mixture consisted of 500 ng of DNA, 2.5 µL of 10X NEB Buffer, 0.2 µL rAlbumin (20 mg/mL), 0.5 µL of the P1-Sce1 enzyme and sterile water. The mixture was incubated at 37 ℃ for 1 h and run on a 0.8 % agarose gel along with undigested pDMS plasmid as a negative control to visualize.

3.7 Integration of the *vma* gene into JM109 *tel*$^{+}$ recombinant bacteria using P1 transduction

3.7.1 P1 transduction

P1 bacteriophage is a temperate phage belonging to the *Myoviridae* family and is composed of an icosahedral head and a contractile tail. This bacteriophage transfers gene fragments from one *E. coli* bacterial strain to another via a process called 'transduction'(Ikeda & Tomizawa, 1965;Saragliadis et al., 2018). Although the P1 phage is a temperate bacteriophage, it does not integrate its double-stranded DNA genome into the host chromosome during transduction. Instead, its genome circularizes and forms a plasmid-like structure and replicates. The P1 bacteriophage can accommodate ~110 kb of DNA in its capsid. During P1 infection, virus-like particles inside the bacterial host cells are formed which contain nonphage DNA. These virus-like particles contain DNA fragments of plasmids and chromosomal DNA of the bacterial host cell. The P1 phage can aberrantly package fragments anywhere along the bacterial chromosome and inject into recipient *E. coli* host cells and facilitate integration of the non-phage DNA into the bacterial chromosome via homologous recombination. If there is selection for a gene of interest, any region of the host's chromosomal DNA can be transferred from one bacterial strain to another through a process referred to as "generalized transduction". P1 transduction is carried out by first infecting donor host cells with P1 phage, resulting in the generation of transducing particles. After sterilization, lysate is used to infect recipient *E. coli* cells where non-phage DNA is injected into the host cells and integrated into the bacterial chromosome via homologous recombination. The positive recombinants were selected by growing them in selectable media. For transduction, P1 vir phage was used, which is a lytic mutant of the P1 phage and enters the lytic life cycle upon infection (Saragliadis et al., 2018).

P1 transduction was performed to transfer the PI-Sce1 homing endonuclease enzyme-expressing *vma* gene from W3110 *vma*$^{+}$ to JM109 *tel*$^{+}$ (Cm resistant) recombinant bacteria. The Km resistance gene linked with the *vma* expression cassette was used as a selectable marker to isolate positive recombinants.

3.7.2 Regeneration of the P1rev6

The variant of P1*vir* used for the transduction experiment is called "P1rev6". To prepare fresh P1rev6 stock solution, BB4 *E. coli* cells were grown overnight at 37 ℃ in 5 mL of LB broth. The overnight culture was subcultured at a 1:100 ratio and grown at 37 ℃ until A600=0.4. The 300 µL of mid log phase culture was then gently mixed with 100 µL of P1rev6 old stock suspension. Subsequently, 3 mL of top agar (0.7 % agar in LB) was added to the cell suspension (Bacto Tryptone and Bacto Agar from Difco Laboratories, Sparks, MD). The mixture was poured into prewarmed LB + agar plates and incubated overnight at 37 ℃. The next day, 10 mL of TN buffer (0.01 M Tris-HCl and 0.1 M NaCl, pH 7.8; Fisher Scientific, USA) was added to the plate and kept at 4 ℃ for 24 h. The soft top agar layer was then scrapped off into prechilled falcon tubes. Chloroform (50 µL) was added into the mixture and incubated for 5 min. The mixture was centrifuged at 10,000 RPM for 20 min using the Beckman Coulter (Beckman Coulter Life Sciences, Mississauga, CA) Avanti J-E Centrifuge. The supernatant was filtered through a 0.45 µm filter.

The phage titer was determined by performing the spot plate technique. A dilution series of the P1rev6 phage stock was prepared, as shown in Figure 24. The 300 µL of cells was taken from a BB4 overnight culture and mixed with 3 mL of top agar. The cell suspension was poured onto a petri plate and incubated at 30 °C for 30 min. From the P1 rev6 dilution series (1 -10^{-8}), 10 µL was added onto the bacterial lawn and incubated at 37 ℃ overnight.

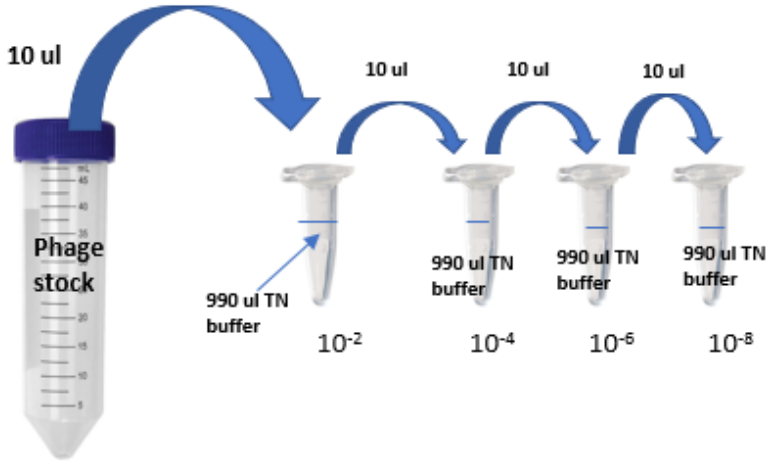
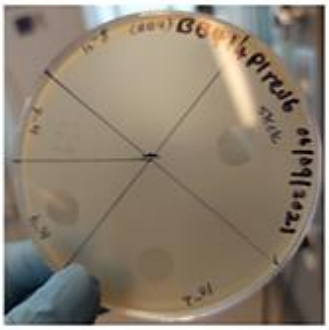


Figure 24:Preparation of phage dilution series. A culture dilution series of P1rev6 was prepared by adding 10 µL from the P1rev6 stock solution into 990 µL TN buffer, which gave a 10^{-2} dilution. 10 µL from the 10^{-2} was added into 990 µL of TN buffer, making a 10^{-4} dilution. This procedure was continued for the rest of the dilution series. A culture dilution series was prepared from undiluted to 10^{-8}.

3.7.3 Spot plate technique to determine the phage titer

P1 rev6 phage stock solution was regenerated, as outlined in section 3.7.2, in BB4 *E. coli* cells. The spot plate technique was used to determine the phage titer in Plaque Forming Unit/mL (PFU/mL) of the P1 rev6 stock. At a 10^{-6} dilution, three plaques were observed, resulting in a phage titer of 3×10^{9} PFU/mL of P1rev6 stock (Figure 25).



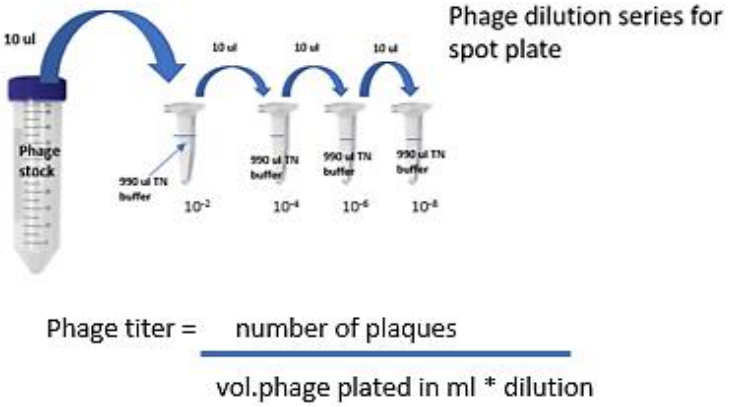


Figure 25:Spot Plate Technique to determine the phage titer. A. In order to determine the phage titer, the lawn of BB4 *E. coli* cells were prepared by mixing the 300 µL of mid log phase culture with 3 mL of top agar and poured into agar plates and incubated at 37 ^{0}C for 30 min. The 10 µL from the dilution series 1-10^{-8} were spotted onto the BB4 bacterial lawn and incubated overnight at 37 ℃. The number of plaques were counted to determine the phage titer. **B**. Phage Titer in PFU/mL were calculated by adding data into the equation.

3.7.4 Preparation of P1 *vma*$^{+}$ transducing lysate

"Generalized transduction," carried out by the P1 phage, is used to transfer the *vma* gene expression cassette from W3110 *vma*$^{+}$ recombinant bacteria to JM109 *tel*$^{+}$ recombinant bacteria. The P1 *vma*$^{+}$ phage lysate (this is referred to as *vma* lysate) was prepared by infecting W3110 *vma*$^{+}$ recombinant donor bacteria with 100 µL of the P1rev6 stock lysate (3x10^{9} PFU/mL). This resulted in a multiplicity of infection (MOI) at around 1 phage for each bacterium. Phage titer of the *vma* lysate was determined using the spot plate technique on a lawn of BB4 cells. The phage titer of the *vma* lysate was 5x10^{9} PFU/mL.

3.7.5 Generation of JM109 *tel*$^+$*vma*$^+$ double integrants

The recipient recombinant bacteria JM190 *tel*$^+$ was grown overnight at 30 °C in 1 mL of LB. The overnight culture was centrifuged at 3000 ×g for 3 min using Benchmark Scientific (Brea, USA) compact centrifuge. The cell pellet was resuspended in 500 µL of P1-LB (1 mL of LB supplemented with 5 µL of $MgCl_2$ (~12.5 mM), 5 µL of $CaCl_2$ (~5 mM), and 5 µL of 20 % glucose (~0.1 %)). To conduct the transduction step, 100 µL of the *vma* lysate was transferred into a clean microcentrifuge tube and incubated at 37 °C for 30 min with the cap opened to evaporate the added chloroform. The 100 µL of the cell suspension was added to the 100 µL of *vma* lysate and mixed well. This cell suspension was kept at a 37 °C shaking incubator for 30 min. Subsequently, 200 µL of 1 M sodium citrate (pH 5.5) was added. On top of this, 1 mL of fresh LB broth was added and incubated at 30 °C for 2 h. This was then centrifuged at 5,000 ×g for 5 min. The supernatant was discarded, and the pellet was resuspended in 100 µL of LB-Citrate (LB broth supplemented with 100 mM of Na-Citrate, pH 5.5). The entire content was poured into both Km and Cm supplemented LB agar plates and incubated overnight at 30 °C. The concentrations of Cm (6 µg/mL) and Km (10 µg/mL) added are as indicated in section 3.1.1. The positive transduced colonies were re-streaked on 5 mM citrate selective media and incubated at 30 °C overnight. Colony PCR was performed with positive transduced colonies to confirm the presence of the *vma* gene. The *vma* gene was amplified using forward (5’ GATGGGCCCTTAGCAATTATGG 3’) and reverse (5’ AAAAGAAAGCCCCCGACG 3’) primers. The PCR reaction contained 12.5 µL of NEB (Ipswich, USA) 2X Taq PCR master mix, 10 ng of plasmid DNA, 0.2 µM of forward and reverse primers, and sterile water within a final volume of 25 µL. PCR amplification was performed for 30 cycles, starting with the denaturing step for 30 s at 94 °C, the annealing step for 1 min at 52 °C, and the final extension for 1 min at 68 °C.

3.8 Designing of the CRISPR Cas 3 pre-crRNA encoding gene sequence

The pre-crRNA encoding gene sequence targeting the unwanted bacterial backbone of the msDNA synthesizing precursor plasmid, pNN9/pDMS, was designed based on the journal article published by Gomaa et al., (2014). The 90 bp pre-crRNA encoding gene sequence consists of a 29 bp repeat - 32 bp spacer - 29 bp repeat sequence (Figure 26; this sequence will be referred to as pre-crRNA). The spacer sequence in the pre-crRNA is complementary with the protospacer sequence within the target sequence of the plasmid. The repeat sequence in the pre-crRNA is a unique sequence identified by the bacterial CRISPR cas3 system. The 32 bp protospacer sequence is located downstream of the AAG protospacer adjacent motif (PAM) sequence of the bacterial backbone portion of the precursor plasmid. Two protospacer sequences were designed. The first target sequence is located in the Ap^R gene of the plasmid, and the second target sequence is located in the *ori* of the plasmid (Figure 26; A and B respectively).

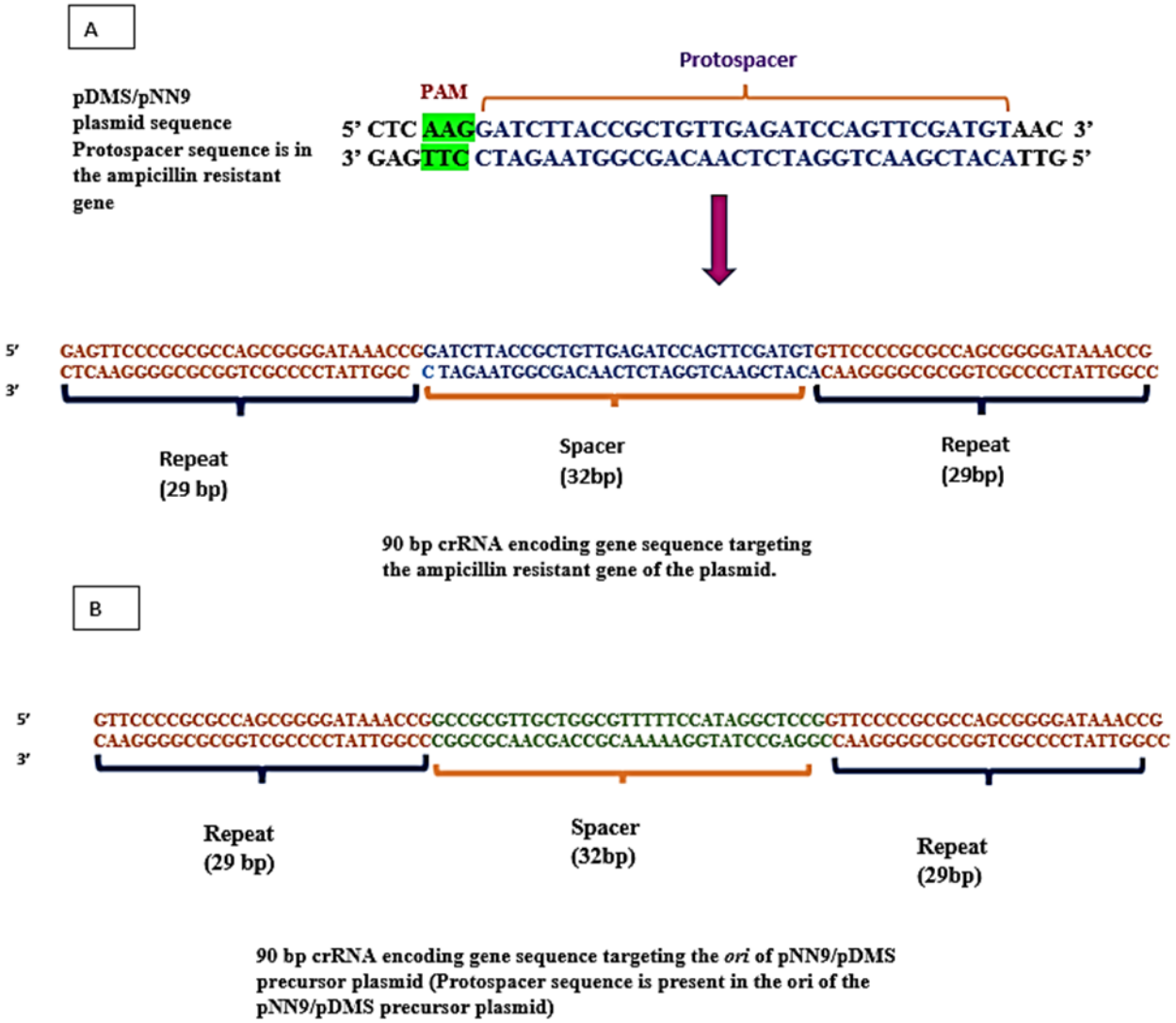


Figure 26:Designing of pre-crRNA encoding gene sequences. The 90 bp pre-crRNA sequence consists of repeat (29 bp)-spacer (32 bp)-repeat (29 bp) sequences. The 32 bp protospacer sequence is located downstream of the AAG PAM sequence of pDMS/pNN9 precursor plasmids. **A.** Target 1-The protospacer sequence is in the Ap resistant gene of pDMS/pNN9 precursor plasmids. **B.** Target 2-The protospacer sequence is present in the *ori* of pNN9/pDMS precursor plasmids. Adapted from Gomaa et al., 2014.

4. Results

4.1 Integration of the PI-SceI homing endonuclease encoding gene (*vma*) into the bacterial chromosome

The *vma* gene was cloned into the pAH120 integrating plasmid and placed under the control of the L-arabinose-inducible P_{BAD} promoter (this construct is referred to as pSW3; this work was done by Dr. Shirley Wong in 2016). The pSW3 CRIM plasmid was successfully integrated into the chromosome of the following bacterial strains: W3110, W3110 *tel*$^+$, BW25113 and BW25113 *tel*$^+$. Both W3110 and BW25113 *E. coli* K-12 bacteria used in this study are *pir*$^-$. Since the CRIM plasmid bears the R6K replicon, (which requires a *trans*-acting π enzyme for replication, encoded by *pir*$^+$ bacterial cells) the plasmid cannot propagate in the *pir*$^-$ target cell. It exists as a suicide plasmid in a *pir*$^-$ cell that can only be replicated if it recombines into the chromosome of the cell.

The pSW3 integrative plasmid was integrated into the bacterial chromosome as a single copy. CRIM plasmids have a conditional-replication origin and a phage-attachment (*attP*) site, where site-specific recombination takes place between the phage-attachment site and the corresponding *attB* site present in the bacterial chromosome. This results in recombinant bacterial strains that carry the *vma* endonuclease gene with stability. Initially, W3110, BW25113, W3110 *tel*$^+$ and BW25113 *tel*$^+$ strains were transformed using a low-copy number pINT helper plasmid via electroporation as described in section 3.3. The expression of the integrase gene, *int*, within the pINT plasmid, is controlled by a temperature sensitive λ repressor, ts, whereby the induction of the *int* is accomplished by shifting the temperature from 30 °C to 42 °C. The integrase (Int) enzyme encoded by the pINT helper plasmid, facilitated the integration of the CRIM plasmid into the bacterial chromosome in *pir*$^-$ *E. coli*. Since the pINT helper plasmid is temperature-sensitive, the positive integrants (*vma* integrated recombinant bacteria) are cured of the helper plasmid.

Upon successful integration of the pSW3 plasmid, positive integrants were selected based on antibiotic resistance and further screened for positive integrants by colony PCR. As indicated in Table 8 and described in section 3.3, colony PCR of the positive single copy integrants should yield two bands corresponding to 577 bp and 666 bp when run on an agarose gel. Figure 27 (lanes;1-4) shows the colony PCR results on a 1 % agarose gel, indicating the positive integrant bearing the *vma* gene. The presence of the *vma* gene in positive integrants was confirmed by performing a colony PCR using *vma* gene-specific primers. Obtaining fragment sizes of 1832 bp upon amplification confirmed the presence of the *vma* gene in positive integrants (Figure 27; lanes 5-7).

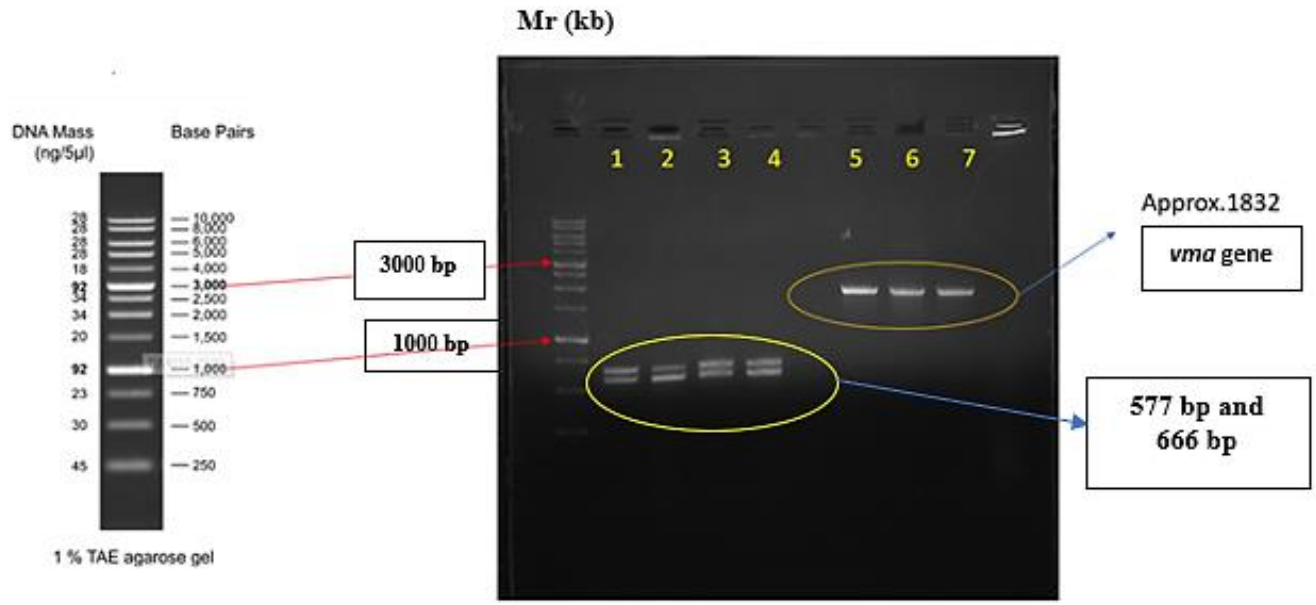


Figure 27:Colony PCR test results on a 1% agarose gel confirmed the integration of pSW3 and the presence of the integrated *vma* gene in the bacterial chromosome. 1. W3110 *vma*$^{+}$; **2.** W3110 *tel*$^{+}$*vma*$^{+}$; **3.** BW25113 *tel*$^{+}$*vma*$^{+}$; **4.** BW25113 *vma*$^{+}$; **5.** Positive control of *vma* gene in pSW3 plasmid; **6.** BW25113 *tel*$^{+}$*vma*$^{+}$; **7.** W3110 *tel*$^{+}$*vma*$^{+}$. Mr: 1 kb ladder (FroggaBio).

4.2 Expression of the *vma* homing endonuclease gene upon integration into the *E. coli* genome

To determine whether the introduced homing endonuclease has any adverse effects on the synthesis of msDNA, the msDNA-synthesizing precursor plasmid pDMS was introduced into the *vma*-integrated single and double integrants (both *tel* and *vma* integrated). This pDMS msDNA-synthesizing precursor plasmid contains the PI-Sce1 homing endonuclease recognition sequence cloned into it. Initially, the digestion of the pDMS plasmid upon induction of *vma* gene overexpression was tested in both single and double integrants. Then, msDNA synthesis was induced in double integrants to assess the Tel protelomerase enzyme activity in *vma*-integrated double integrants.

4.2.1 Transformation of the pDMS precursor plasmid into the positive integrant

To confirm the presence of PI-SceI homing endonuclease recognition sequences in the pDMS msDNA synthesizing precursor plasmid, we performed in vitro digestion using commercially available PI-SceI endonuclease (New England Bio Labs) enzyme. The locations of the PI-SceI recognition sites are shown in Figure 9 A. The *in vitro* digestion of the pDMS plasmid with PI-SceI endonuclease should result in two bands corresponding to 4.3 kb and 1.9 kb in size on an agarose gel. As shown in Figure 28, these digestion products were observed on the gel, confirming the presence of the PI-SceI recognition sequences in the pDMS plasmid. After confirming that pDMS carries the correct PI-SceI recognition sequences, the recombinant bacterial strains including W3110 tel^+, W3110 tel^+ vma^+, W3110 vma^+, BW25113 tel^+, BW25113 tel^+ vma^+, and BW25113 vma^+ were transformed with the pDMS plasmid.

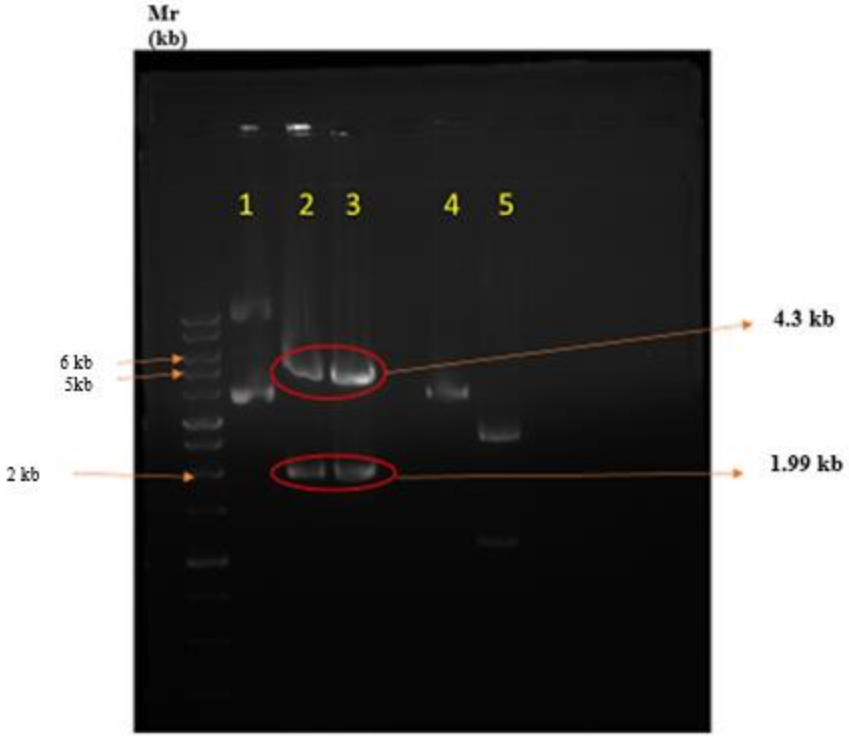


Figure 28:pDMS precursor plasmid digested with the PI-SceI endonuclease enzyme. 1. Un-digested pDMS plasmid; **2.** Digested pDMS plasmid; **3.** Digested pDMS plasmid; **4.** Un-digested control plasmid provided by the manufacturer; **5.** Digested control plasmid. Mr: 1 kb ladder (FroggaBio).

4.2.2 Induction of the *vma* overexpression in recombinant bacteria

The recombinant strains (W3110 *vma*$^{+}$ *and* BW25113 *vma*$^{+}$) were transformed by the pDMS precursor plasmid and *vma* gene over expression was induced as outlined in section 3.4.1. The plasmids were extracted after 3, 6, and 16 hours of growth at 37 °C, and equal volumes were loaded onto gels. Compared to the control (Figure 29; lanes 7-10; un-induced), the induced samples showed lower band intensities (Figure 29; lanes 1-6). The band intensities also decreased with induction time. Based on densitometry analysis, after 3 h of induction inside W3110, there was a 26 % decrease in band intensity, and also after 16 h of induction, there was a 29 % decrease in band intensity. In BW25113, based on densitometry analysis, after 3 h of induction, the band

intensity decreased by 49 %, and after overnight induction, the band intensity decreased by 53 %. Equal volumes of the plasmid extracts were loaded and run on a 0.8 % agarose gel. Based on these results, it is evident that optimal endonuclease activity can be obtained at 37 ^{0}C with an induction period of 16 h. Even though the data is not shown, the following negative controls were run along the experiment. The W3110 *vma*$^{+}$ and BW25113 *vma*$^{+}$ recombinant bacteria carrying the pNN9 precursor plasmid acted as negative controls and didn't show digestion of the plasmids upon *vma* gene overexpression, as pNN9 precursor plasmid doesn't have the PI-Sce1 recognition sequence cloned into it. Another negative control was the W3110 and BW2511 bacteria carrying the pDMS plasmid. Since there is no *vma* gene integrated, there was no observed digestion of the plasmid. Compared to the W3110 *vma*$^{+}$ strain, BW25113 *vma*$^{+}$ recombinant bacteria exhibited a higher digestion of the pDMS precursor plasmid. This is likely due to the BW25113 bacterial cells not being able to utilize the added L-arabinose within the media owing to the deletion mutation in the *araB-araD* region of the *araBAD* Δ*(araD-araB567*) gene in their chromosome (Grenier et al., 2014). As such, a higher concentration of L-arabinose is left within the media for efficient and consistent induction of the P_{BAD} promoter, compared to W3110.

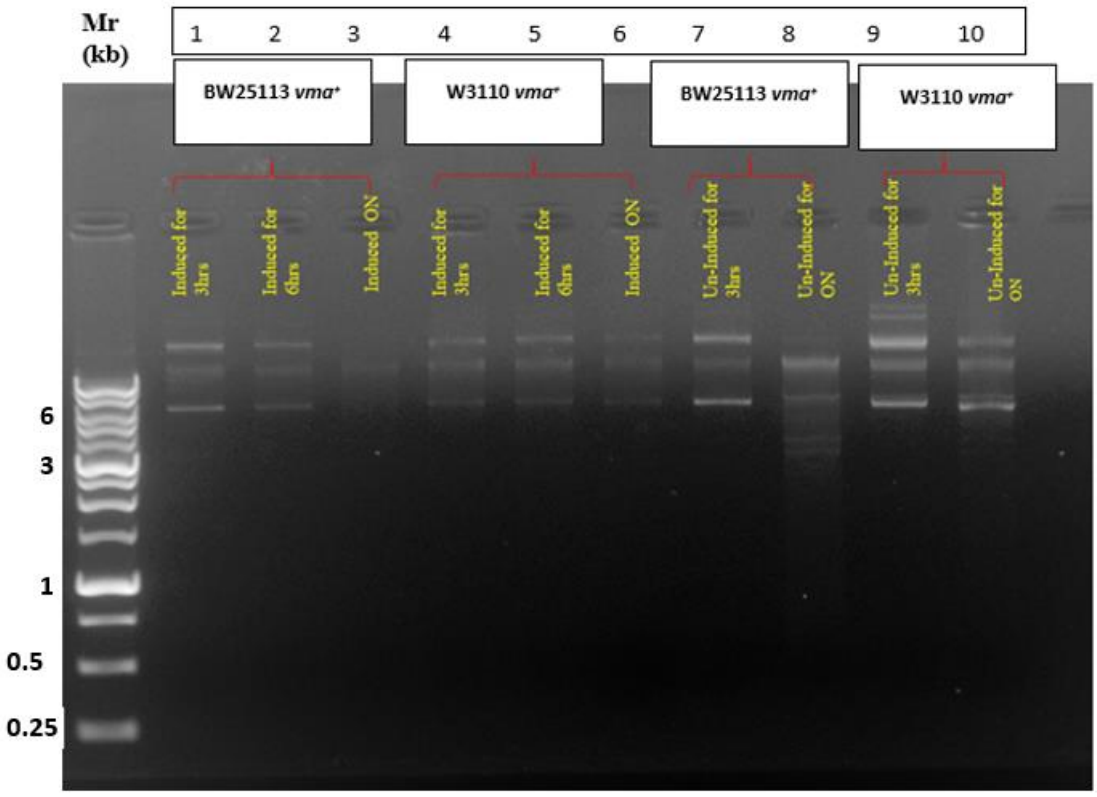


Figure 29:Extractions of plasmids after induction of the PI-SceI enzyme in *vma*-integrated W3110 and BW25113 recombinant bacteria. overexpression of the PI-SceI enzyme was induced in W3110 and BW25113 single integrants carrying the pDMS precursor plasmid. Overnight refers to 16 h of incubation. Lanes:**1**. BW25113 *vma*$^{+}$+pDMS (induced for 3 h);**2.** BW25113 *vma*$^{+}$+pDMS (induced for 6 h);**3.** BW25113 *vma*$^{+}$+pDMS (induced for 16 h);**4.** W3110 *vma*$^{+}$ + pDMS (induced for 3 h);**5.** W3110 *vma*$^{+}$+pDMS (induced for 6 h);**6.**W3110 *vma*$^{+}$+pDMS (induced for 16 h);**7.**BW25113 *vma*$^{+}$+pDMS (Un-induced for 3 h);**8.**BW25113 *vma*$^{+}$+pDMS (Un-induced for 16 h);**9.**W3110 *vma*$^{+}$+pDMS (Un-induced for 3 h);**10.**W3110 *vma*$^{+}$+pDMS (Un-induced for 16 h). Mr: 1 kb ladder (FroggaBio).

4.2.3 Induction of the *vma* gene over expression in both *tel* and *vma* integrated recombinant bacteria

Both *vma* and *tel* integrated BW25113 and W3110 double integrants were transformed with pDMS precursor plasmids and induced the *vma* gene overexpression as described in section 3.4.1.

As indicated in Figure 30 (lanes 1,3,5, and 7), upon induction of endonuclease enzyme activity, the precursor plasmid concentration decreased compared to the un-induced cultures (Figure 30;

lanes 2, 4 ,6, and 8). According to densitometric analysis, after 16 h of induction, the band intensity in W3110 double integrants has decreased by 46 %, and the band intensities in BW25113 double integrants after 16 h of induction have decreased by 76 %. This confirms that the activity of the homing endonuclease enzyme has digested the precursor plasmid pDMS, which was present in the double-integrated recombinant bacteria.

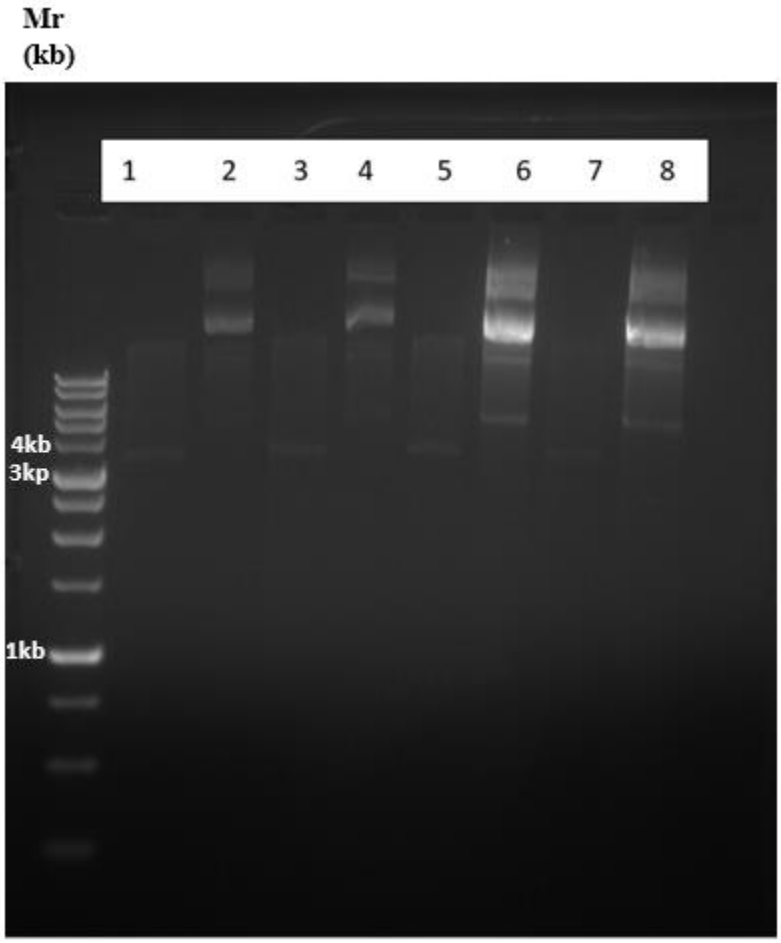


Figure 30:The induction of PI-SceI enzyme expression in double integrants carrying the pDMS plasmid.1. W3110 *tel*$^{+}$*vma*$^{+}$(Induced); **2.** W3110 *tel*$^{+}$*vma*$^{+}$(Un-Induced); **3.**W3110 *tel*$^{+}$*vma*$^{+}$(Induced); **4.** W3110 *tel*$^{+}$*vma*$^{+}$(Un-Induced); **5.** BW25113 *tel*$^{+}$*vma*$^{+}$(Induced); **6.** BW25113 *tel*$^{+}$*vma*$^{+}$ (Un-Induced);**7.** BW25113 *tel*$^{+}$*vma*$^{+}$(Induced)**;8.** BW25113 *tel*$^{+}$ *vma*$^{+}$(Un-Induced). Mr: 1 kb ladder (FroggaBio).

4.2.4 Ministring DNA synthesis in *vma* and *tel* integrated double integrants

In order to determine whether pSW3 integration into the bacterial chromosome has an adverse effect on msDNA synthesis, pDMS carrying recombinant bacteria W3110 *tel*$^+$ *vma*$^+$, W3110 *tel*$^+$, BW25113 *tel*$^+$ and BW25113 *tel*$^+$*vma*$^+$ were induced for msDNA synthesis (Figures 31 and 32). The *tel*-integrated W3110 *tel*$^+$ and BW25113 *tel*$^+$ bacterial strains were used as positive controls. Positive controls, upon heat induction, synthesized msDNA (Figure 31; lanes 11 & 12; Figure 32; lanes 1,4,6). The presence of a 3.1 kb LCC bacterial backbone and 3.3 kb msDNA fragment confirms msDNA synthesis. Recombinant bacterial cells with integrated *vma* and *tel* genes did not synthesize msDNA; however, the positive controls did. The integration of pSW3 (plasmid carrying the *vma* gene expression cassette) into the bacterial chromosome appears to have disrupted the Tel protelomerase enzyme activity. To further investigate the reason msDNA is not being synthesized, msDNA synthesis using the pNN9 precursor plasmid in double integrants was conducted.

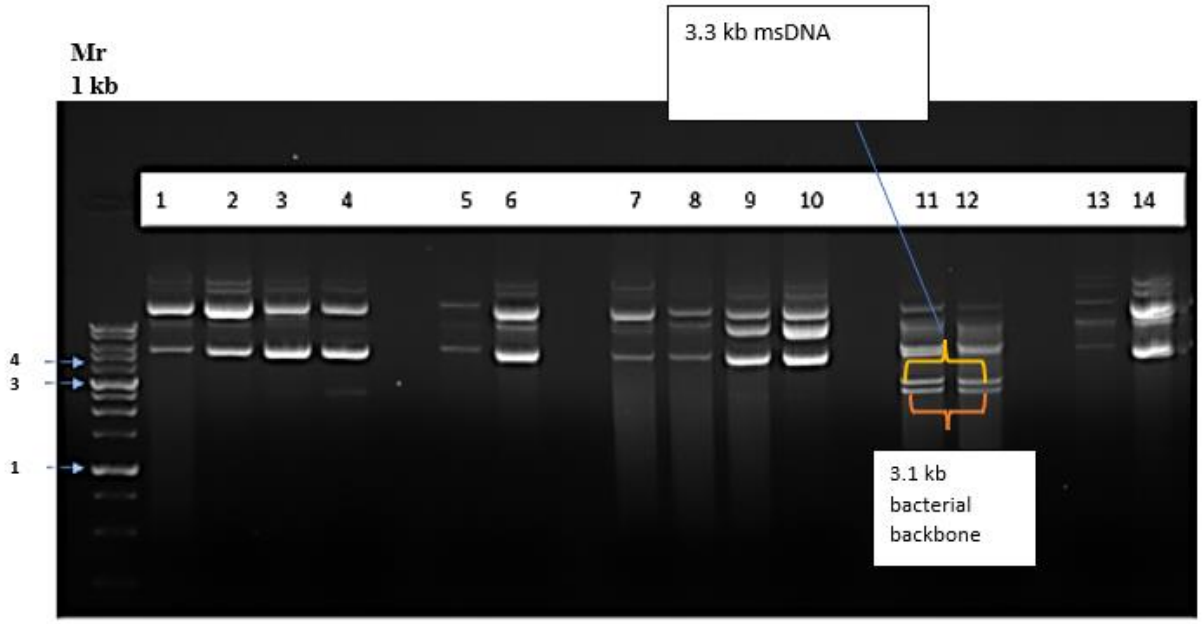


Figure 31:Heat Induction of msDNA synthesis in pDMS carrying W3110 *tel*$^{+}$*vma*$^{+}$, BW25113 *tel* $^{+}$*vma*$^{+}$ and W3110 *tel*$^{+}$ recombinant bacteria. 1. W3110 *tel* $^{+}$*vma*$^{+}$ (Induced); **2.** W3110 *tel* $^{+}$*vma*$^{+}$ (Induced); **3.** W3110 *tel* $^{+}$*vma*$^{+}$ (Induced); **4.** W3110 *tel* $^{+}$*vma*$^{+}$ (Un-Induced); **5.** W3110 *tel* $^{+}$*vma*$^{+}$ (Induced); **6.** W3110 *tel* $^{+}$*vma*$^{+}$ (Un-Induced); **7.** BW25113 *tel* $^{+}$*vma*$^{+}$ (Induced); **8.** BW25113 *tel* $^{+}$*vma*$^{+}$ (Induced); **9.** BW25113 *tel* $^{+}$*vma*$^{+}$ (Un-Induced);**10.** BW25113 *tel* $^{+}$*vma*$^{+}$ (Un-Induced); **11.** W3110 *tel*$^{+}$ (Induced); **12.** W3110 *tel*$^{+}$ (Induced); **13.** W3110 *tel*$^{+}$ (Un-Induced);**14.** W3110 *tel*$^{+}$ (Un-Induced). Mr: 1kb (FroggaBio).

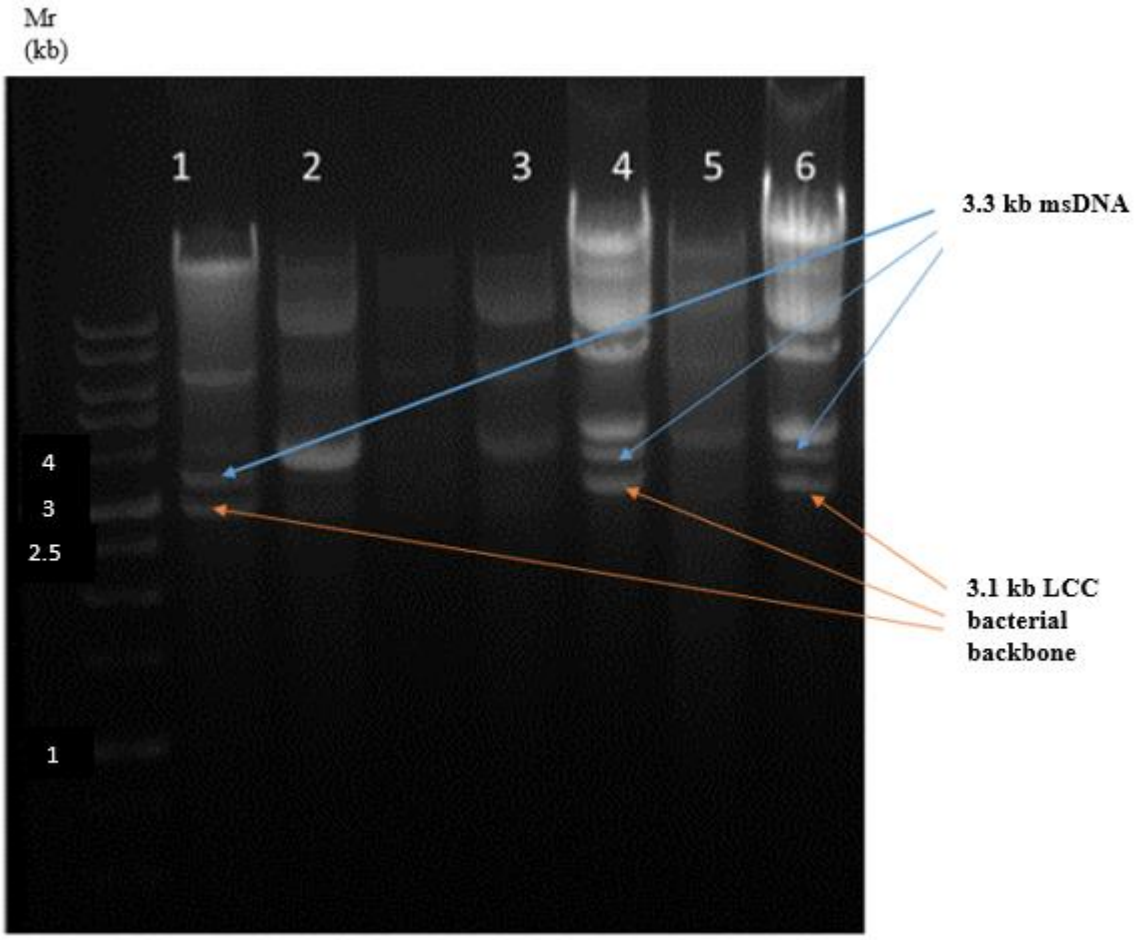


Figure 32:Heat induction of msDNA synthesis in BW25113 *tel*$^+$ recombinant bacteria carrying pDMS precursor plasmid. 1. BW25113 *tel*$^+$ (Induced); **2.** BW25113 *tel*$^+$ (Un-Induced); **3.** BW25113 *tel*$^+$ (Un-Induced); **4.** BW25113 *tel*$^+$ (Induced); **5.** BW25113 *tel*$^+$ (Un-Induced); **6.** BW25113 *tel*$^+$ (Induced). Mr: 1kb (FroggaBio).

4.2.5 Ministring DNA synthesis in double integrants with pNN9 precursor plasmid

Both *tel* and *vma* integrated double integrants were transformed with pNN9 msDNA synthesizing precursor plasmid and heat induction was performed to induce msDNA synthesis as described in section 3.4.3. Since there was no msDNA synthesis observed in double integrants in the presence of pDMS precursor plasmid, the msDNA synthesis with pNN9 precursor plasmid was conducted. As observed in Figure 33, lanes 3 and 4 show no difference in banding pattern between induced and un-induced cultures of double integrants upon heat induction. Only the positive controls were

able to synthesize msDNA, with the double integrants being unable to in the presence of the pNN9 precursor plasmid. Even though the data is not shown here, similar to BW25113 double integrants, W3110 double integrants didn't synthesize msDNA with the pNN9 plasmid.

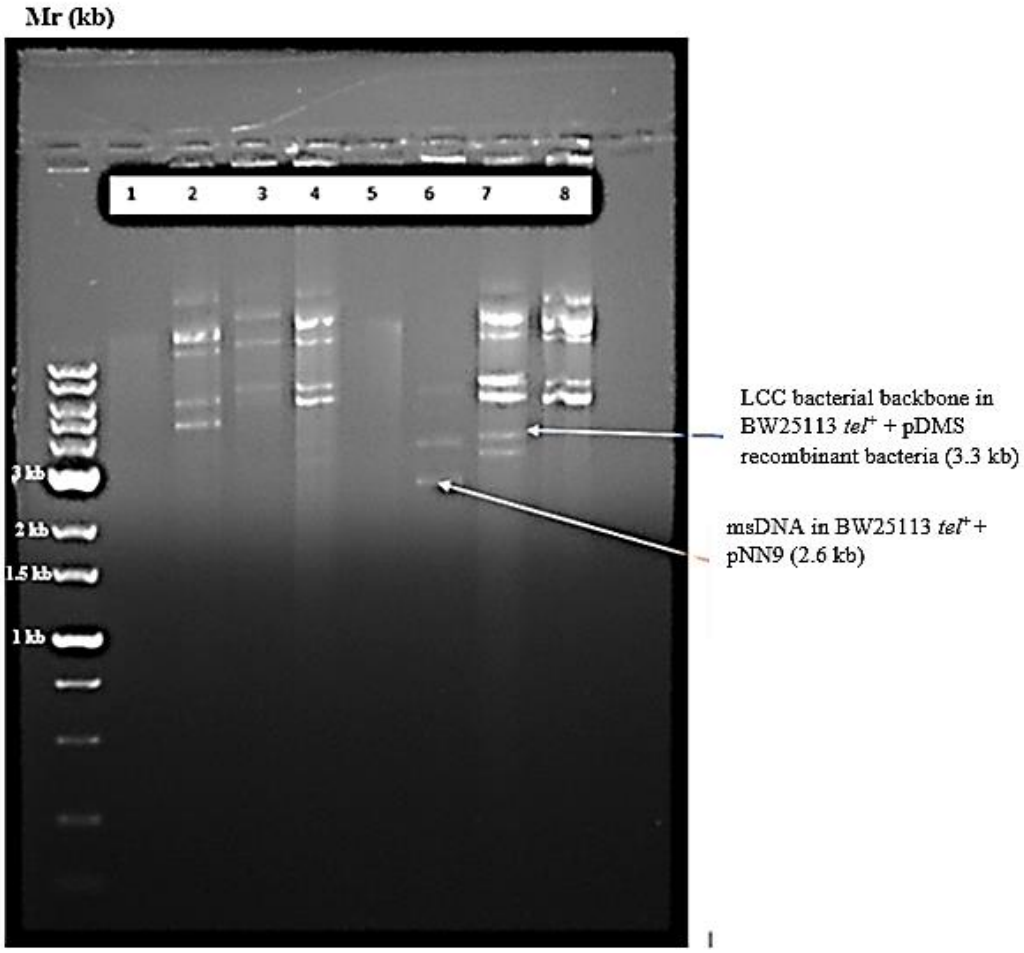


Figure 33:msDNA synthesis upon heat induction in double integrants carrying pNN9 plasmid. msDNA synthesis was induced upon heat induction in BW25113 *tel* $^{+}$*vma*$^{+}$+ pNN9 recombinant bacteria. Lanes:**1.** & **2.** Not applicable (N/A);**3.** BW25113 *tel* $^{+}$*vma* $^{+}$+ pNN9 (Induced);**4.** BW25113 *tel* $^{+}$*vma* $^{+}$+pNN9 (Un-Induced);**5.** N/A;**6**. BW25113 *tel*$^{+}$+pNN9 (Induced);**7.** BW25113 *tel*$^{+}$+pDMS (Induced);**8.** BW25113 *tel*$^{+}$+pNN9 (Un-Induced). Mr: 1kb (FroggaBio).

As described in section 1.4, the main difference between pDMS and pNN9 plasmids is the presence of the PI-Sce1 homing endonuclease enzyme recognition sequence in the pDMS plasmid (Figure 9 A). PI-SceI binds tightly to its target site and remains bound to the DNA after cleavage of the DNA occurs. The Tel protelomerase enzyme is also a DNA binding protein and may locate the Tel protelomerase enzyme binding target site by sliding along the non-specific portion of the plasmid until the specific target is found. If PI-SceI is bound to the same DNA at its recognition sites, it might be impeding the protelomerase from finding its target at a high frequency. The experiment was conducted to determine whether the leaky overexpression of *vma* gene has an adverse effect on msDNA synthesis, as indicated in the earlier statement. Since double integrants were unable to synthesize msDNA even in the presence of pNN9 (which does not contain the PI-Sce1 recognition sequences) does not support the above argument. As shown in Figure 34 the chromosomal locations of *tel* and *vma* genes in double integrants are not in close enough proximity to conclude that *vma* gene integration disrupted the chromosomally integrated *tel* gene sequence. The *tel* gene expression gene cassette was integrated into the *lacZ* gene of the W3110 bacteria via homologous recombination using the pBRINT-cat integrating plasmid by Dr. Nafissi—a previous member of the Slavcev lab—in 2012. The *vma* gene expression cassette was integrated into the *λattP* site in the W3110 tel^+ bacterial chromosome via site-specific recombination using pSW3 integrating plasmid. Both *tel* and *vma* genes were integrated via site-specific recombination where chromosomal locations are known. A different approach was used in the generation of BW25113 tel^+vma^+ double integrants, where P1 transduction was used to integrate the *tel* gene expression gene cassette into the bacteria. Despite P1 transduction having been used to transfer gene fragments from one bacterial strain to another, and the gene integration being through homologous recombination, the chromosomal location of *tel* gene is similar to that of W3110 tel^+ recombinant

bacteria. The *vma* gene was integrated into BW25113 *tel*$^+$ recombinant bacteria, similarly to that of W3110 *tel*$^+$*vma*$^+$ double integrants. Since both BW25113 and W3110 double integrants did not synthesize msDNA, JM109 double integrants were generated. A second approach to generate JM109 double integrants was employed by integrating the *vma* gene via P1 transduction into JM109 *tel*$^+$ recombinant bacteria, where *tel* gene was already integrated into the JM109 bacterial chromosome via P1 transduction. W3110, BW25113 and JM109 bacterial cells with different stabilities were used to develop a system that favors msDNA synthesis in double integrants.

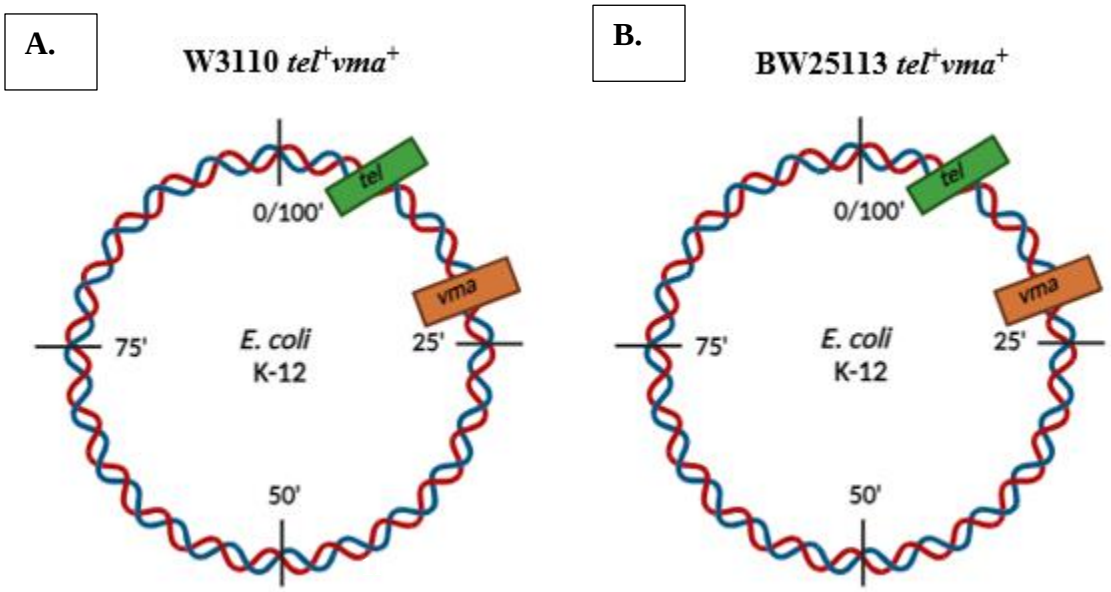


Figure 34:Chromosomal location of the *tel* and *vma* genes in double integrants. A. The location of the *tel* and *vma* genes in W3110 *tel*$^+$*vma*$^+$ double integrants. The *tel* gene expression cassette where the *tel* gene is placed under the control of the heat labile lambda repressor CI-857 was integrated into the *lacZ* gene in the bacterial chromosome via homologous recombination using the pBRINT-cat plasmid by Dr. Nafissi in 2012. The *vma* gene expression cassette was integrated into the *λattP* site in the bacterial chromosome via site-specific recombination using pAH120 integrating plasmid. **B.** The chromosomal location of the *tel* and *vma* integrated double integrants of BW25113 *tel*$^+$*vma*$^+$ recombinant bacteria. The *tel* gene expression cassette was integrated into BW25113 bacterial chromosome via P1 transduction and the *vma* gene cassette was integrated into the chromosome using pSW3 integrating plasmid via site-specific recombination.

4.3 Generation of JM109 *tel*$^{+}$*vma*$^{+}$ double integrants via P1 transduction

In JM109 *tel*$^{+}$ recombinant bacteria, the *tel* gene expression cassette contained a Cm^{R} selectable marker. Generalized transduction using P1rev6 was used to transfer the *vma* gene expression cassette linked with Km resistance into the JM109 *tel*$^{+}$ recombinant bacteria. When conducting the transduction, two negative controls were mixed, as shown in Figure 35.

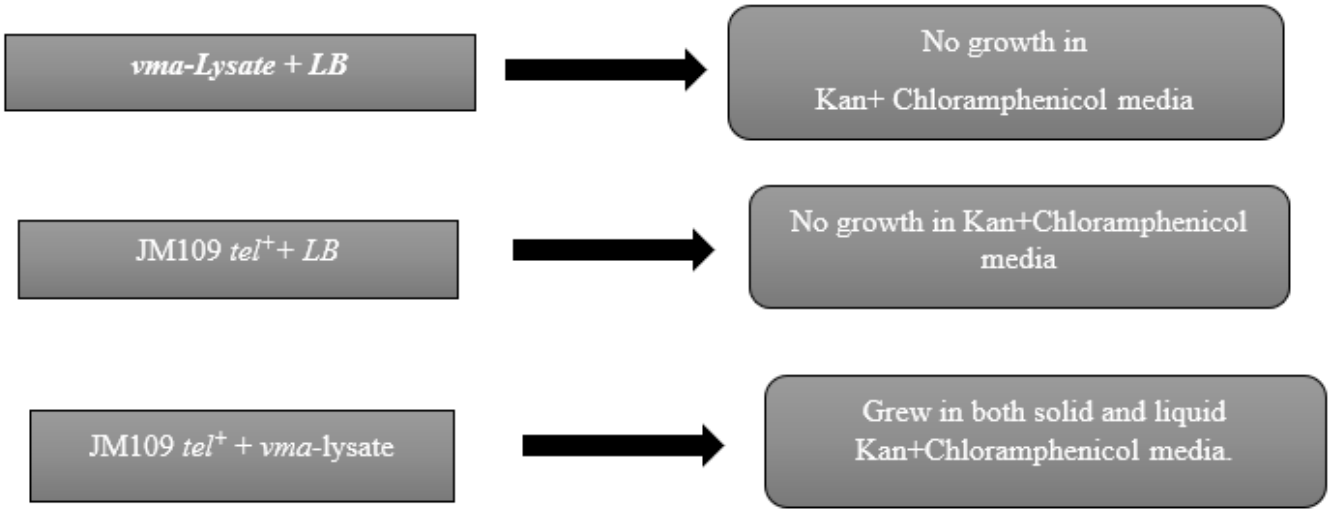


Figure 35:Transducing JM109 *tel*$^{+}$ with *vma* lysate. The *vma* lysate was mixed with JM109 *tel*$^{+}$ recombinant bacteria. Two negative controls were additionally run for the experiment to confirm the validity of the positive recombinant colonies in both Km and Cm containing LB plates. As shown in the diagram, two negative controls (*vma* lysate and JM109 *tel*$^{+}$) should not exhibit growth in LB media containing Km and Cm.

After the transduction the cell suspensions were streaked on both chloramphenicol and kanamycin antibiotic containing LB plates. The concentrations of the antibiotics added are as outlined in section 3.1.1. As expected, there were colonies observed in JM109 *tel*$^{+}$ transduced suspension-streaked plates, not in the negative controls. The positive transductant colonies were picked and re-streaked on LB+ Km and Cm containing plates. Colony PCR was performed to further

confirmed the presence of *vma* gene by amplifying the segment of *vma* gene using PCR. As in Figure 36, were able to obtain 1046 bp band in colonies tested. This confirmed the presence of *vma* gene in the chromosome. Since the positive transductant grew on LB plates supplemented with both chloramphenicol and kanamycin antibiotics and also the colony PCR confirmed the presence of *vma* gene, confirmed the generation of JM109 *tel*$^{+}$*vma*$^{+}$ double integrants.

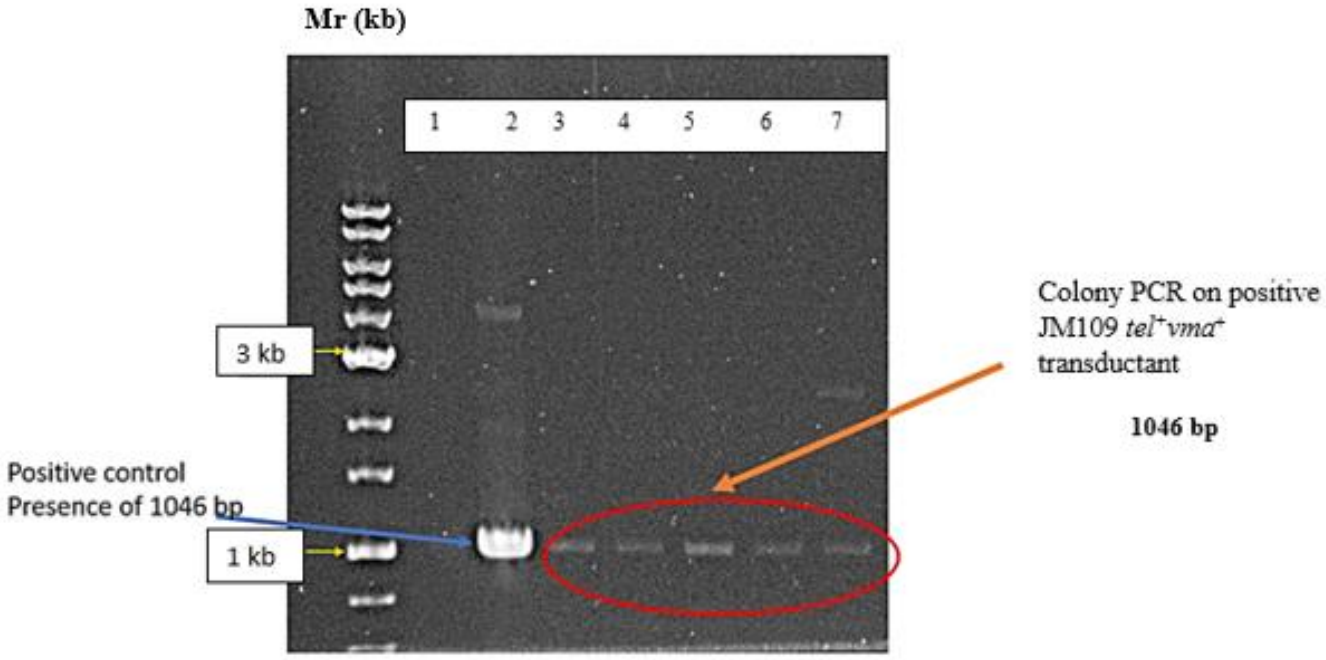


Figure 36:Colony PCR for positive transductant colonies. Colony PCR was conducted for colonies that grew on both Cm and Km added plates. Lanes;**1.** Ignore;**2**. The positive control (W3110 *vma*$^{+}$); **3-7.** The colonies that tested positive and showed an ~1046 bp band. Mr: 1kb (FroggaBio).

4.3.1 Induction of msDNA synthesis and homing endonuclease enzyme overexpression in JM109 *tel*$^{+}$*vma*$^{+}$ double integrants

JM109 *tel*$^{+}$*vma*$^{+}$ double integrants carrying the pNN9 and pDMS msDNA synthesizing precursor plasmids were subjected to msDNA synthesis via heat induction. The positive controls, JM109 *tel*$^{+}$

and W3110 *tel*$^{+}$ recombinant bacteria carrying the pDMS plasmid, synthesized msDNA upon heat induction; however, the double integrants did not (Figure 37 and Figure 38 (A and B)). The synthesis of msDNA was further confirmed by digesting the plasmid extract with EcoR1. Tables 10 and 11 summarize the expected DNA fragment sizes upon digestion with EcoR1. Since pDMS and pNN9 have only one EcoR1 recognition site present in the plasmid, they should give 6.4 and 5.6 kb bands, respectively. If msDNA is synthesized with pDMS and pNN9 precursor plasmids, upon EcoR1 digestion, band sizes of approximately 0.6, 2.7, 3.1, 1.2, and 1.4 kb should be observed, respectively, as shown in Table 10. As indicated in Figure 35, after digesting the plasmid extract with EcoR1 no bands corresponding to msDNA synthesis were observed. Instead, 5.6 kb and 6.4 kb bands (Figure 37, yellow highlighted arrows) corresponding to pNN9 and pDMS were visible. This confirms that there was no msDNA synthesis observed in double integrants. Even though the double integrants couldn't synthesize msDNA, they were able to over-express the PI-Sce1 homing endonuclease enzyme and digest the pDMS precursor plasmid. As shown in Figure 38 (C; lanes 6-9), both induced and un-induced conditions showed digestion of the pDMS msDNA synthesizing precursor plasmid.

P**lasmids**	**ms**DNA (**kb**)	B**acterial backbone** (**kb**)
pNN9	2.6	3.0
pDMS	3.3	3.1

Table 10:Sizes of msDNA and bacterial backbone fragments in the presence of pNN9 and pDMS.

Plasmid DNA	**Sizes of the DNA fragments when digested with EcoR1 (kb)**
pNN9	5.6
pDMS	6.4
msDNA with pDMS + bacterial backbone	0.6, 2.7, 3.1
msDNA with pNN9++ bacterial backbone	1.2, 1.4, 3.0

Table 11:Plasmid and msDNA fragment sizes after digestion with EcoR1.

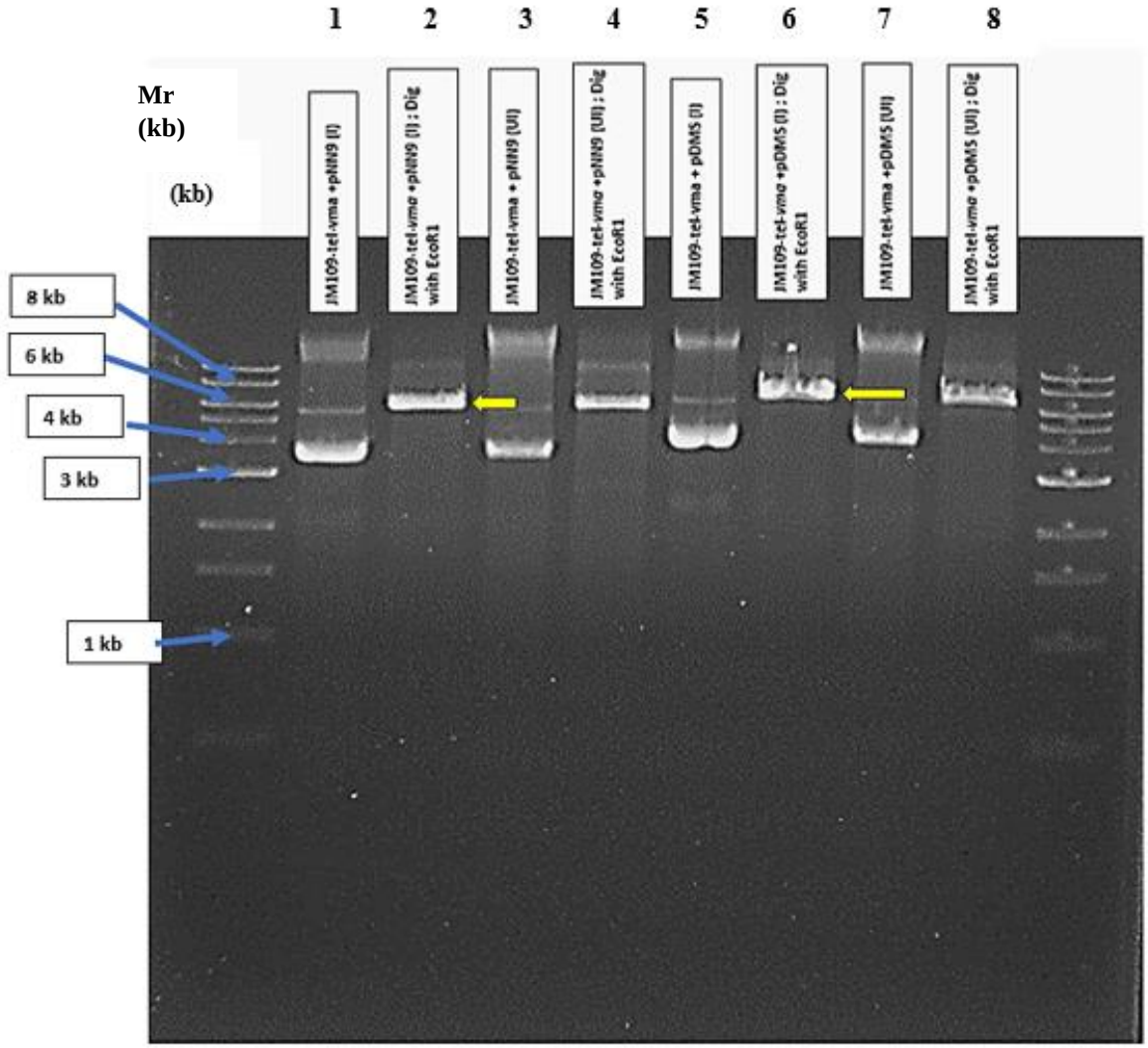


Figure 37:msDNA Synthesis in JM109 *tel*$^{+}$*vma*$^{+}$ double integrants with pNN9 and pDMS msDNA synthesizing precursor plasmids. Plasmids were extracted after msDNA synthesis and run on 0.8 % agarose gel. Lanes;**1.** JM109 *tel*$^{+}$*vma*$^{+}$+pNN9 (Induced), **2.** JM109 *tel*$^{+}$*vma*$^{+}$+ pNN9 (Induced) Digested with EcoR1;**3.**JM109 *tel*$^{+}$*vma*$^{+}$+pNN9 (Un-Induced);**4.**JM109 *tel*$^{+}$*vma*$^{+}$+pNN9 (Un-Induced) digested with EcoR1;**5.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Induced);**6.** JM109 *tel*$^{+}$*vma*$^{+}$ +pDMS (Induced) digested with EcoR1;**7.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS(Un-induced);**8.**JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Un-Induced) digested with EcoR1.Mr 1kb Ladder (FroggaBio).

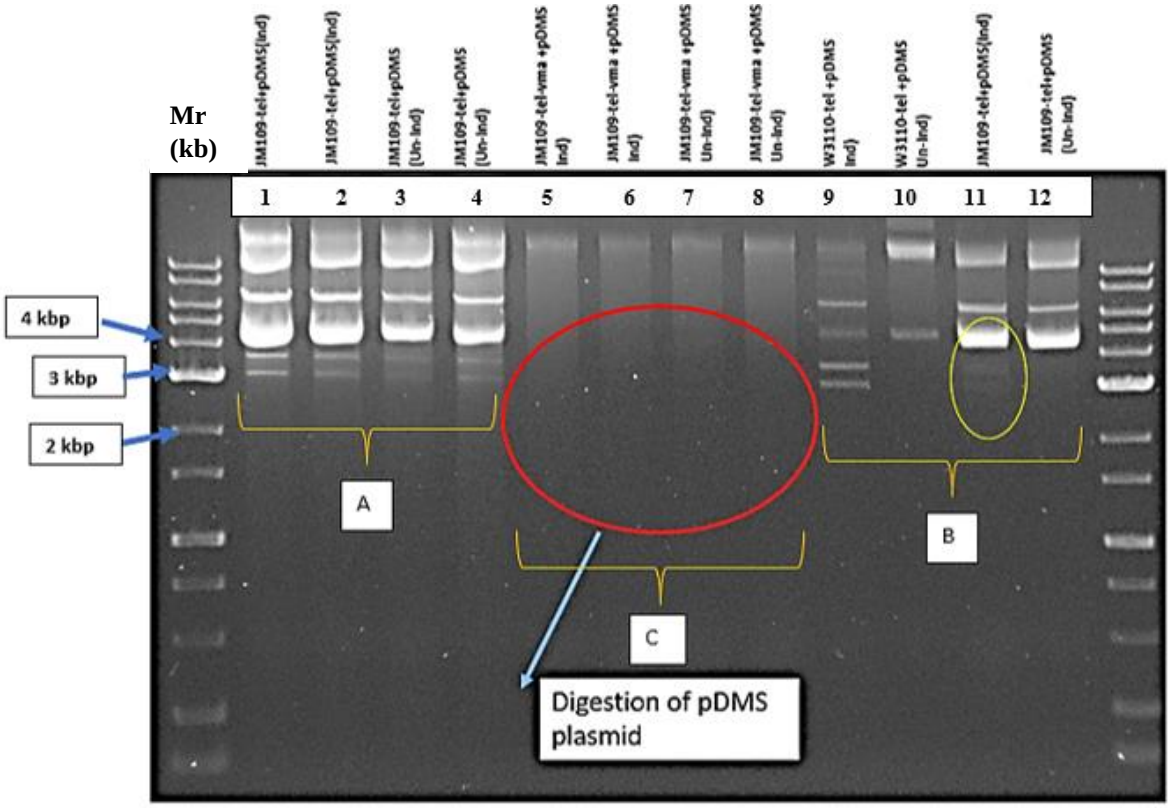


Figure 38: Induction of msDNA synthesis and homing endonuclease enzyme over expression. PI-Sce1 homing endonuclease enzyme over expression in pDMS carrying JM109 *tel*$^{+}$*vma*$^{+}$ double integrant recombinant bacteria. Plasmids were extracted and run on a 0.8 % agarose gel. **A.** Heat induction was performed to induce msDNA synthesis; Lanes; **1.** JM109 *tel*$^{+}$+pDMS (Ind), **2.** JM109 *tel*$^{+}$+pDMS (Ind), **3.** JM109 *tel*$^{+}$+pDMS (Ind), **4.** JM109 *tel*$^{+}$+pDMS (Ind), **B.** Heat Induction was performed to induce msDNA synthesis. **9.** W3110 *tel*$^{+}$+pDMS (Ind), **10.**W3110 *tel*$^{+}$+pDMS (Un-Induced), **11.** JM109 *tel*$^{+}$+pDMS (Induced), **12.** JM109 *tel*$^{+}$+pDMS (Un-Induced) **C.** *vma* gene over expression **5.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Induced), **6.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Un-Induced), **7.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Induced), **8.** JM109 *tel*$^{+}$*vma*$^{+}$+pDMS (Un-Induced)**.** Mr 1kb Ladder (FroggaBio).

4.4 Episomal expression of the Tel protelomerase enzyme

As described in section 4.2.4, double integrants, both *tel* and *vma* integrated recombinant bacteria didn't synthesize msDNA upon *tel* gene over expression. Therefore, experiments were conducted to express the Tel protelomerase enzyme episomally instead of integrating *tel* gene into the bacterial chromosome. In order to enhance the msDNA synthesis episomal expression of the *tel*

gene was achieved by cloning the *tel* gene expression gene cassette placed under the control of different inducible promoters, such as heat-inducible lambda repressor, P_{BAD} promoter, and T7 promoter into low copy number plasmids. At first wild-type bacteria carrying Tel plasmid and precursor plasmids were induced for msDNA synthesis (Figure 39). If msDNA synthesis was observed, induction of msDNA synthesis in *vma* integrated recombinant bacteria was conducted. Following msDNA synthesis L-arabinose was added to induce *vma* gene over-expression to digest unwanted bacterial backbone and the unprocessed precursor plasmids, leaving only msDNA and the Tel plasmid (Figure 40).

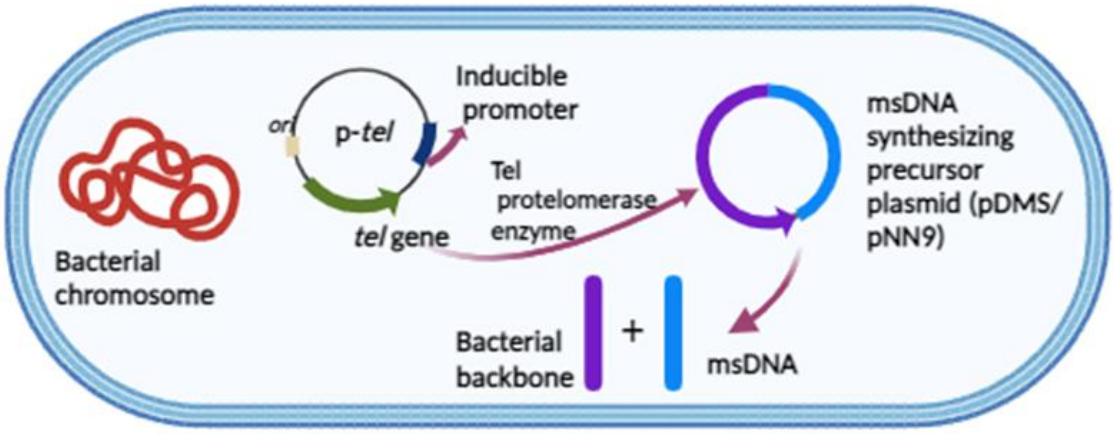


Figure 39:Induction of msDNA synthesis in recombinant wild type bacteria carrying p-Tel plasmids and pNN9/pDMS precursor plasmids. The W3110 bacteria were transformed with p-Tel plasmids and msDNA-synthesizing precursor plasmids. The *tel* gene is placed under the control of different inducible promoters (heat inducible, T7, P_{BAD}). Upon induction, the p-Tel plasmid overexpresses the Tel protelomerase, which synthesizes msDNA by acting on the precursor plasmids.

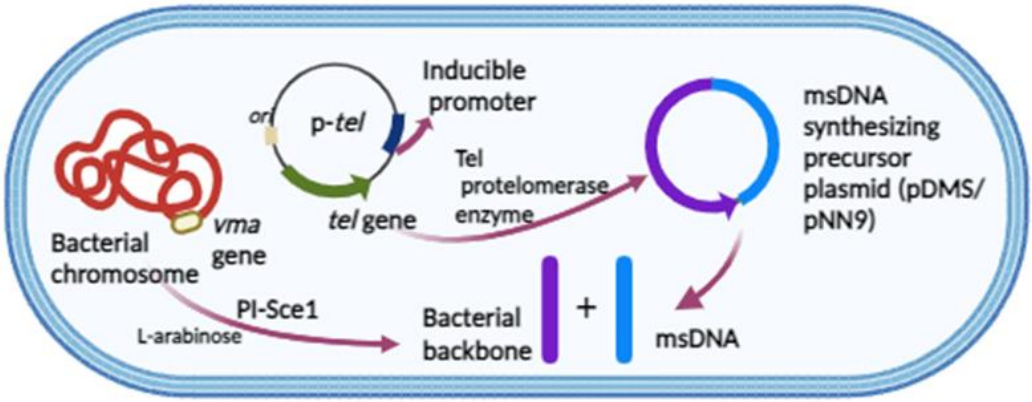


Figure 40:Induction of msDNA synthesis in *vma* integrated recombinant bacteria. After msDNA synthesis, the *vma*-integrated recombinant bacteria were induced for *vma* gene expression to digest non-msDNA species, such as unwanted bacterial backbone and unprocessed precursor plasmids. This process resulted in leaving only msDNA and p-Tel recombinant plasmids.

4.4.1 Induction of msDNA synthesis in W3110/JM109 recombinant bacteria carrying the pACYC-CI-857-tel plasmid in the presence of pDMS/pNN9 precursor plasmids

W3110 and JM109 bacteria were transformed with pACYC-CI-857-tel and pNN9/pDMS msDNA synthesizing precursor plasmids as described in section 3.4.5.1. Upon heat induction, *tel* gene should over-express the Tel protelomerase enzyme and act on pNN9/pDMS to synthesize msDNA. The expected test results are summarized in Table 12 and Table 13. However, the actual test results obtained differed from the expected test results (Table 14 and 15). As shown in Figure 41, pACYC-CI-857-tel and pNN9 plasmids carrying W3110 recombinant bacteria upon heat induction did not synthesize msDNA, only positive control, W3110 tel^+ + pNN9 recombinant bacteria, synthesized msDNA. If msDNA was synthesized there should be 2.6 kb and 3.0 kb DNA fragments corresponding to msDNA and bacterial backbone respectively. As depicted in Figure 41, under both induced and un-induced conditions W3110+pACYC-CI-857-tel+pNN9 recombinant bacteria upon heat induction synthesized ~2.2 kb fragment and a faint 3 kb DNA fragment. These bands do not match with msDNA and there is no difference in the banding pattern between the induced

and un-induced cultures. These bands could represent different isoforms of the precursor plasmid and p-Tel plasmid. Similarly, as shown in Figure 42 upon heat induction JM109+pDMS+pACYC-CI-857-tel recombinant bacteria didn't synthesize msDNA. If msDNA was synthesized there should be 3.3 kb and 3.1 kb bands corresponding to msDNA and bacterial backbone respectively. Even though the insert, *CI-857-tel*, was successfully cloned into the pACYC184 vector, the over-expressed Tel protelomerase enzyme might have formed highly aggregated proteins called inclusion bodies. The following factors favor the argument. Due to *pL-pR* being a strong promoter, this could have resulted in high expression of the *tel* gene. As the *tel* gene was cloned into a low copy number plasmid, this likely resulted in the synthesis of a higher copy number of the target *tel* gene. This strong promoter is induced by temperature, and hence, continuous supply of heat resulted in higher expression of the recombinant Tel protelomerase enzyme which could have led to inclusion body formation.

Recombinant Bacteria-generated	**Induced (42 ℃)**	**Un-induced (30 ℃)**
W3110+ pACYC-CI-857-tel	No msDNA	No msDNA
W3110+ pACYC-CI-857-tel + pNN9	msDNA synthesis	No msDNA
W3110+pNN9	No msDNA	No msDNA
W3110 + pACYC	No msDNA	No msDNA
W3110 + pACYC + pNN9	No msDNA	No msDNA
W3110 *tel*$^{+}$ + pNN9	msDNA synthesis	No msDNA

Table 12:Expected test results upon heat induction in W3110 bacteria carrying p-Tel plasmid and precursor plasmid.

Recombinant Bacteria generated	Induced (42 ℃)	Un-induced (30 ℃)
JM109+ pACYC-CI-857-tel	No msDNA	No msDNA
JM109+ pACYC-CI-857-tel + pDMS	msDNA synthesis	No msDNA
JM109+pDMS	No msDNA	No msDNA
JM109 + pACYC	No msDNA	No msDNA
JM109 + pACYC + pDMS	No msDNA	No msDNA
W3110 *tel*$^{+}$ + pDMS	msDNA synthesis	No msDNA

Table 13:Expected test results upon heat induction in JM109 bacteria carrying pDMS and p-Tel plasmids.

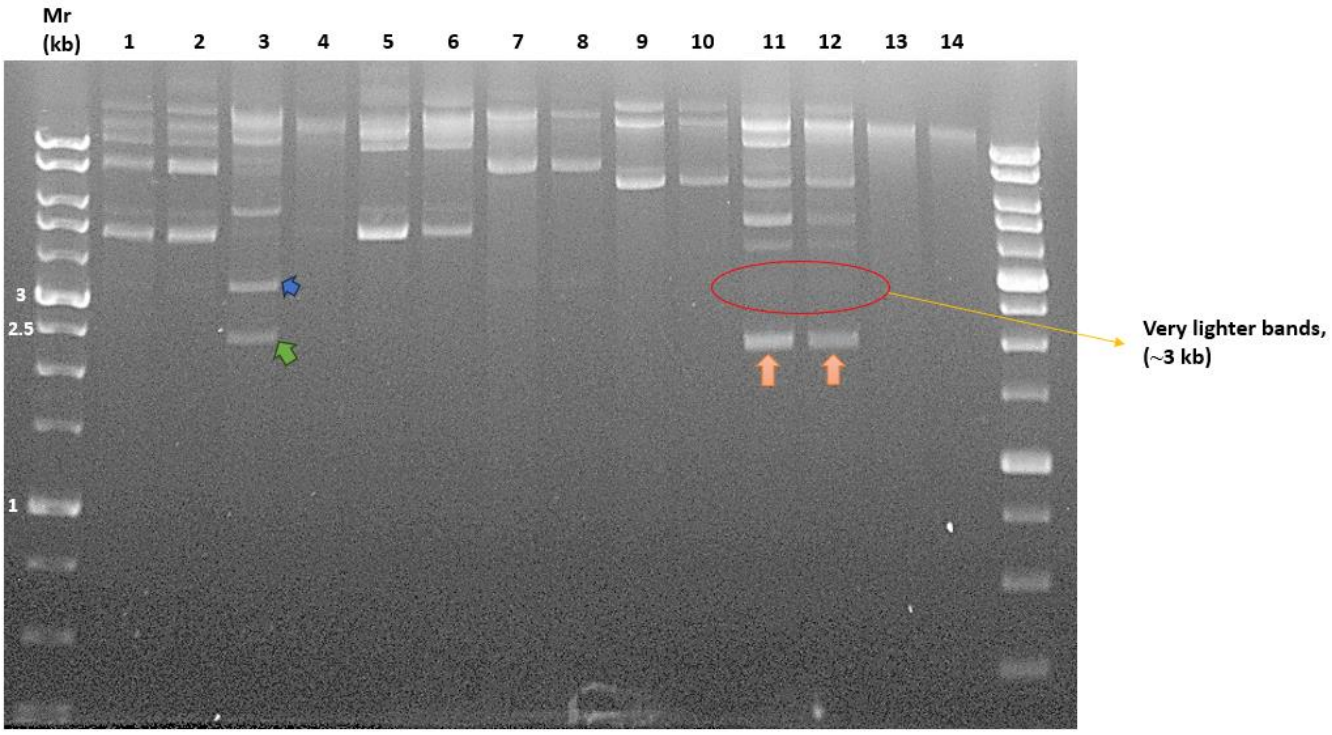


Figure 41:Induction of msDNA synthesis in W3110 bacteria carrying pACYC-CI-857-tel and pNN9 plasmids. msDNA synthesis was induced upon heat induction and the plasmids were extracted and run on 0.8 % Agarose gel. Lanes; **1.**W3110+pACYC+pNN9 (Induced), **2.** W3110+pACYC+pNN9 (Un-Induced); **3.** W3110 *tel*$^{+}$+pNN9 (Induced); **4.** W3110 *tel*$^{+}$+pNN9

(Un-Induced); **5.** W3110+pNN9 (Induced); **6.** W3110+pNN9 (Un-Induced); **7.**W3110+pACYC (Induced); **8.** W3110+pACYC (Un-Induced);**9.**W3110+pACYC-tel(Induced);**10.** pACYC-tel(Un-Induced); **11.**W3110+pACYC-tel+pNN9 (Induced);**12.**W3110+pACYC-tel+pNN9 (Un-Induced);**13.** W3110 (Induced). **14.**W3110 (Un-Induced). Arrows indicate bands of interest; blue (3.0 kb), green (2.6 kb), orange (2.2 kb). Mr 1 kb ladder (FroggaBio).

Recombinant bacteria	**Expected fragment sizes upon msDNA synthesis (kb)**	**The sizes of the observed fragments (kb)**
W3110+pACYC+pNN9 (Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pACYC+pNN9(Un-Induced)	No msDNA synthesis	No msDNA synthesis
W3110 *tel*$^{+}$ + pNN9 (Induced)	2.6, 3.0	2.6, 3.0
W3110 *tel*$^{+}$ + pNN9 (Un-Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pNN9 (Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pNN9 (Un-Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pACYC (Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pACYC (Un-Induced))	No msDNA synthesis	No msDNA synthesis
W3110+pACYC-CI-857-tel + pNN9 (Induced)	2.6, 3.0	~2.2, 3.0
W3110+pACYC-CI-857-tel + pNN9 (Un-Induced)	No msDNA synthesis	~2.2, 3.0

Table 14:Summary of the expected fragment sizes versus the observed fragment sizes upon heat-dependent induction of msDNA synthesis in W3110 recombinant bacteria.

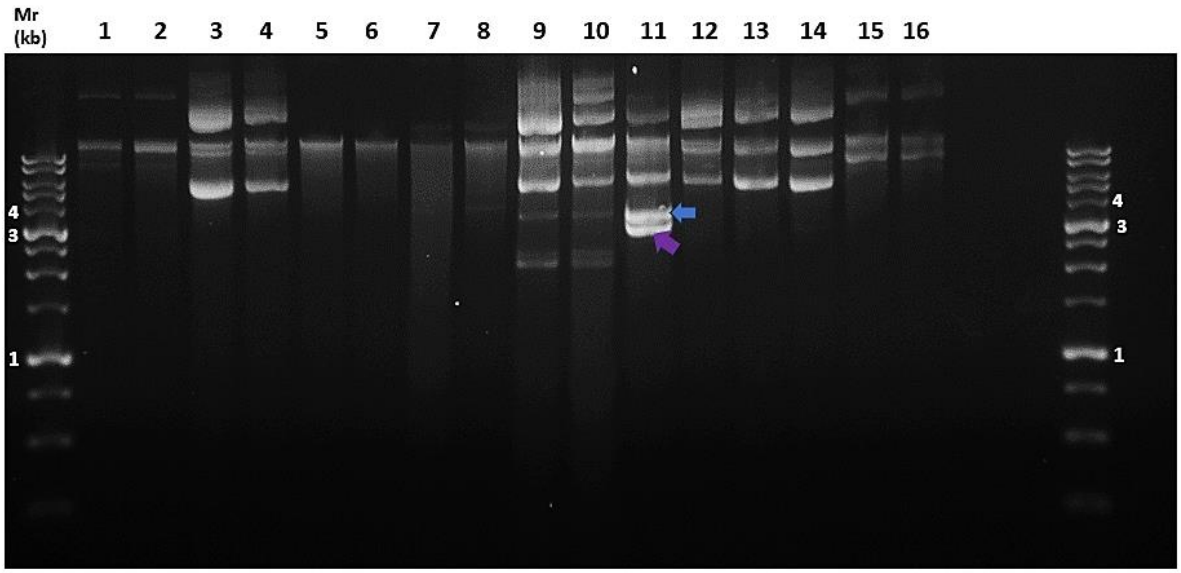


Figure 42:Induction of msDNA Synthesis in JM109+pACYC-CI-857-tel+pDMS recombinant bacteria. msDNA synthesis is induced upon heat induction and the plasmids were isolated and ran on 0.8 % agarose gel to visualize.Lanes;**1.**JM109+pACYC (Induced);**2.** JM109+pACYC(Un-Induced);**3.**JM109+pDMS(Induced);**4.**JM109+pDMS(Un-Induced);**5.**JM109(Induced);**6.**JM109(Un-Induced);**7.**N/A;**8.**N/A;**9.**JM109+pACYC-CI-857-tel+pDMS(Induced);**10.**JM109+pACYC-CI-857-tel+pDMS (Un-Induced);**11.**JM109 *tel*$^{+}$+pDMS (Induced);**12.**JM109*tel*$^{+}$+pDMS(Un- Induced);**13.**JM109+pACYC+pDMS(Induced);**14.**JM109+ pACYC+pDMS (un-induced); **15.**JM109+pACYC-CI-857-tel (Induced); **16.** JM109+pACYC-*CI*-857-tel (Un-Induced). Arrows indicate bands of interest; blue (3.3 kb), purple (3.1 kb). Mr 1 kb (FroggaBio).

Recombinant bacteria	Expected fragment sizes upon msDNA synthesis (kb)	The sizes of the observed fragments (kb)
JM109+pACYC (Induced)	No msDNA synthesis	No msDNA synthesis
JM109+pACYC (Un-Induced)	No msDNA synthesis	No msDNA synthesis
JM109+pDMS (Induced)	No msDNA synthesis	No msDNA synthesis
JM109+pDMS (Un-Induced)	No msDNA synthesis	No msDNA synthesis
JM109+pACYC-CI-857-tel + pDMS (Induced)	3.3, 3.1	2.4,2.6, 3.1
JM109+pACYC-CI-857-tel (Un-Induced)	No msDNA synthesis	2.4,2.6, 3.1
JM109 *tel*$^{+}$+pDMS (Induced)	3.3,3.1	3.3,3.1
JM109 *tel*$^{+}$+pDMS (Un-Induced)	No msDNA synthesis	No msDNA synthesis

Table 15:Summary of the expected fragment sizes versus the observed fragment sizes upon heat-dependent induction of msDNA synthesis in JM109 recombinant bacteria.

4.4.2 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-pBAD and pDMS/pNN9 precursor plasmids

The pACYC-tel-pBAD recombinant plasmid was a gift from Mediphage Bioceuticals (Toronto, CA). This low copy number recombinant plasmid has the p15 *ori* and the *tel* gene which was placed under the control of an inducible P_{BAD} promoter. W3110 bacteria were transformed with the recombinant Tel plasmid and the msDNA synthesizing precursor plasmid as outlined in section 3.4.6.3. Since the pACYC-tel-pBAD and the msDNA synthesizing pDMS/pNN9 (ColE1 *ori*) precursor plasmids have compatible *ori*, the two plasmids should coexist in W3110 recombinant bacteria upon successful transformation of pACYC-tel-pBAD and pDMS/pNN9 precursor plasmids via electroporation. After the successful transformation of W3110 with recombinant plasmids, bacterial cells grew on both Cm and Ap supplemented LB agar plates. Even though the positive transformants grew on selective LB-agar plates, there was no growth observed in

overnight cultures. Due to an unknown reason the two plasmids didn't coexist in W3110 recombinant bacteria upon transformation and couldn't continue the experiment.

4.4.3 Induction of msDNA synthesis in recombinant bacteria carrying pACYC-tel-T7 and pDMS/pNN9 recombinant plasmids

BL21(DE3) and W3110 bacteria were transformed with the pACYC-tel-T7 **(**T7-tel) and precursor plasmids and induced msDNA synthesis as described in section 3.4.6.2. BL21(DE3) bacteria has the chromosomally integrated prophage DE3 which carries the T7 RNA Polymerase encoding gene placed under the control of lacUV5 promoter (Jeong et al., 2015). BL21(DE3) bacteria is commonly used for high-level recombinant protein expression and upon addition of IPTG activates the lacUV5 promoter which induces the expression of T7 RNA polymerase enzyme. T7 RNA polymerase upon activation of the T7 promoter transcribes *tel* gene eight times faster than the *E. coli* RNA polymerase (Du et al., 2021). As shown in Figure 43 at 0.1 mM and 0.4mM IPTG concentrations both induced and un-induced cultures in the presence of pNN9 and pDMS precursor plasmids synthesized msDNA. The presence of 2.6 kb (orange arrow) and 3.0 kb (red arrow) bands in recombinant bacteria carrying T7-tel and pNN9 plasmids confirm the synthesis of msDNA. The presence of 3.1 kb (purple arrow) and 3.3 kb (blue arrow) confirms the synthesis of msDNA in recombinant bacteria carrying T7-tel and pDMS recombinant plasmids (Table 16). There was no significant difference observed between the added IPTG concentration vs. the amount of msDNA synthesized. Overall, msDNA was synthesized at significantly lower concentrations compared to msDNA synthesis inside W3110 and BW25113 with msDNA synthesizing precursor plasmids. This could be due to the fact that leaky overexpression of the T7 promoter resulted in the overexpression of Tel protelomerase enzyme even in the absence of inducer IPTG. This might have resulted in converting the precursor plasmid to msDNA and bacterial backbone even under

uninduced conditions (at early stages of bacterial growth), which halted the amplification of the msDNA synthesizing precursor plasmid. According to Du et al. (2021), the lacUV5 promoter is a leaky promoter that leads to the expression of target genes even under uninduced conditions. If the early expression and accumulation of the target protein are toxic to the cell, as in this experiment where early expression of the Tel protelomerase enzyme converts precursor plasmids into msDNA, there might not be enough precursor plasmids to synthesize msDNA abundantly.

W3110 bacteria were transformed with T7-tel and msDNA synthesizing precursor plasmids as described in section 3.4.6.2 to determine whether a tight regulation of the T7 promoter can be achieved in the absence of T7 RNA Polymerase enzyme. IPTG was added to a final concentration of 1 mM to induce protelomerase enzyme over-expression. As shown in Figure 44 there was no msDNA synthesis observed either with pNN9 nor with pDMS precursor plasmids (Figure 44; lanes 1- 4). The expected band sizes corresponding to msDNA and bacterial backbone upon msDNA synthesis with pNN9 and pDMS precursor plasmids as depicted in Table 17 don't match with actual results obtained in induced cultures. The T7 promoter is designed to be recognized and transcribed by T7 RNA polymerase enzyme. RNA polymerase in W3110 bacteria is not the same as bacteriophage T7 RNA polymerase, therefore T7 promoter was not recognized by *E. coli* RNA polymerase for its activation and efficient transcription of the *tel* gene.

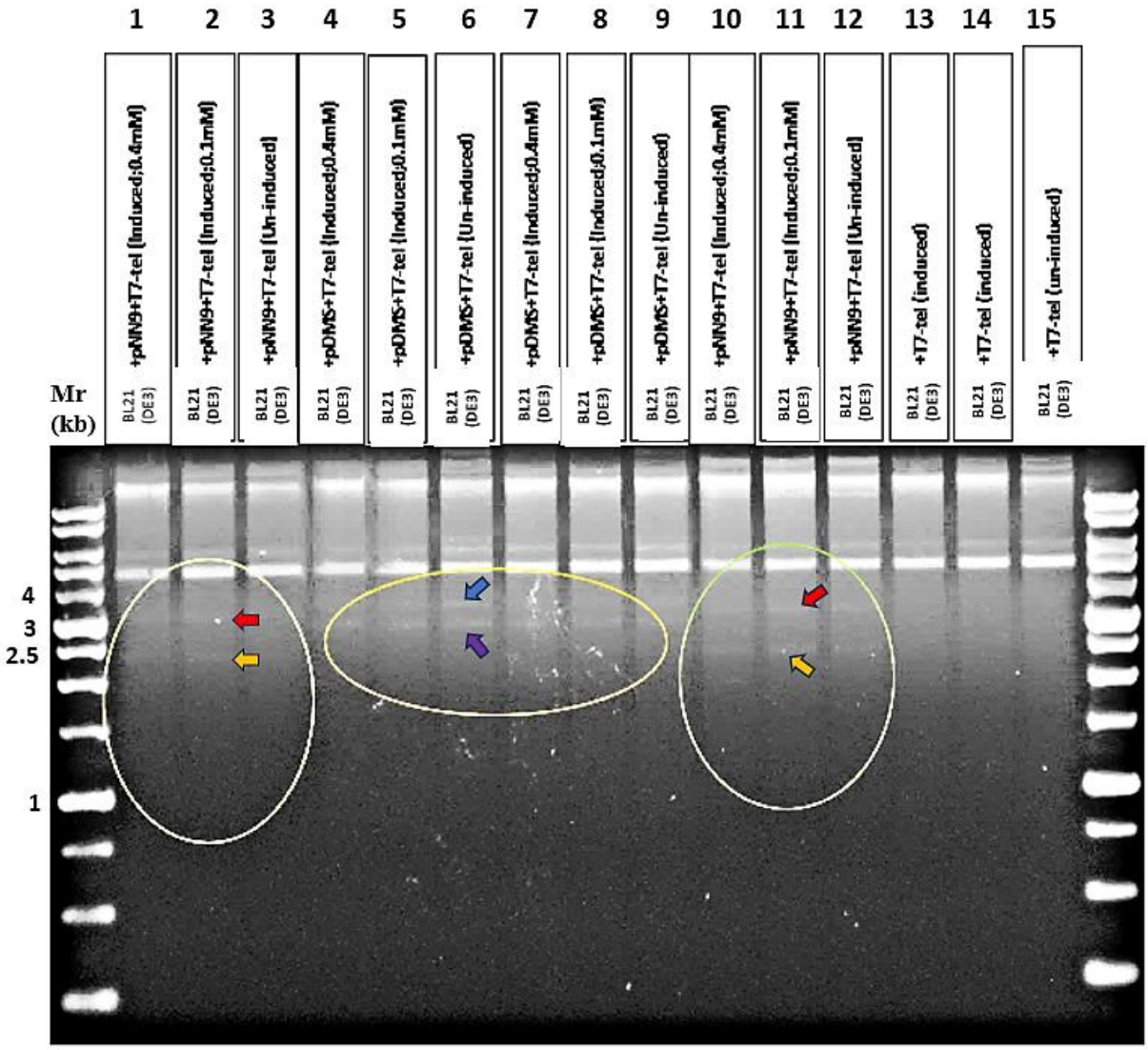


Figure 43:Induction of msDNA synthesis in BL21(DE3) recombinant bacteria. BL21(DE3) recombinant bacteria carrying T7-tel and pNN9/pDMS precursor plasmids were induced for msDNA synthesis by adding 0.4 mM and 0.1mM IPTG. Lanes **1.**BL21(DE3)+T7-tel+pNN9(Induced;0.4 mM);**2.**BL21(DE3)+T7-tel+pNN9(Induced;0.1mM);**3.**BL21(DE3)+T7-tel+pNN9(Un-induced);**4.**BL21(DE3)+T7-tel+ pDMS (Induced;0.4mM); **5.**BL21(DE3)+T7-tel+pDMS (Induced;0.1mM);**6.**BL21(DE3)+T7-tel+pDMS(Un-Induced); **7.**BL21(DE3)+T7-tel+pDMS(Induced;0.4 mM); **8.**BL21(DE3)+T7-tel+ pDMS(Induced;0.1 mM); **9.**BL21(DE3)+T7-tel+pDMS(Un-Induced);**10**.BL21(DE3)+T7-tel+pNN9(Induced;0.4 mM); **11.**BL21(DE3)+T7-tel+pNN9(Induced;0.1 mM);**12.**BL21(DE3)+T7-tel+pNN9(Un-Induced);**13.**BL21(DE3)+T7-tel(Induced);**14.**BL21(DE3)+T7-tel(Induced);**15.**BL21(DE3)+T7-tel(Un-Induced); Arrows indicate bands of interest; red (3.0 kb), orange (2.6 kb), blue (3.3 kb), purple (3.1kb). Mr: 1 kbp ladder (FroggaBio).

Recombinant bacteria	Expected fragment sizes upon msDNA synthesis (kb)	The sizes of the observed fragments (kb)
BL21(DE3)+T7-tel+pNN9 (Induced;0.4 mM)	2.6,3	2.6,3
BL21(DE3)+T7-tel+pNN9 (Induced;0.1 mM)	2.6,3	2.6,3
BL21(DE3)+T7-tel+pNN9 (Un-Induced)	No msDNA synthesis	2.6,3
BL21(DE3)+T7-tel+pDMS (Induced;0.4 mM)	3.1,3.3	3.1,3.3
BL21(DE3)+T7-tel+pDMS (Induced;0.1 mM)	3.1,3.3	3.1,3.3
BL21(DE3)+T7-tel+pDMS (Un-Induced)	No msDNA synthesis	3.1,3.3
BL21(DE3)+T7-tel (Induced)	No msDNA synthesis	No msDNA synthesis
BL21(DE3)+T7-tel (Un-Induced)	No msDNA synthesis	No msDNA synthesis

Table 16:Summary of the expected fragment sizes versus the observed fragment sizes upon IPTG-dependent induction of msDNA synthesis in BL21(DE3) recombinant bacteria.

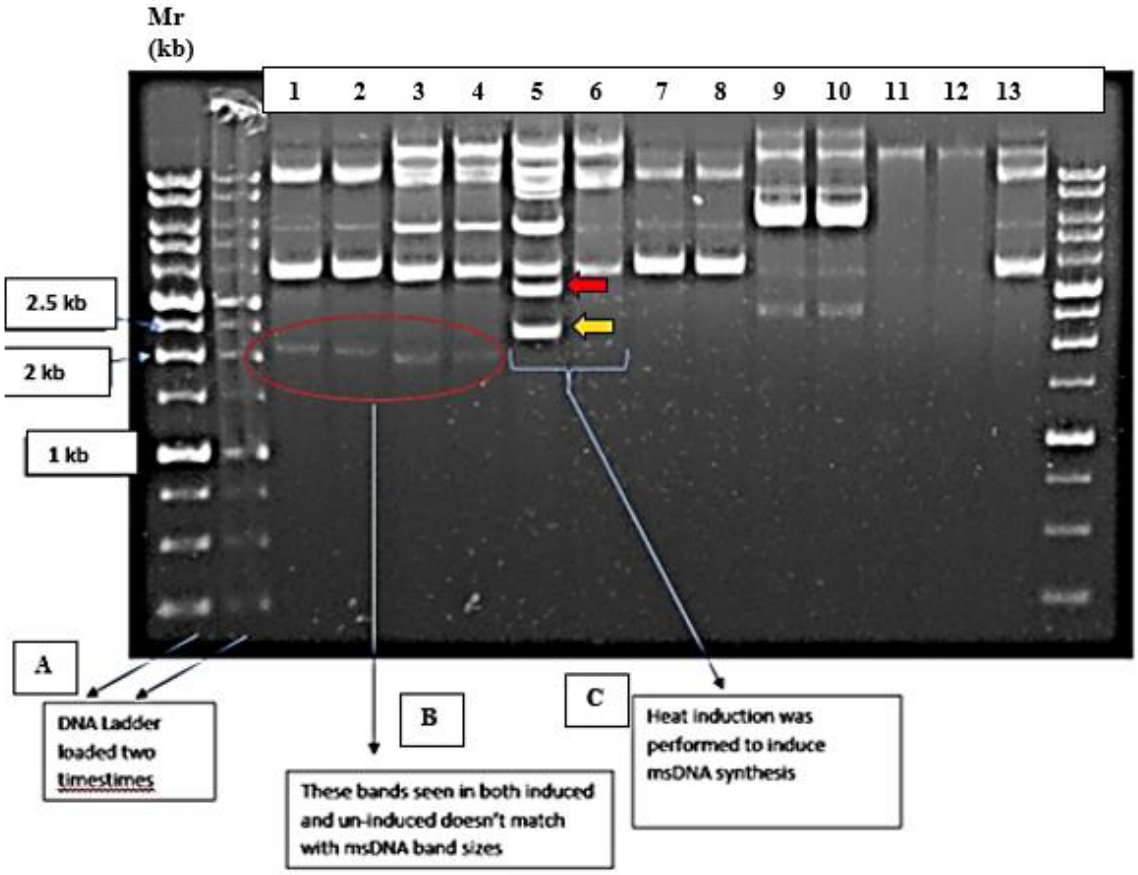


Figure 44:Induction of msDNA synthesis in W3110 recombinant bacteria. msDNA synthesis was induced by adding IPTG at a final concentration of 1 mM. Lanes;**1**.W3110+T7-tel+pDMS(Induced);**2.**W3110+T7-tel+pDMS(Un-Induced); **3.**W3110+T7-tel+pNN9(Induced);**4.** W3110+T7-tel+pNN9(Un-Induced);**5.**W3110 *tel*$^{+}$+pNN9(Induced);**6.**W3110*tel*$^{+}$+pNN9(Un-Induced);**7.**W3110+T7-tel(Induced);**8.**W3110+T7-tel(Un-Induced); **9.**W3110+pACYC+pNN9(Induced);**10.**W3110+pACYC+pNN9(Un-Induced); **11.**W3110(Induced); **12.**W3110(Un-Induced); 13.W3110+pNN9 (Induced); **A.** 1 kb ladder was loaded twice. **B.** The bands highlighted are visible in both induced and induced conditions but don't match with the banding pattern of msDNA synthesis. **C.**W3110 *tel*$^{+}$+pNN9 is the positive control where msDNA synthesis was induced by heat induction. Arrows indicate bands of interest; orange (2.6 kb), red (3.0 kb); Mr: 1 kbp ladder (FroggaBio).

Recombinant bacteria	Expected fragment sizes upon msDNA synthesis (kb)	The sizes of the observed fragments (kb)
W3110+T7-tel+pDMS (Induced)	3.1,3.3	~ 2.2,3.4
W3110+T7-tel+pDMS (Un-Induced)	No msDNA synthesis	~ 2.2,3.4
W3110+T7-tel+pNN9 (Induced)	2.6,3.0	~ 2.2,3.4
W3110+T7-tel+pNN9 (Un-Induced)	No msDNA synthesis	~ 2.2,3.4
W3110 *tel*$^{+}$+pNN9(Induced)	2.6,3.0	2.6,3.0
W3110 *tel*$^{+}$+pNN9 (Un-Induced)	No msDNA synthesis	No msDNA synthesis
W3110+T7-tel (Induced)	No msDNA synthesis	No msDNA synthesis
W3110+T7-tel (Un-Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pACYC+pNN9 (Induced)	No msDNA synthesis	No msDNA synthesis
W3110+pACYC+PNN9 (Un-Induced)	No msDNA synthesis	No msDNA synthesis

Table 17:Summary of the expected fragment sizes versus the observed fragment sizes upon IPTG-dependent induction of msDNA synthesis in W3110 recombinant bacteria.

4.4.4 Induction of msDNA Synthesis in W3110+pXG-pBAD-tel+pNN9/pDMS carrying recombinant bacteria

As described in section 3.4.6 after W3110 bacteria were transformed with pXG-pBAD-tel ***(***pXG-tel) and msDNA synthesizing precursor plasmids pNN9/pDMS, msDNA synthesis was induced by adding L-Arabinose to a final concentration of 1 %. In pXG-tel plasmid *tel* gene is placed under the control of an inducible P_{BAD} promoter. The bacterial cultures were grown in LB media supplemented with 1 % glucose in order to turn off the P_{BAD} promoter expression (due to the leakiness of the P_{BAD} promoter) until the bacterial culture reached the late log phase (A_{600}=0.8). The pXG-tel recombinant plasmid is a gift from Mediphage Bioceuticals, Toronto, and it is 6.9 kb

in size. Upon induction of the P_{BAD} inducible promoter as shown in Figure 45 (A and B) in the presence of both pNN9 and pDMS msDNA synthesizing precursor plasmids, W3110 recombinant bacteria synthesized msDNA. Obtaining band sizes of 2.6 kb and 3.0 kb corresponding to msDNA and bacterial backbone with pNN9 precursor plasmid (Figure 45, A; lanes 1-2) compared to un-induced cultures (Figure 45, A; lanes 3-4) confirms the msDNA synthesis after 6 hr of induction. Similarly, in the presence of pDMS precursor plasmid, W3110 recombinant synthesized msDNA. Obtaining bands of 3.1 kb (blue arrow) and 3.3 kb (purple arrow) as in Figure 45, B, lanes 1-4, corresponding to the bacterial backbone and msDNA respectively confirmed the msDNA synthesis.

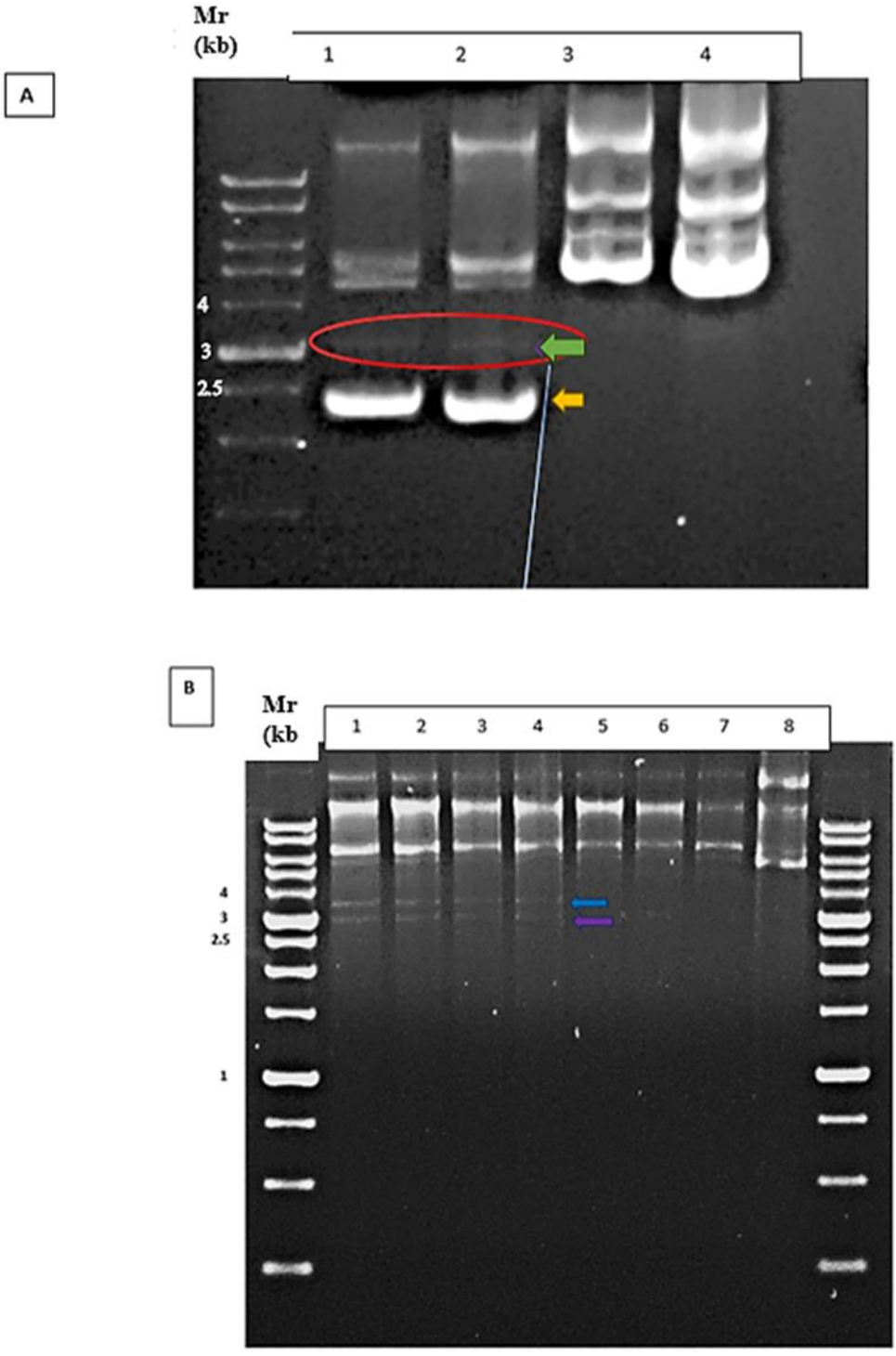


Figure 45: Induction of msDNA synthesis in W3110+pXG-tel+pNN9/pDMS recombinant bacteria. A. Induction of msDNA synthesis with pNN9 precursor plasmid. Lanes;**1 – 2.** W3110+pXG-tel+pNN9 (Induced); **3– 4.** W3110+pXG-tel+pNN9 (Un-Induced). **B.** Induction of msDNA synthesis with pDMS precursor plasmid. Lanes; **1 – 4.** W3110+pXG-tel+pDMS

(Induced); **5-8.** W3110+pXG-tel+pDMS (Un-Induced). Arrows indicate bands of interest; **A.** yellow (2.6 kb), green (3.0 kb); **B.** blue (3.3 kb), purple (3.1kb) Mr: 1 kb ladder (FroggaBio).

Since there was msDNA synthesis observed in W3110 recombinant bacteria, we moved on with transforming W3110 *vma*$^+$ integrated recombinant bacteria with pXG-tel and precursor plasmids as outlined in section 3.4.6.5. Successful chromosomal integration of *vma* gene was confirmed by performing a colony PCR. As shown in Figure 46 (lanes 1-2 + 4-6) obtaining two DNA fragments of 577 ,666 bp confirmed the successful chromosomal integration of *vma* gene expression cassette in W3110 *vma*$^+$ single copy integrants (Table 8). The *vma* gene is placed under the control of an inducible P_{BAD} promoter. As shown in Figure 47 upon induction W3110 *vma*$^+$ recombinant bacteria synthesized msDNA with pNN9 precursor plasmid (in comparison with uninduced cultures in lane 7). Obtaining DNA fragments of 2.6 kb and 3.0 kb as highlighted in lane 5 confirmed the synthesis of msDNA in induced cultures. msDNA synthesis was further confirmed by digesting the plasmid extract with EcoR1 restriction enzyme. As indicated in Table 18 upon EcoR1 digestion of the plasmid extract obtaining 2.9, 3, 3.9,5.6, 1.4, and 1.2 kb DNA fragments confirmed msDNA synthesis (Figure 47; lane 6). As there is no significant size difference between 2.9 & 3 and 1.4 & 1.2 kb DNA fragments it appeared as one thick band on agarose gel. This is the reason why only 4 fragments appeared in lane 6 instead of 6 fragments. But the gel could have been run longer to see better band separation. The uninduced plasmid extract upon digestion (lane 8) showed 3 bands corresponding to pNN9 (5.6 kb) and pXG-tel (2.9 and 3.9 kb) plasmids respectively. In the presence of pDMS precursor plasmid, induced cultures (lane 1) compared to uninduced cultures in lane 3, showed two DNA fragments of 3.1 and 3.3 kb corresponding to the bacterial backbone and msDNA respectively. Upon digestion of the plasmid extract with EcoR1 didn't show 5 fragments (6.4, 2.7, 0.6, 3.9, 2.9, and 3.1 kb) as shown in Table 19. There was no difference

between the banding pattern obtained between induced-digested and un-induced digested (lanes 2 and 4). Overall, 500 ng of DNA were loaded on each well from both induced and un-induced cultures of recombinant bacteria carrying pDMS/pNN9 precursor plasmids. Compared to pNN9 carrying bacterial cultures, pDMS carrying bacterial cultures have lower pDMS plasmid concentrations (lanes 1,3 and 5 and 6). This is because both *tel* and *vma* are placed under the control of P_{BAD} promoter and there is no tight regulation of *vma* gene overexpression. Both *tel* and *vma* gene overexpression takes place at the same time. This resulted in digesting the pDMS plasmid at early stages of bacterial growth. The presence of lower concentrations of pDMS plasmid resulted in the synthesis of msDNA at very low concentrations which are not detectable on an agarose gel.

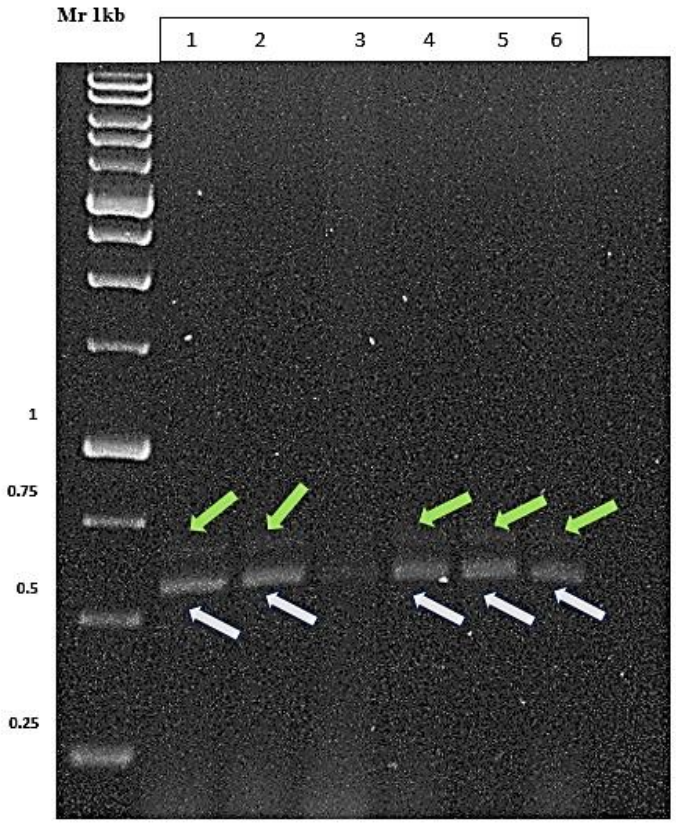


Figure 46:Colony PCR to confirm chromosomal integration of *vma* gene in W3110 *vma*$^+$ single integrants. Lanes:**1-2 and 4-6;** colony PCR on five different colonies of W3110 *vma*$^+$

single integrants grew on Km plates. **lane 3:** N/A. Arrows indicate bands of interest; green (666 bp), grey (577 bp) Mr: 1 kb ladder (FroggaBio).

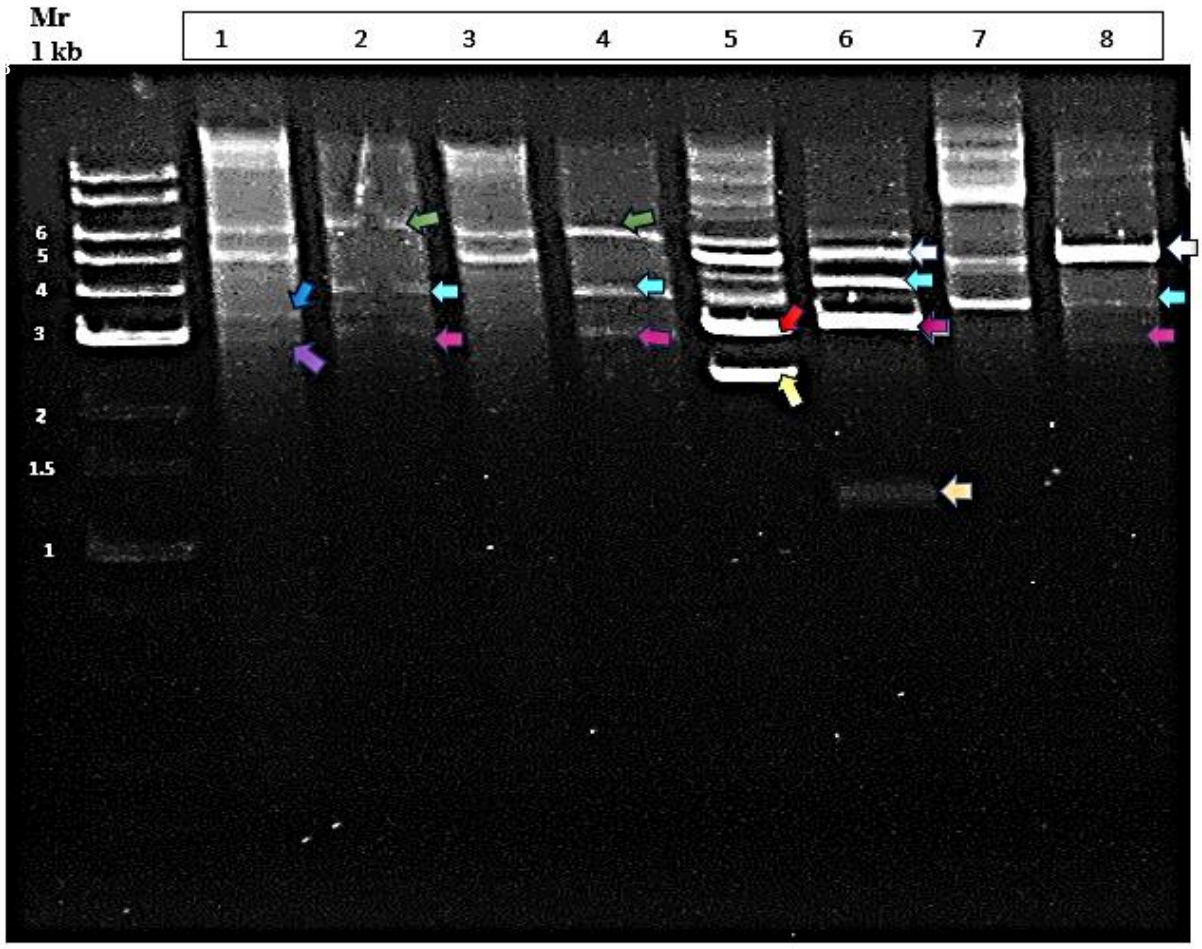


Figure 47:Induction of msDNA synthesis in W3110 vma++pXG-tel+pNN9/pDMS recombinant bacteria. Lanes;**1.** W3110 *vma*$^{+}$+pXG-*tel*+pDMS (Induced); **2.** Lane1 digested with EcoR1; **3.** W3110 *vma*$^{+}$+pXG-*tel*+pDMS (Un-Induced); **4.** Lane3 digested with EcoR1; **5.** W3110 *vma*$^{+}$+pXG-*tel*+pNN9(Induced);**6.** Lane 5 digested with EcoR1;**7.** W3110 *vma*$^{+}$+pXG-*tel*+pNN9(Un-Induced); **8.** Lane 7 digested with EcoR1. Arrows indicate bands of interest; purple;3.1; dark blue;3.3; pink;2.9; blue;3.9; dark green;6.4; yellow;2.6; red;3; orange;1.2,1.4; grey;5.6; Mr; 1 kb plus ladder from FroggaBio.

Plasmid	The size of the DNA fragments when digested with EcoR1
pGX -tel	2.9, 3.9
pNN9	5.6
pDMS	6.4
msDNA with pDMS + BB*	0.7, 2.6, 3.1
msDNA with pNN9 + BB*	1.2,1.4, 3

Table 18:DNA fragment sizes obtained after digesting the plasmid extract with EcoR1.BB* denotes-bacterial backbone.

Recombinant bacteria	Expected fragment sizes upon msDNA synthesis (kb)	The sizes of the observed fragments (kb)	Expected fragment sizes upon EcoR1 digestion (kb)	Observed fragment sizes uponEcoR1 digestion (kb)
W3110 *vma*$^{+}$+pXG-tel+pDMS (Induced)	3.1, 3.3	3.1, 3.3	0.7, 2.6, 3.1, 2.9, 3.9, 6.4	2.9, 3.9, 6.4
W3110 *vma*$^{+}$+pXG-tel+pDMS (Un-Induced)	No msDNA synthesis	No msDNA synthesis	2.9, 3.9, 6.4	2.9, 3.9, 6.4
W3110 *vma*$^{+}$+pXG-tel+pNN9 (Induced)	2.6, 3.0	2.6, 3.0	1.2*, 1.4*, 3.0**, 2.9**, 3.9, 5.6	1.2*, 1.4*, 3.0**, 2.9**, 3.9, 5.6
W3110 *vma*$^{+}$+pXG-tel+pNN9 (un-induced)	No msDNA synthesis	No msDNA synthesis	2.9, 3.9, 5.6	2.9, 3.9, 5.6

Table 19:Summary of the expected fragment sizes versus the observed fragment sizes upon L-arabinose-dependent induction of msDNA synthesis in W3110 *vma*$^{+}$ recombinant bacteria. The * and ** denote the fragments that appeared as single thick bands due to the very low difference in size between the bands.

4.5 Digestion of the msDNA synthesizing precursor plasmid upon over expression of the pre-crRNA targeting the bacterial backbone

Another approach to digest the msDNA-synthesizing precursor plasmid was developed, utilizing the CRISPR-Cas3 RNA-dependent immune system naturally present in *E. coli* K-12 bacteria. In this approach, pre-CRISPR RNA targeting the *ori* and the Ap-resistant gene of the bacterial backbone part of the msDNA-synthesizing precursor plasmid was developed. Initially, we tested the digestion of the msDNA-synthesizing precursor plasmid in W3110 bacteria. We then introduced the pre-CRISPR RNA-expressing plasmid into W3110 *tel*$^{+}$ recombinant bacteria to test msDNA synthesis, followed by CRISPR RNA overexpression to digest the unwanted by-products of msDNA synthesis.

4.5.1 pre-crRNA over expression in W3110 recombinant bacteria

W3110 bacteria were transformed with pcrRNA90/120 and msDNA synthesizing pDMS precursor plasmids as outlined in section 3.4.6.1. As pre-crRNA encoding gene is placed under the control of P_{BAD} promoter L-arabinose was added in order to induce the promoter activation as outlined in section 3.5.6. The purpose of this experiment is to determine whether digestion of the pDMS/pNN9 msDNA synthesizing precursor plasmids is achieved upon over-expression of pre-crRNA. Lanes 5,6,9 and 10 of Figure 48 represent induced and un-induced cultures of W3110+pDMS (negative controls). The pattern of bands in the mentioned lanes illustrates the visualization of the intact pDMS plasmid on an agarose gel. A high intense band within this pattern is emphasized by a yellow circle. Lanes 7,8,11 and 12 of Figure 48 on the agarose gel display the pattern of pcrRNA90/120 plasmids, serving as negative controls that demonstrate the appearance of an intact CRISPR plasmids inside W3110 bacteria (high intense band within this pattern is emphasized by a red circle). As indicated in Figure 48; lane 3, upon pre-crRNA overexpression pDMS precursor

plasmid concentration, in induced cultures of W3110+pcrRNA120+pDMS recombinant bacteria has decreased compared to un-induced cultures (Figure 48; lane 4). Based on densitometric analysis, there is a 71 % decrease in pDMS band intensity in the induced cultures compared to the uninduced cultures (Figure 48; lanes 3 and 4 compared). When contrasted with the negative controls in lanes 4,5 and 6 of Figure 48, it is evident that both induced and uninduced cultures exhibit significantly reduced levels of pDMS plasmid. Based on densitometric analysis there is a 77 % decrease in pDMS plasmid band intensities in induced cultures compared to negative controls. The digestion observed in both induced and un-induced cultures could be due to the leaky overexpression of the pre-crRNA resulting from the leakiness of the P_{BAD} promoter. There was no digestion of pDMS plasmid observed in induced and uninduced cultures of W3110+pcrRNA90+pDMS (Figure 48; lanes 1 and 2).

The digestion of the pDMS plasmid observed in induced and un-induced cultures of W3110+pcrRNA120+pDMS was further confirmed by digesting the plasmid extract with EcoRV restriction enzyme. Since there is only one EcoRV recognition sequence present in both pDMS and pcrRNA120 plasmid, upon EcoRV digestion 6.4 and 5 kb DNA fragments should be visible in an agarose gel. As shown in Figure 49 (lanes 5 and 6) only one band of 5 kb representing pcrRNA120 appeared on the gel. The disappearance of the precursor plasmid further confirmed the successful degradation of pDMS by crRNA120 (Table 20). The complete disappearance of the pDMS plasmid in the induced cultures is due to the fact that partially digested fragments were completely digested by the exonucleases present in bacteria. Since two fragments of 5 and 6.4 kb corresponding to pcrRNA90 and pDMS appeared in lanes 3 and 4 of Figure 49 confirms that upon expression of pre-crRNA90 in W3110+pDMS+pcrRNA90 didn’t digest the pDMS precursor

plasmid. Mutations generated during the construction of the pcrRNA90 plasmid might have resulted in aberrant crRNA expression (Appendix A; Figure A1).

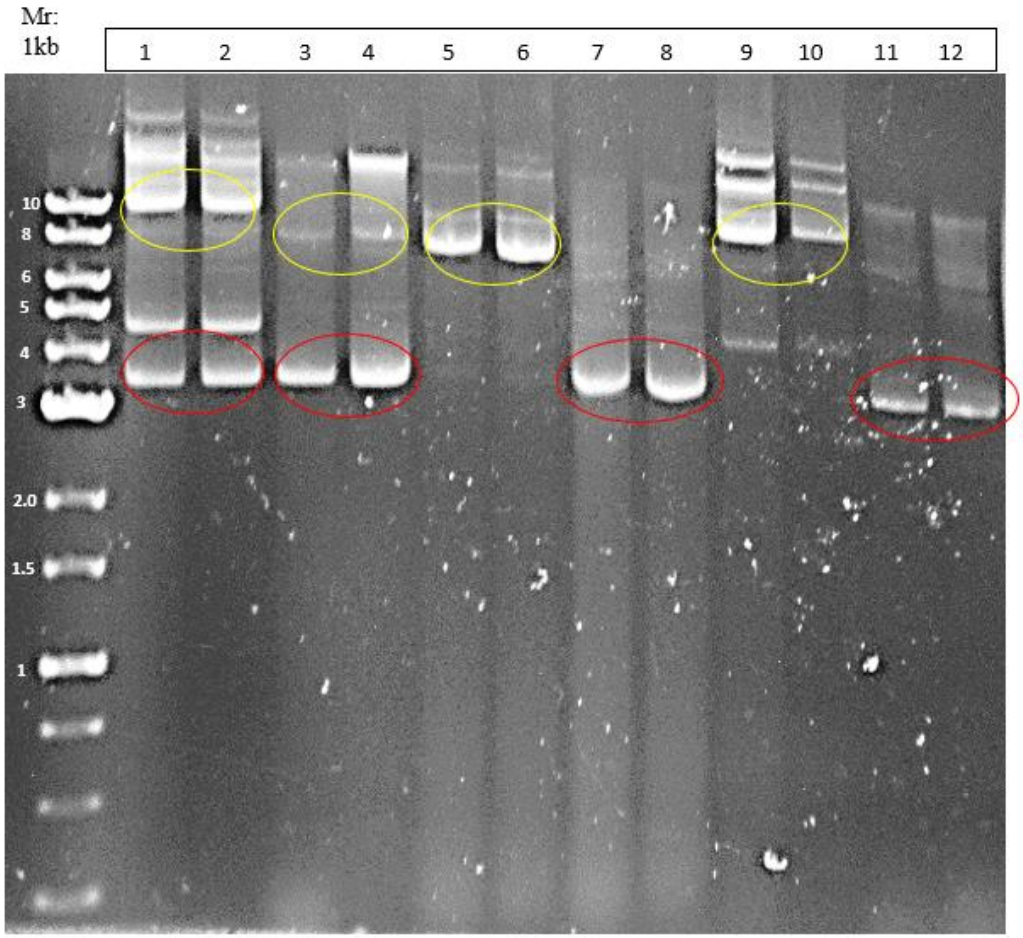


Figure 48:Induction of pre-crRNA over expression. W3110 recombinant bacteria carrying pcrRNA (90/120) and pDMS plasmid was induced by adding L-arabinose into the media. The induction was performed for 6 hr and the plasmid was extracted and ran on 0.8 % agarose gel. Lanes;**1.**W3110+pDMS+pcrRNA90(Induced);**2.**W3110+pDMS+pcrRNA90(Un-Induced);**3.**W3110+pDMS+pcrRNA120(Induced);**4.**W3110+pDMS+pcrRNA120(Un-Induced);**5.**W3110+pDMS(Induced);**6.**W3110+pDMS(Un-Induced);**7**.W3110+pcrRNA120(Induced);**8.**W3110+pcrRNA120(Un-Induced);**9.**W3110+pDMS(Induced);**10.**W3110+pDMS(Un-Induced);**11.**W3110+pcrRNA90(Induced);**12**.W3110+pcrRNA90(Un-Induced).Highlighted circles indicate bands of interest; yellow; pDMS plasmid, red; pcrRNA90/120. Mr;1 kb plus ladder (FroggaBio).

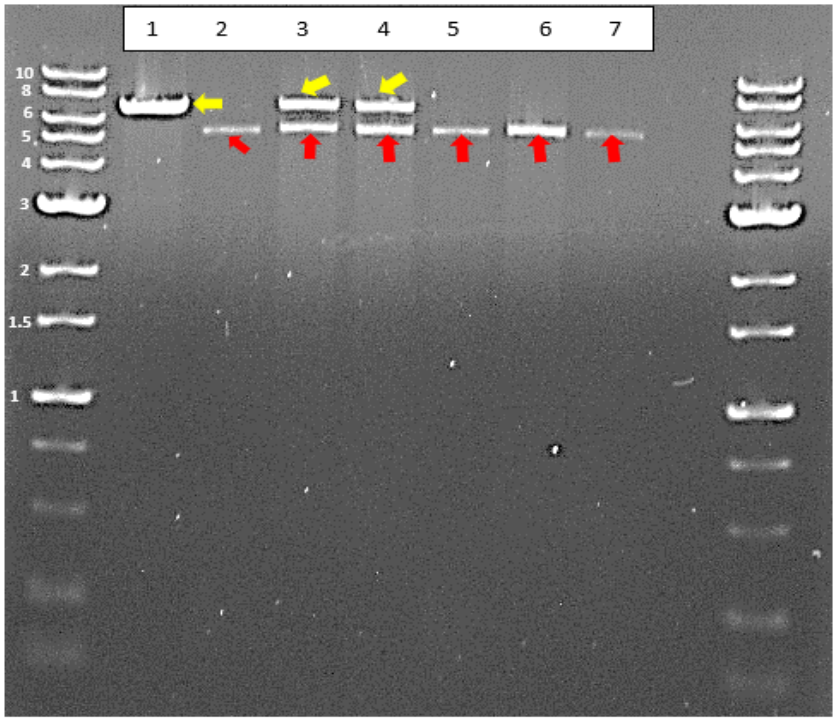


Figure 49:Digestion of the plasmid extract with EcoRV. The plasmid extract isolated after crRNA over expression was digested with EcoRV restriction enzyme and ran on 0.8 % agarose gel. Lanes;**1.** W3110+ pDMS digested with EcoRV;**2.** W3110+pcrRNA90 digested with EcoRV;**3**. W3110+ pcrRNA90+pDMS (Induced) digested with EcoRV;**4.** W3110+ pcrRNA90+pDMS (Un-Induced) digested with EcoRV;**5.** W3110+pcrRNA120+pDMS (Induced)digested with EcoRV;**6.** W3110+pcrRNA120+pDMS (Un-Induced) digested with EcoRV;**7.** W3110+pcrRNA120 digested with EcoRV. Arrows indicate the bands of interest; yellow-6.4 kb; red-5 kb. Mr;1 kb plus ladder (FroggaBio).

Recombinant Bacteria	Expected fragment sizes upon EcoRV digestion (kb)	Observed fragment sizes upon EcoRV digestion (kb)
W3110 +pDMS	6.4	6.4
W3110+pcrRNA90	5	5
W3110+ pcrRNA90+pDMS (Induced)	5	5,6.4
W3110+ pcrRNA90+pDMS (Un-Induced)	5,6.4	5,6.4
W3110+pcrRNA120+pDMS (Induced)	5	5
W3110+pcrRNA120+pDMS (Un-Induced)	5,6.4	5
W3110+pcrRNA120	5	5

Table 20: The expected fragment sizes versus observed fragment sizes upon CRISPR RNA overexpression. The crRNA over expression was induced in W3110 recombinant bacteria carrying pDMS and pcrRNA90/120 plasmids. The plasmid extract was isolated after induction and digested with EcoRV to evaluate test results. W3110+pDMS, W3110+pcrRNA90 and W3110+ pcrRNA120 act as negative controls of this experiment.

4.6 Determination of msDNA synthesis in pre-crRNA synthesizing recombinant bacteria

Since there was digestion observed of the pDMS precursor plasmid upon overexpression of the pre-CRISPR RNA targeting the *ori* of the plasmid, we moved on to transform W3110 *tel*$^+$ recombinant bacteria with the CRISPR plasmid and msDNA-synthesizing precursor plasmids to induce msDNA synthesis. This was followed by the overexpression of pre-CRISPR RNA to digest the by-products of msDNA synthesis.

4.6.1 Expression of msDNA Synthesis in pre-crRNA expressing W3110 *tel*$^+$ recombinant bacteria

Having observed pDMS plasmid degradation during CRISPR120 expression, we proceeded to induce msDNA synthesis in W3110 *tel*$^+$ that harbored both pcrRNA120 and pNN9 plasmids. W3110 *tel*$^+$ recombinant bacteria were transformed with pNN9 and pcrRNA120 plasmids as outlined in section 3.4.6. Sequential transformation was performed to introduce recombinant

plasmids into W3110 *tel*$^{+}$ recombinant bacteria. Since both CRISPR plasmids and W3110 *tel*$^{+}$ possess Cm resistance, the positive transformants, which are W3110 *tel*$^{+}$ carrying pcrRNA120 plasmids, exhibited a diffuse pattern when grown on LB agar plates supplemented with the selective antibiotics. Due to this reason, it became practically impossible to isolate positive colonies that carry both pNN9 and CRISPR plasmids. Therefore, the Cm^{R} gene in the pcrRNA120 plasmid was replaced with the Km^{R} gene. This plasmid construct, expressing the pre-crRNA targeting the *ori* of the precursor plasmid and conferring Km resistance, will be referred to as p120k.

4.6.2 Induction of msDNA synthesis

W3110 *tel*$^{+}$ recombinant bacteria were transformed with pNN9 and p120k plasmids as described in section 3.5.7., followed by heat induction to induce msDNA synthesis. As depicted in Figure 50, upon heat induction W3110 *tel*$^{+}$ recombinant bacteria containing both pNN9 and p120k plasmids synthesized msDNA (Lane 4). The presence of 2.6 and 3.0 kb DNA fragments in induced cultures, in contrast to the uninduced cultures, affirmed the synthesis of msDNA. The 2.6 kb and 3.0 kb DNA fragments on the agarose gel were denoted by blue and green arrows, respectively. W3110 *tel*$^{+}$ with pNN9 plasmid served as a positive control of this experiment (lanes 5 and 6).

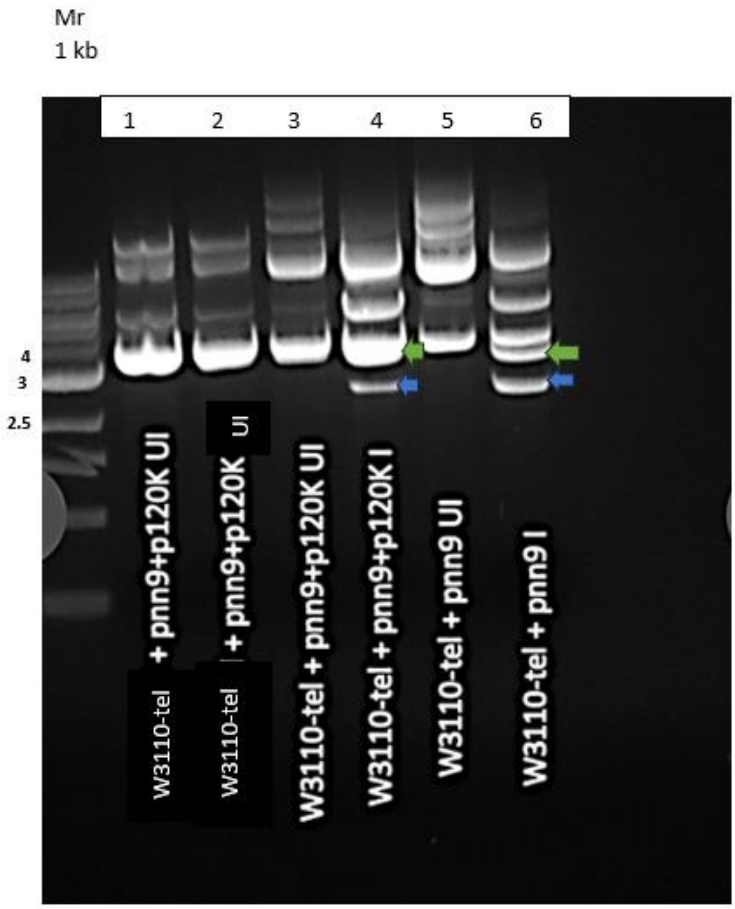


Figure 50:Induction of msDNA synthesis in recombinant bacteria carrying pNN9 and pcrRNA120 plasmids. Heat induction was performed to induce msDNA synthesis. lanes;1-3; W3110 *tel*$^{+}$+pNN9+p120k (Un-Induced); lane 4; W3110 *tel*$^{+}$+pNN9+p120k(Induced); lane 5; W3110 *tel*$^{+}$+pNN9(Un-Induced); lane 6; W3110 *tel*$^{+}$+pNN9(Induced). Arrows indicate the bands of interest. Green (3.0 kb); blue;2.6 kb. Mr 1 kb from FroggaBio.

4.6.3 Induction of msDNA synthesis followed by CRISPR RNA over expression

The p120k overexpression was induced following msDNA synthesis as outlined in section 3.4.6. After msDNA synthesis crRNA over expression was induced for 4hr and 24 hr. As depicted in Figure 51, even though there was msDNA synthesis (Figure 51-lanes 7 and 8), there was not a significant digestion of the pNN9 plasmid observed upon crRNA expression. As shown in lanes

9,11,13 and 15, upon induction of crRNA expression for 4 hr and 24 hr following msDNA synthesis, compared uninduced cultures, induced cultures show a slight degradation of the plasmid. In order to further confirm the degradation, the plasmid extract was digested with EcoR1 restriction enzyme. The digestion reaction contained 500 ng of DNA from induced and uninduced plasmid extracts. Upon digestion of the plasmid extract with EcoR1, the intensity of the DNA fragment corresponding to the pNN9 plasmid remains consistent in both induced and uninduced cultures (Figure 51; lanes 10,12,14 and 16). As summarized in Table 21, upon heat induction and CRISPR RNA overexpression, we should not observe a 5.6 kb band corresponding to pNN9 in induced cultures after digestion of the plasmid extract with EcoR1.The pNN9 plasmid remained intact in W3110 *tel*$^{+}$ cells that possess p120k and pNN9 plasmids, despite observing pDMS plasmid degradation in W3110 cells harboring pcrRNA120 and pDMS plasmids. Because the pcrRNA120 plasmid targets the origin of replication in both pNN9 and pDMS plasmids, it should lead to the digestion of both plasmids.

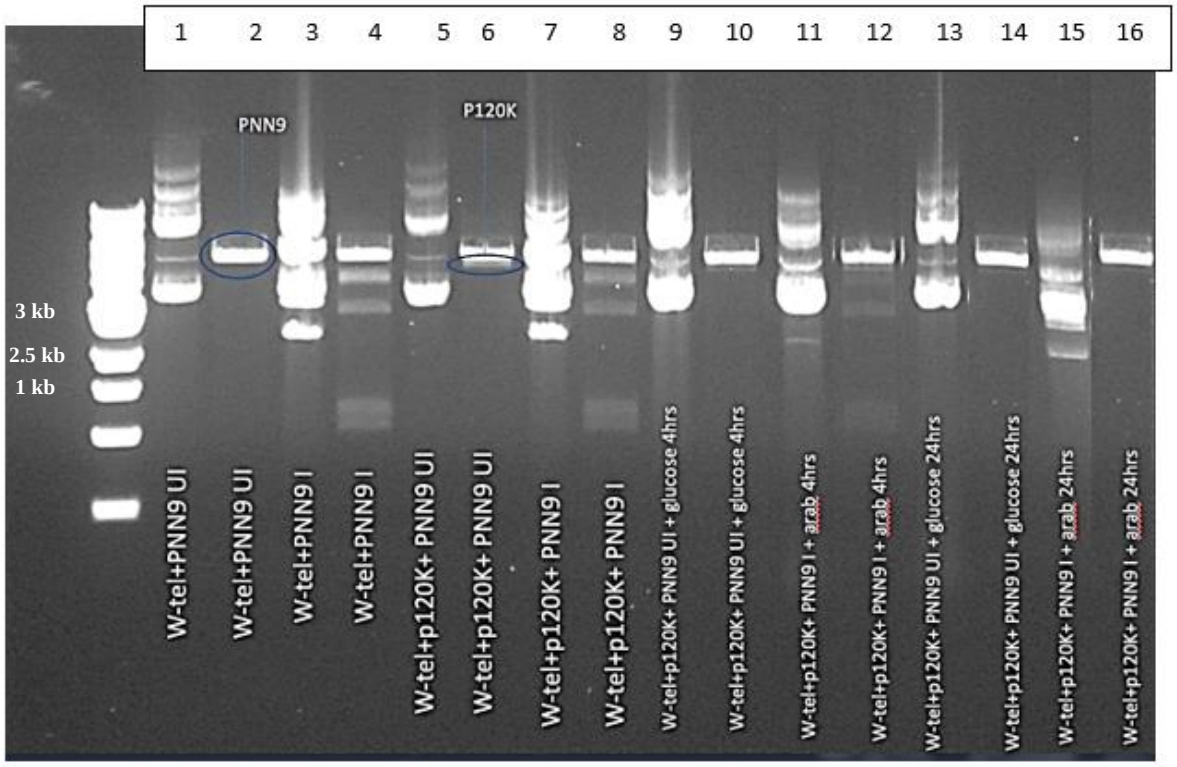


Figure 51:Induction of p120k over expression followed by msDNA synthesis. The pre-crRNA was over expressed for 4 hr and 24 hr after msDNA synthesis. The plasmid extract was digested with EcoR1 to confirm msDNA synthesis and digestion of pNN9 precursor plasmid.Lanes;**1.**W3110 *tel*$^{+}$+pNN9(Un-Induced);**2.**W3110 *tel*$^{+}$+pNN9(Un-Induced)-digested;**3.**W3110 *tel*$^{+}$+pNN9(Induced);**4.**W3110 *tel*$^{+}$+pNN9(Induced)-digested;**5.**W3110 *te*l^{+}+pNN9+pcrRNA120k(Un-Induced);**6.**W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-digested;**7.**W3110 *tel*$^{+}$+pNN9+p120k(Induced);**8.**W3110 *tel*$^{+}$+pNN9+p120k(Induced)-digested;**9.**W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-4hr;**10.**W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-4hr-digested;**11.**W3110 *tel*$^{+}$+pNN9+p120k(Induced)-4hr;**12.**W3110 *tel*$^{+}$+pNN9+p120k(Induced)-4hr-digested;**13.**-W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-24hr;**14.**W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-24hr-digested;**15.**W3110 *tel*$^{+}$+pNN9+p120k(Induced)-24hr;**16.**W3110 *tel*$^{+}$+pNN9+p120k(Induced)-24hr-digested.Mr 1kb from FroggaBio.

Recombinant bacteria	Expected fragment sizes upon EcoR1 digestion (kb)	Observed fragment sizes upon EcoR1 digestion (kb)
W3110 *tel*$^{+}$+pNN9(Un-Induced); heat UI	5.6	5.6
W3110 *tel*$^{+}$+pNN9(Induced); heat Ind	1.2,1.4,3,5.6	1.2,1.4,3,5.6
W3110 *tel*$^{+}$+pNN9+pcrRNA120k(Un-Induced); heat UI	5.6,5.7	5.6,5.7
W3110 *tel*$^{+}$+pNN9+p120k(Induced); heat Ind	1.2,1.4,3,5.6,5.7	1.2,1.4,3,5.6,5.7
W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-4hr; heat+ araC	5.6,5.7	5.6,5.7
W3110 *tel*$^{+}$+pNN9+p120k(Induced)-4hr; heat+ araC	1.2,1.4,5.7	1.2,1.4,3,5.6,5.7
W3110 *tel*$^{+}$+pNN9+p120k(Un-Induced)-24hr; heat+ araC	5.6,5.7	5.6,5.7
W3110 *tel*$^{+}$+pNN9+p120k(Induced)-24hr; heat+ araC	1.2,1.4,5.7	1.2,1.4,3,5.6,5.7

Table 21: The expected fragment sizes versus observed fragment sizes upon msDNA synthesis followed by CRISPR RNA overexpression. The W3110 *tel+* recombinant bacteria carrying the pNN9 precursor plasmid and p120K CRISPR plasmid were first induced for msDNA synthesis through heat induction and then induced for pre-crRNA overexpression to digest the unwanted by-products of msDNA synthesis. The table summarizes the fragment sizes upon digestion of the plasmid extract with the EcoR1 restriction enzyme.

5. Discussion

Gene therapy is a powerful medical approach used to treat or prevent diseases. In order to target the desired cells, vectors are used to efficiently deliver genes of interest. These vectors are broadly categorized into two groups: viral and non-viral vectors. The LCC msDNA employed in this study falls within the non-vector category and is synthesized through a cost-effective *E. coli* based *in vivo* production process. This msDNA therapeutic platform has the potential to be applied towards a wide range of applications and diseases and the requirement for additional scale is growing. In order to move msDNA toward clinical trials, enhanced synthesis and purification methods are required on a larger scale (Talebnia et al., 2023b). Therefore, this research was primarily focused on enhancing the purification of msDNA by utilizing a recombinant system involving the PI-Sce1 homing endonuclease enzyme and the CRISPR-Cas3 system, whereby by-products of msDNA synthesis are digested upon the expression of PI-Sce1 homing endonuclease or pre-crRNA after msDNA synthesis.

The msDNA synthesizing precursor plasmid carries the therapeutic gene expression cassette flanked with Tel protelomerase enzyme recognition super sequences (SS). Upon heat induction, the overexpressed Tel protelomerase acts on the *pal* sequence in the SS and generates LCC msDNA and the LCC bacterial backbone. To isolate the synthesized msDNA, additional steps such as treating the plasmid extract with restriction enzymes to digest the unwanted by products, is performed. Running the plasmid extract through an agarose gel and subsequently conducting a gel purification step to isolate msDNA, result in low yields. These methods are also time-consuming and expensive. Purification using the PI-Sce1 homing endonuclease enzyme, as discussed in the current research, allows for isolation of msDNA directly from the bacterial lysate using commercially available affinity columns. Chen et al., 2005 developed a one-step mini circle

purification method using homing endonuclease I-Sce1. The therapeutic gene expression cassette was cloned into the mini circle synthesizing precursor plasmid. This expression cassette, placed under the control of an inducible P_{BAD} promoter, is flanked with ØC31 integrase enzyme recognition sequences. The ØC31 integrase enzyme encoding gene cassette was placed under the control of the P_{BAD} promoter and was integrated into the same mini circle synthesizing precursor plasmid. Upon induction at 30 ℃, the ØC31 integrase enzyme was overexpressed and targeted the recognition sequences present in the mini circle generating precursor plasmid. This resulted in the generation of unwanted bacterial backbone and mini circle DNA via homologous recombination. By changing the temperature to 37 ℃, homing endonuclease I-Sce1 induction digests the unwanted by products of mini circle DNA synthesis (Chen et al., 2005). Similarly, the PI-Sce1 homing endonuclease enzyme was used in the current research to digest the unwanted by products of msDNA synthesis. The homing endonucleases are highly specific DNA cleaving endoribonucleases that digest internal phosphodiester bonds of DNA. The recognition sequence of the PI-SceI homing endonuclease enzyme is long, which increases specificity. The presence of homing endonuclease enzyme recognition sites in the *E. coli* genome is very rare. Homing endonucleases, unlike restriction enzymes, are less stringent when binding to the recognition sequences. Even a mutation in one base of the recognition site alters binding for restriction enzymes. On the other hand, homing endonucleases are not stringent; even a base change in the recognition site doesn't affect binding with the target sequence. These are the primary reasons for selecting this homing endonuclease enzyme for digestion (Steuer et al., 2004).

The PI-Sce1 homing endonuclease encoding *vma* gene was integrated into the *λattP* site in the bacterial chromosome via site specific recombination (Haldimann & Wanner, 2001). The *vma* gene was placed under the control of an inducible P_{BAD} promoter. As expected, single integrants

(*vma* integrated recombinant bacteria) and double integrants (both *vma* and *tel* integrated recombinant bacteria) were able to overexpress the PI-Sce1 homing endonuclease enzyme and digest the msDNA synthesizing pDMS precursor plasmid upon induction of the *vma* gene over expression. The pDMS precursor plasmid contains the homing endonuclease enzyme recognition sequence cloned into the bacterial backbone part of the plasmid. Optimal enzyme activity was observed at 37 ℃ in the presence of Mg^{2+} (Steuer et al., 2004). Even though the double integrants were able to over express the homing endonuclease enzyme and digest the precursor plasmid, Tel protelomerase activity was not observed. As shown in Figure 32, the location of the two genes are not in close enough proximity to conclude that integration of the *vma* gene disrupts the *tel* gene sequence. Both Tel protelomerase and PI-Sce1 are DNA binding enzymes. They bind to DNA and slide across until they reach the target sequence. Due to leaky over expression of the PI-Sce1 homing endonuclease enzyme, this may have led to binding to the precursor plasmid, consequently halting the Tel protelomerase enzyme from binding to the plasmid. To verify this, double integrants were introduced with the pNN9 msDNA synthesizing precursor plasmid, which does not contain integrated PI-Sce1 homing endonuclease enzyme recognition sequences. Even in the presence of the pNN9 plasmid, double integrants did not synthesize msDNA, indicating the unlikelihood of the suggested possibility mentioned earlier. Double integrants of W3110, BW25113 and JM109 with different stabilities were generated to develop a recombinant system that favors Tel protelomerase enzyme activity. W3110, *E. coli* K-12 bacteria used in the experiment, is RecA+, and this RecA protein is essential for homologous recombination. In the generation of W3110 $tel^{+}vma^{+}$ double integrants, the *tel* gene was integrated into the *lac* operon in the bacterial chromosome via homologous recombination using the pBRINT-cat integrating plasmid, integrating the *tel* gene into the lacZ gene of the bacterial chromosome through RecA-

dependent homologous recombination (Nafissi & Slavcev, 2012). Additionally, the *vma* gene was integrated into *tel*$^+$ recombinant bacteria via site specific recombination using integrating plasmid (Haldimann & Wanner, 2001). Even though the *tel*$^+$*vma*$^+$ double integrants were able to overexpress the PI-Sce1 homing endonuclease enzyme and digest the msDNA synthesizing precursor plasmid, they did not synthesize msDNA. The BW25113 *E. coli* bacteria used in the study have deletion mutations in the *araD* and *araB* genes, which results in the inability to utilize added L-arabinose (Grenier et al., 2014). Compared to W3110 *vma*$^+$ recombinant bacteria, BW25113 *vma*$^+$ recombinant bacteria exhibited enhanced digestion of the pDMS precursor plasmid upon *vma* gene overexpression due to an elevated supply of L-arabinose (Figure 29). In the generation of BW25113 double integrants, the *tel* gene was integrated into the bacterial chromosome via P1 transduction using W3110 *tel*$^+$ as the donor cells of the *tel* gene expression cassette. Since generalized transduction was used to transfer the *tel* gene, and it has been integrated into the *tel*-integrated recombinant bacteria via homologous recombination, the chromosomal locations are very similar to that of W3110 *tel*$^+$. The *vma* gene was integrated into the *tel*-integrated BW25113 *tel*$^+$ bacteria via site-specific recombination, resulting in double integrants with similar locations as W3110 *tel*$^+$*vma*$^+$. The JM109 *E. coli* strain used in the study has mutations in the *endA1*, *recA1*, and *gyrA96* genes at the positions indicated by the numbers. These genes are involved in the expression of enzymes related to DNA repair, recombination, and cleavage of DNA. Therefore, the introduced extrachromosomal DNA is more stable inside JM109 (Stratagene Technical Services, 2003). Both *tel* and *vma* integrated double integrants of JM109 were generated via P1 transduction. Although different methods were used in the generation of double integrants, since the chromosomal locations are the same, none of the recombinant bacteria were able to

synthesize msDNA, even though they were able to overexpress the PI-Sce1 homing endonuclease enzyme.

Another possibility that could explain why the double integrants were not synthesizing msDNA, is potential interaction of the PI-Sce1 enzyme with the Tel protelomerase enzyme. To test this possibility, episomal expression of the Tel enzyme was performed using low copy number plasmids. When the Tel protelomerase enzyme encoding gene is integrated into the bacterial chromosome as a single copy, the expressed Tel protelomerase concentration is not enough to outcompete the PI-Sce1 enzyme. When the *tel* gene is expressed episomally through low copy number plasmids, it produces a higher quantity of enzymes, effectively surpassing the inhibitory effect of the single copy PI-Sce1 enzyme. The single copy integrants of W3110 *vma*$^{+}$ recombinant bacteria were transformed with the pXG-pBAD-tel plasmid, and upon induction, were able to synthesize msDNA in the presence of the pNN9 precursor plasmid. Since both the *tel* gene and *vma* gene are placed under the control of the P_{BAD} promoter, there is no individual regulation of each gene's expression. We observed the digestion of the pDMS precursor plasmid compared to pNN9, which does not contain PI-Sce1 recognition sequences. Although we observed msDNA synthesis in induced cultures compared to uninduced cultures upon heat induction with the pDMS precursor plasmid, we were unable to confirm msDNA synthesis upon digesting the plasmid extract with EcoR1 restriction enzyme. This could have resulted in the appearance of two bands, 0.6 bp and 2.7 bp, corresponding to msDNA upon digestion with the EcoR1 enzyme. This could be due to the fact that early digestion of the pDMS precursor plasmid resulted in synthesizing msDNA at very low concentrations, which were not detectable in an agarose gel. Placing the *tel* gene under the control of different promoters could help in obtaining the tight regulation of msDNA synthesis and subsequent PI-Sce1 enzyme over expression to digest the unwanted by

products. In the Chen et al.,2005 paper, it is mentioned both the ØC31 integrase and I-Sce1 were placed under the control of the P_{BAD} promoter. To overcome the fact that digestion of the precursor plasmid decreases its availability, three strategies were developed. This includes increasing the number of ØC31 integrase gene copies to increase the quantity of minicircle synthesis. Initially, the bacterial culture was grown at 32 ℃ for 2 hr, which is the optimal temperature of the ØC31 integrase enzyme. The temperature was then increased to 37 ℃, which is the optimal temperature of the I-Sce1 enzyme. Changing the pH of the LB broth to 8 optimizes the I-Sce1 enzyme activity. This enhances the digestion of the bacterial backbone after minicircle synthesis. The *tel*/PI-Sce1 system can follow a similar approach to alleviate the issue. This is achieved by increasing the *tel* gene copy numbers in the plasmid to enhance the turnover of msDNA synthesis. Overexpressing PI-Sce1 at 37 ℃, pH 8.5, and in the presence of 2.5 mM $MnCl_2$ optimizes the activity of the PI-Sce1 enzyme (Wende et al., 1996). Another approach to address the issue would be to place the *tel* gene under the control of a different promoter, such as the heat-inducible strong λ *pL-pR* promoters, to regulate gene expression. Meanwhile, we can keep the *vma* gene overexpression under the control of the P_{BAD} promoter. This way, we can achieve more controlled overexpression of the genes.

Another technique employed in this study to digest the precursor plasmid is the CRISPR-Cas 3 system (Type 1 CRISPR Cas system) found within W3110 *E. coli* K-12 bacteria. This RNA dependent adaptive immune system in *E. coli* K-12 bacteria cleaves and degrades the nucleic acids of invading plasmids and viruses. CRISPR RNA is transcribed from pre-CRISPR RNA consisting of target specific spacer sequences flanked with two identical repeat sequences. This spacer sequence is complementary to the protospacer sequence in the target DNA(Gomaa et al., 2014). The 32 bp protospacer sequence is located downstream of the AAG PAM site of the precursor

plasmid. The pre-CRISPR RNA targeting the ampicillin gene and the *ori* of the bacterial backbone of the msDNA synthesizing precursor plasmids pNN9/pDMS, was designed. The pre-CRISPR RNA encoding gene cassette, placed under the control of the P_{BAD} promoter, was cloned into a low copy number plasmid. First, we tested whether the precursor plasmid gets digested upon over expression of the pre-crRNA. W3110 bacteria was transformed with the msDNA synthesizing precursor plasmid and the CRISPR plasmid. As expected, the crRNA targeting the *ori* of the plasmid, over expressed the pre-crRNA upon induction and digested the precursor plasmid. Due to the tendency of the P_{BAD} promoter to leak, digestion was observed in both induced and uninduced cultures. The pre- crRNA targeting the ampicillin gene didn't digest the precursor plasmid upon over expression, due to mutations generated when constructing the crRNA plasmid (Appendix A; Figure A1). As we observed the digestion of the precursor plasmid, we moved on to the subsequent step of heat induction to induce msDNA synthesis. Synthesis of msDNA was observed using the pNN9 precursor plasmid in a crRNA expressing background within W3110 *tel*$^{+}$ recombinant bacteria. Even though there was msDNA synthesis, we did not observe the digestion of the precursor plasmid upon induction of crRNA over expression after msDNA synthesis. Eight CRISPR cas genes are expressed in the CRISPR-Cas 3 system, which are involved in processing of the pre-CRISPR RNA and cleavage and degradation of target nucleic acids (Brouns et al., 2008). Under normal growth conditions, the eight cas genes expressing two operons, are repressed by the histone-like nucleoid structuring (H-NS) system in *E. coli* K-12 bacteria(Pul et al., 2010). Therefore, in Gomaa et al., 2014 the CRISPR genes were expressed episomally using two plasmids which express cas*ABCDE* and *cas1/2/3* genes to obtain optimal activity. The chromosomal CRISPR Cas3 genes can be forcibly expressed by deleting the *hns* gene. These methods have yielded significant CRISPR-Cas3 activity. Although in this study we did not express

the CRISPR-Cas3 genes episomally, by observing the digestion of the msDNA synthesizing precursor plasmid in both induced and uninduced growth conditions, suggested that the CRISPR-Cas3 system is active in plasmids carrying W3110 recombinant bacteria. Testing the digestion of the precursor plasmid upon episomal expression of the CRISPR Cas genes would be a good strategy to determine whether enhanced digestion is observed. The lack of digestion of the msDNA synthesizing precursor plasmid in W3110 *tel*$^{+}$ recombinant bacteria carrying the CRISPR plasmid and the msDNA synthesizing pNN9 precursor plasmid, could be due to several possibilities. This includes potential mutations in the spacer sequence of the CRISPR plasmid and protospacer sequence. Additionally, a loss of function mutation within the CRISPR genes and the partial immunity provided from the CRISPR system could lead to survival of the target plasmid. There is evidence that the CRISPR plasmid can lose its spacer sequence due to recombination between the identical repeat sequences present in the pre-CRISPR RNA encoding gene sequence (Jiang, Maniv, et al., 2013). By sequencing the CRISPR plasmid isolated after msDNA synthesis, can determined whether the spacer sequence in the CRISPR plasmid is being altered due to recombination. Given that the protospacer sequence is present in the *ori* of the msDNA synthesizing precursor plasmid and there was no adverse effect on the precursor plasmid amplification, it rules out the possibility of escaping precursor plasmid degradation from the CRISPR-Cas3 system due to mutations in the protospacer sequence. Another possibility could be attributed to partial immunity that is provided from the CRISPR-Cas3 system. The plasmid coexists with bacteria due to the presence of beneficial genes such as antibiotic resistance genes provided from the precursor plasmid for W3110 *tel*$^{+}$ bacterial cell survival in the presence of antibiotics. Therefore, the bacteria tolerate the presence of the precursor plasmid and bypasses degradation. According to Jiang, Maniv, et al. 2013, the generation of mutations in the CRISPR-

Cas locus results in either mutations or deletion of the spacer sequence targeting the plasmid carrying an antibiotic-resistant gene. As a result, the CRISPR system doesn't recognize and digest the plasmid, which is beneficial for bacterial survival. Another possibility is that the precursor plasmid is able to tolerate partial digestion by CRISPR-Cas. It is possible that precursor plasmids still can survive even after partial digestion. Some plasmids can escape CRISPR immunity and survive inside the recombinant bacteria. This could be due to spontaneous mutations that arise within the CRISPR Cas genes in the genomic DNA of the bacteria (Palmer & Gilmore, 2010).

6. Conclusion and future directions

The main objective of this research project was to enhance msDNA purification using the PI-Sce1 homing endonuclease enzyme system and CRISPR-Cas3 recombinant system. The PI-Sce1 homing endonuclease enzyme encoding the *vma* gene cassette, placed under the control of an inducible P_{BAD} promoter, was successfully integrated into *tel* integrated W3110 and BW25113 recombinant bacteria via site specific recombination. The *vma* integrated recombinant bacteria were able to over express the homing endonuclease enzyme and digest the msDNA synthesizing pDMS precursor plasmid by binding to PI-Sce1 recognizing sequences in the pDMS plasmid. Although both *tel* and *vma* integrated double integrants were able to over express the PI-Sce1 homing endonuclease enzyme and digest the pDMS msDNA synthesizing precursor plasmid, they did not synthesize msDNA. The double integrants were generated with the purpose to synthesize msDNA upon heat induction, followed by L-arabinose addition to induce *vma* gene over expression to digest the unwanted by products of msDNA synthesis, as stated in the hypothesis. Additionally, different approaches were followed to synthesize double integrants, including P1 transduction. However, since the locations of the *tel* and *vma* genes in the chromosome of double integrants are the same, none of the double integrants showed Tel protelomerase enzyme activity.

Therefore, we moved on with expressing the Tel protelomerase enzyme episomally by integrating the *tel* gene expression cassette into a low copy number plasmid. Inside W3110 vma^+ integrated recombinant bacteria carrying the pXG-pBAD-tel plasmid, msDNA was synthesized from the pNN9 msDNA synthesizing precursor plasmid. The induced cultures of W3110 vma^+ + pXG-pBAD-tel + pDMS recombinant bacteria showed msDNA synthesis compared to uninduced

cultures. However, upon digestion with EcoR1, fragments corresponding to msDNA were not observed on a gel. This could be due to the fact that both the *vma* and *tel* genes were placed under the control of the P_{BAD} promoter and no regulation expression timing of the two genes was achieved. Since both were expressed at the same time due to early degradation of the pDMS precursor plasmid, this resulted in msDNA synthesis at a very lower yield which was not detectable in an agarose gel. Since the *vma* gene is integrated into the bacterial chromosome while Tel protelomerase is expressed episomally, this allows for msDNA synthesis by Tel protelomerase and digestion of the unwanted byproducts of msDNA synthesis by homing endonuclease (expressed from the *vma* gene).

The present study lays the groundwork for improving msDNA purification using the PI-Sce1 homing endonuclease enzyme system. However, several avenues for further investigation and optimization exist to bolster the efficiency of the system. Future research should focus on implementing the following improvements to enhance the overall efficiency and applicability of the system. Since both the *vma* gene and the *tel* gene are under the control of the P_{BAD} inducible promoter, there is no regulation of the timing of gene expression for each gene. This can be rectified by placing the *vma* gene under the control of a different promoter, such as heat-inducible strong λ *pL-pR* promoters, which are regulated by the heat-inducible λ repressor, CI857. The optimum temperature of PI-Sce1 homing endonuclease enzyme activity is 37 ℃. Using the heat-inducible promoter facilitates optimal expression at 37 ℃ as well (H. Wang et al., 2022). Since the Tel protelomerase enzyme is active at temperatures less than 37 ℃, having a contrasting optimal activation temperature compared to PI-Sce1, minimizes the digestion of the precursor plasmid in case of leaky overexpression of the *vma* gene at early stages of bacterial growth. In this system, msDNA synthesis always takes place first, followed by *vma* gene overexpression to digest

the unwanted byproducts of msDNA synthesis. This recombinant system requires enhancement, such as introducing PI-Sce1 homing endonuclease enzyme recognition sequences into the Tel protelomerase-encoding plasmid. This way, after msDNA synthesis, both the Tel plasmid and the unwanted by products of msDNA synthesis will get digested, leaving only msDNA, which is easy to purify using chromatographic columns. Another modification that can be implemented is increasing the number of PI-Sce1 homing endonuclease recognition sequences present in both the Tel plasmid and the pDMS precursor plasmid, which would shorten the required PI-Sce1 enzyme induction period. Having more than one copy of the *tel* and *vma* genes and recognition sites of the homing endonuclease enzyme within the plasmids, increases the turnover of msDNA synthesis and unwanted by product digestion. An alternative homing endonuclease that can be used to enhance the system, is I-Sce1, which is known to be more active than PI-Sce1 and has an 18 bp recognition sequence that is not present in *E. coli* chromosomal DNA.

Another method that was employed to digest the unwanted by products of msDNA synthesis was the *E. coli* K-12 CRISPR-Cas3 system that is naturally present in bacteria. The pre-CRISPR RNA expression plasmid, targeting the *ori* of the precursor plasmid, was successfully constructed. The pre-CRISPR RNA expressing gene was placed under the control of an inducible P_{ABD} promoter. Upon induction, the CRISPR plasmid over expressed the pre-CRISPR RNA. The CRISPR cascade protein complex binds to the CRISPR RNA and directs it to the target sequence. Inside W3110 recombinant bacteria, upon CRISPR RNA expression, digestion of the precursor plasmid occurred. When W3110 tel^+ recombinant bacteria were transformed with the msDNA synthesizing precursor plasmid and CRISPR plasmid, msDNA synthesis occurred, however, digestion of the msDNA synthesizing precursor plasmid was not observed. This could be due to disruption of the CRISPR locus inside W3110 tel^+ recombinant bacteria. By episomal expression of the CRISPR cas genes

inside W3110 *tel*$^{+}$ recombinant bacteria along with the precursor plasmid, this might enhance digestion of the precursor plasmid. Further modifications such as mutating the histone-like nucleoid structuring (H-NS) repressor protein expressing genes, which are known to be repressing CRISPR-Cas3 cascade protein gene expression, will improve the degradation of unwanted by-products of msDNA synthesis.

7. References

Al-Dosari, M. S., & Gao, X. (2009). Nonviral gene delivery: Principle, limitations, and recent Progress. In *AAPS Journal* (Vol. 11, Issue 4, pp. 671–681). https://doi.org/10.1208/s12248-009-9143-y

Anguela, X. M., & High, K. A. (2018). Entering the Modern Era of Gene Therapy. *Annu. Rev. Med. 2019, 70*, 273–288. https://doi.org/10.1146/annurev-med-012017

Arabi, F., Mansouri, V., & Ahmadbeigi, N. (2022). Gene therapy clinical trials, where do we go? An overview. *Biomedicine & Pharmacotherapy*, *153*, 113324. https://doi.org/10.1016/j.biopha.2022.113324

Argos, P., Landy1, A., Abremski2, K., Egan3, J. B., Haggard-Ljungquist4, E., Hoess2, R. H., Kahn5', M. L., Kalionis3, B., Narayana7, S. V. L., Pierson, L. S., Sternberg2, N., & Leong1, J. M. (1986). The integrase family of site-specific recombinases: regional similarities and global diversity. In *The EMBO Journal* (Vol. 5, Issue 2).

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, *315*(5819), 1709–1712. https://doi.org/10.1126/science.1138140

Barzel, A., Privman, E., Peeri, M., Naor, A., Shachar, E., Burstein, D., Lazary, R., Gophna, U., Pupko, T., & Kupiec, M. (2011). Native homing endonucleases can target conserved genes in humans and in animal models. *Nucleic Acids Research*, *39*(15), 6646–6659. https://doi.org/10.1093/nar/gkr242

Benjamin, R., Berges, B. K., Solis-Leal, A., Igbinedion, O., Strong, C. L., & Schiller, M. R. (2016). TALEN gene editing takes aim on HIV. In *Human Genetics* (Vol. 135, Issue 9, pp. 1059–1070). Springer Verlag. https://doi.org/10.1007/s00439-016-1678-2

Berling, E., Nicolle, R., Laforêt, P., & Ronzitti, G. (2023). Gene therapy review: Duchenne muscular dystrophy case study. *Revue Neurologique*, *179*(1–2), 90–105. https://doi.org/10.1016/j.neurol.2022.11.005

Bicho D, Queiroz JA, & Tomaz CT. (2015). Influenza Plasmid DNA Vaccines: Progress and Prospects. *Current Gene Therapy*, *15*(6), 541–549. https://doi.org/10.2174/1566523215666150929111048

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, *151*(8), 2551–2561. https://doi.org/10.1099/mic.0.28048-0

Borges, A. L., Davidson, A. R., & Bondy-Denomy, J. (2017). The Discovery, Mechanisms, and Evolutionary Impact of Anti-CRISPRs. *Annual Review of Virology*, *4*(1), 37–59. https://doi.org/10.1146/annurev-virology-101416-041616

Boudes, P. F. (2014). Gene therapy as a new treatment option for inherited monogenic diseases. In *European Journal of Internal Medicine* (Vol. 25, Issue 1, pp. 31–36). https://doi.org/10.1016/j.ejim.2013.09.009

Boyd, D., Weiss, D. S., Chen, J. C., & Beckwith, J. (2000). Towards Single-Copy Gene Expression Systems Making Gene Cloning Physiologically Relevant: Lambda InCh, a Simple Escherichia coli Plasmid-Chromosome Shuttle System. In *JOURNAL OF BACTERIOLOGY* (Vol. 182, Issue 3). http://jb.asm.org/

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & van der Oost, J. (2008a). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, *321*(5891), 960–964. https://doi.org/10.1126/science.1159689

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., Dickman, M. J., Makarova, K. S., Koonin, E. V., & van der Oost, J. (2008b). Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes. *Science*, *321*(5891), 960–964. https://doi.org/10.1126/science.1159689

Brown, C. E., Alizadeh, D., Starr, R., Weng, L., Wagner, J. R., Naranjo, A., Ostberg, J. R., Blanchard, M. S., Kilpatrick, J., Simpson, J., Kurien, A., Priceman, S. J., Wang, X., Harshbarger, T. L., D'Apuzzo, M., Ressler, J. A., Jensen, M. C., Barish, M. E., Chen, M., … Badie, B. (2016). Regression of Glioblastoma after Chimeric Antigen Receptor T-Cell Therapy. *New England Journal of Medicine*, *375*(26), 2561–2569. https://doi.org/10.1056/NEJMoa1610497

Bryant, L. M., Christopher, D. M., Giles, A. R., Hinderer, C., Rodriguez, J. L., Smith, J. B., Traxler, E. A., Tycko, J., Wojno, A. P., & Wilson, J. M. (2013). Lessons Learned from the Clinical Development and Market Authorization of Glybera. *Human Gene Therapy Clinical Development*, *24*(2), 55–64. https://doi.org/10.1089/humc.2013.087

Buchholz, F. (2009). Engineering DNA processing enzymes for the postgenomic era. In *Current Opinion in Biotechnology* (Vol. 20, Issue 4, pp. 383–389). https://doi.org/10.1016/j.copbio.2009.07.005

Calvin, N. M., & Hanawalt, P. C. (1988). High-efficiency transformation of bacterial cells by electroporation. *Journal of Bacteriology*, *170*(6), 2796–2801. https://doi.org/10.1128/jb.170.6.2796-2801.1988

CAMPBELL, A. (1992). Chromosomal Insertion Sites for Phages and Plasmids. *JOURNAL OF BACTERIOLOGY*, *174*(23), 7495–7499.

Carte, J., Wang, R., Li, H., Terns, R. M., & Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes & Development*, *22*(24), 3489–3496. https://doi.org/10.1101/gad.1742908

Chen, Z. Y., He, C. Y., Ehrhardt, A., & Kay, M. A. (2003). Minicircle DNA vectors devoid of bacterial DNA result in persistent and high-level transgene expression in vivo. *Molecular Therapy*, *8*(3), 495–500. https://doi.org/10.1016/S1525-0016(03)00168-0

Chen, Z.-Y., He, C.-Y., & Kay, M. A. (2005). Improved Production and Purification of Minicircle DNA Vector Free of Plasmid Bacterial Sequences and Capable of Persistent Transgene Expression In Vivo. In *HUMAN GENE THERAPY* (Vol. 16). www.liebertpub.com

Chevalier, B. S. (2001). Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research*, *29*(18), 3757–3774. https://doi.org/10.1093/nar/29.18.3757

Cooney, A. L., Abou Alaiwa, M. H., Shah, V. S., Bouzek, D. C., Stroik, M. R., Powers, L. S., Gansemer, N. D., Meyerholz, D. K., Welsh, M. J., Stoltz, D. A., Sinn, P. L., & McCray, P. B. (2016). Lentiviral-mediated phenotypic correction of cystic fibrosis pigs. *JCI Insight*, *1*(14). https://doi.org/10.1172/jci.insight.88730

Darquet, A.-M., Cameron, B., Wils, P., Scherman, D., & Crouzet, J. (1997). A new DNA vehicle for nonviral gene delivery: supercoiled minicircle. In *Gene Therapy* (Vol. 4).

Dong, D., Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., & Huang, Z. (2017). Structural basis of CRISPR–SpyCas9 inhibition by an anti-CRISPR protein. *Nature*, *546*(7658), 436–439. https://doi.org/10.1038/nature22377

Du, F., Liu, Y.-Q., Xu, Y.-S., Li, Z.-J., Wang, Y.-Z., Zhang, Z.-X., & Sun, X.-M. (2021). Regulating the T7 RNA polymerase expression in E. coli BL21 (DE3) to provide more host options for recombinant protein production. *Microbial Cell Factories*, *20*(1), 189. https://doi.org/10.1186/s12934-021-01680-6

Friedmann, T. (1992). *A brief history of gene therapy*. http://www.nature.com/naturegenetics

Gasiunas, G., Barrangou, R., Horvath, P., & Siksnys, V. (2012). Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, *109*(39). https://doi.org/10.1073/pnas.1208507109

Gimble, F. S., Moure, C. M., & Posey, K. L. (2003). Assessing the Plasticity of DNA Target Site Recognition of the PI-SceI Homing Endonuclease Using a Bacterial Two-hybrid Selection System. *Journal of Molecular Biology*, *334*(5), 993–1008. https://doi.org/10.1016/j.jmb.2003.10.013

Gimble, F. S., & Thorner, J. (1992). *Homing of a DNA endonuclease gene by meiotic gene conversion in Saccharomyces cerevisiae*.

Gimbles, F. S., Thorner, J., & Alkek, A. B. (1993). *THE JOURNAL OF BIOLOGICAL CHEMISTRY Purification and Characterization of VDE, a Site-specific Endonuclease from the Yeast Saccharomyces cerevisiae** (Vol. 268, Issue 29).

Ginn, S. L., Alexander, I. E., Edelstein, M. L., Abedi, M. R., & Wixon, J. (2013). Gene therapy clinical trials worldwide to 2012 - an update. In *Journal of Gene Medicine* (Vol. 15, Issue 2, pp. 65–77). https://doi.org/10.1002/jgm.2698

Ginn, S. L., Amaya, A. K., Alexander, I. E., Edelstein, M., & Abedi, M. R. (2018). Gene therapy clinical trials worldwide to 2017: An update. In *Journal of Gene Medicine* (Vol. 20, Issue 5). Blackwell Publishing Inc. https://doi.org/10.1002/jgm.3015

Glick, B. R. (1995). Metabolic load and heterologous gene expression. *Biotechnology Advances*, *13*(2), 247–261. https://doi.org/10.1016/0734-9750(95)00004-A

Glover, D. J., Lipps, H. J., & Jans, D. A. (2005). Towards safe, non-viral therapeutic gene expression in humans. In *Nature Reviews Genetics* (Vol. 6, Issue 4, pp. 299–310). https://doi.org/10.1038/nrg1577

Gomaa, A. A., Klumpe, H. E., Luo, M. L., Selle, K., Barrangou, R., & Beisel, C. L. (2014). Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems. *MBio*, *5*(1). https://doi.org/10.1128/mBio.00928-13

Gonzales, M. F., Brooks, T., Pukatzki, S. U., & Provenzano, D. (2013). Rapid Protocol for Preparation of Electrocompetent Escherichia coli and Vibrio cholerae; *Journal of Visualized Experiments*, *80*. https://doi.org/10.3791/50684

Grenier, F., Matteau, D., Baby, V., & Rodrigue, S. (2014). Complete genome sequence of *Escherichia coli* BW25113. *Genome Announcements*, *2*(5). https://doi.org/10.1128/genomeA.01038-14

Grigoriev, P. S., & Łobocka, M. B. (2001). *Determinants of segregational stability of the linear plasmid-prophage N15 of Escherichia coli*.

Grindl, W., Wende, W., Pingoud, V., & Pingoud, A. (1998). The protein splicing domain of the homing endonuclease PI-SceI is responsible for specific DNA binding. *Nucleic Acids Research*, *26*(8). https://academic.oup.com/nar/article-abstract/26/8/1857/2901936

Haldimann, A., & Wanner, B. L. (2001a). Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *Journal of Bacteriology*, *183*(21), 6384–6393. https://doi.org/10.1128/JB.183.21.6384-6393.2001

Haldimann, A., & Wanner, B. L. (2001b). Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *Journal of Bacteriology*, *183*(21), 6384–6393. https://doi.org/10.1128/JB.183.21.6384-6393.2001

Hardee, C. L., Arévalo-Soliz, L. M., Hornstein, B. D., & Zechiedrich, L. (2017). Advances in non-viral DNA vectors for gene therapy. In *Genes* (Vol. 8, Issue 2). MDPI AG. https://doi.org/10.3390/genes8020065

Hasan, N., Koob, M., & Szybalski, W. (1994). Escherichia coli genome targeting, I. Cre-lox-mediated in vitro generation of ori- plasmids and their *in vivo* chromosomal integration and retrieval. *Elsevier Science*, *150*, 51–56.

Heinrich, J., Schultz, J., Bosse, M., Ziegelin, G., Lanka, E., & Moelling, K. (2002). Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells. *Journal of Molecular Medicine*, *80*(10), 648–654. https://doi.org/10.1007/s00109-002-0362-2

Hertwig, S., Klein, I., Lurz, R., Lanka, E., & Appel, B. (2003). PY54, a linear plasmid prophage of Yersinia enterocolitica with covalently closed ends. In *Molecular Microbiology* (Vol. 48, Issue 4).

Huang, L.-C., Wood, E. A., & Cox, M. M. (1997). Convenient and Reversible Site-Specific Targeting of Exogenous DNA into a Bacterial Chromosome by Use of the FLP Recombinase: the FLIRT System Downloaded from. In *JOURNAL OF BACTERIOLOGY* (Vol. 179, Issue 19). http://jb.asm.org/

Ikeda, H., & Tomizawa, J. (1965). Transducing fragments in generalized transduction by phage P1. *Journal of Molecular Biology*, *14*(1), 85–109. https://doi.org/10.1016/S0022-2836(65)80232-7

Izsvák, Z., Chuah, M. K. L., VandenDriessche, T., & Ivics, Z. (2009). Efficient stable gene transfer into human cells by the Sleeping Beauty transposon vectors. In *Methods* (Vol. 49, Issue 3, pp. 287–297). https://doi.org/10.1016/j.ymeth.2009.07.001

Jeong, H., Kim, H. J., & Lee, S. J. (2015). Complete Genome Sequence of Escherichia coli Strain BL21. *Genome Announcements*, *3*(2). https://doi.org/10.1128/genomeA.00134-15

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology*, *31*(3), 233–239. https://doi.org/10.1038/nbt.2508

Jiang, W., Maniv, I., Arain, F., Wang, Y., Levin, B. R., & Marraffini, L. A. (2013). Dealing with the Evolutionary Downside of CRISPR Immunity: Bacteria and Beneficial Plasmids. *PLoS Genetics*, *9*(9), e1003844. https://doi.org/10.1371/journal.pgen.1003844

Klinman, D. M., Ylt, A.-K., Beaucaget, S. L., Conover, J., & Kriegt, A. M. (1996). CpG motifs present in bacterial DNA rapidly induce lymphocytes to secrete interleukin 6, interleukin 12, and interferon. In *Immunology* (Vol. 93).

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., & Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, *533*, 420–424. https://doi.org/10.1038/nature17946

Landy, A., & Ross, W. (1977). Viral Integration and Excision: Structure of the Lambda att Sites. In *Source: Science, New Series* (Vol. 197, Issue 4309).

Lesch, H., Pikkarainen, J., Kaikkonen Minna, Taavitsainen Miia, Samaranayake Haritha, Lehtolainen-Dalkilic Pauliina, Vuorio Taina, Määttä Ann-Marie, Wirth Thomas, Airenne Kari, & Ylä-Herttuala, S. (2009). Avidin Fusion Protein-Expressing Lentiviral Vector for Targeted Drug Delivery. *20*.

Lieber, M. R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annual Review of Biochemistry*, *79*(1), 181–211. https://doi.org/10.1146/annurev.biochem.052308.093131

Liu, F., & Huang, L. (2002). Development of non-viral vectors for systemic gene delivery. In *Journal of Controlled Release* (Vol. 78). www.elsevier.com/locate/jconrel

Loring, H., ElMallah, M., & Flotte, T. (2016). Development of rAAV2-CFTR: History of the First rAAV Vector Product to be Used in Humans. *Human Gene Therapy Methods*, *27*.

Lukashev, A. N., & Zamyatnin, A. A. (2016). Viral vectors for gene therapy: Current state and clinical perspectives. In *Biochemistry (Moscow)* (Vol. 81, Issue 7, pp. 700–708). Maik Nauka Publishing / Springer SBM. https://doi.org/10.1134/S0006297916070063

Maguire, A. M., Russell, S., Wellman, J. A., Chung, D. C., Yu, Z.-F., Tillman, A., Wittes, J., Pappas, J., Elci, O., Marshall, K. A., McCague, S., Reichert, H., Davis, M., Simonelli, F., Leroy, B. P., Wright, J. F., High, K. A., & Bennett, J. (2019). Efficacy, Safety, and Durability of Voretigene Neparvovec-rzyl in RPE65 Mutation–Associated Inherited Retinal Dystrophy. *Ophthalmology*, *126*(9), 1273–1285. https://doi.org/10.1016/j.ophtha.2019.06.017

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J. M., Wolf, Y. I., Yakunin, A. F., van der Oost, J., & Koonin, E. V. (2011). Evolution and classification of the CRISPR–Cas systems. *Nature Reviews Microbiology*, *9*(6), 467–477. https://doi.org/10.1038/nrmicro2577

Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., … Koonin, E. V. (2015). An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology*, *13*(11), 722–736. https://doi.org/10.1038/nrmicro3569

Manno, C. S., Pierce, G. F., Arruda, V. R., Glader, B., Ragni, M., Rasko, J. J. E., Ozelo, M. C., Hoots, K., Blatt, P., Konkle, B., Dake, M., Kaye, R., Razavi, M., Zajko, A., Zehnder, J., Rustagi, P., Nakai, H., Chew, A., Leonard, D., … Kay, M. A. (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Medicine*, *12*(3), 342–347. https://doi.org/10.1038/nm1358

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science*, *322*(5909), 1843–1845. https://doi.org/10.1126/science.1165771

Martinez-Morales, F., Borges, A. C., Martinez, A., Shanmugam, K. T., & Ingram, L. O. (1999). Chromosomal Integration of Heterologous DNA in Escherichia coli with Precise Removal of Markers and Replicons Used during Construction Downloaded from. In *JOURNAL OF BACTERIOLOGY* (Vol. 181, Issue 22). http://jb.asm.org/

Metcalf, W., Jiang, W., & Wanner, B. (1994). Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6Ky origin plasmids at different copy numbers. *Elsevier Science B*, *138*, 1–7.

Miller, A. M., & Dean, D. A. (2009). Tissue-specific and transcription factor-mediated nuclear entry of DNA. In *Advanced Drug Delivery Reviews* (Vol. 61, Issues 7–8, pp. 603–613). https://doi.org/10.1016/j.addr.2009.02.008

Murugan, K., Babu, K., Sundaresan, R., Rajan, R., & Sashital, D. G. (2017). The Revolution Continues: Newly Discovered Systems Expand the CRISPR-Cas Toolkit. *Molecular Cell*, *68*(1), 15–25. https://doi.org/10.1016/j.molcel.2017.09.007

Nafissi, N., Alqawlaq, S., Lee, E. A., Foldvari, M., Spagnuolo, P. A., & Slavcev, R. A. (2014). DNA ministrings: Highly safe and effective gene delivery vectors. *Molecular Therapy - Nucleic Acids*, *3*. https://doi.org/10.1038/mtna.2014.16

Nafissi, N., & Slavcev, R. (2012). Construction and Characterization of an in-vivo Linear Covalently Closed DNA Vector Production System. *Microbial Cell Factories*, *11*. https://doi.org/10.1186/1475-2859-11-154

Nafissi, N., Sum, C. H., Wettig, S., & Slavcev, R. A. (2014). Optimization of a one-step heat-inducible in vivo mini DNA vector production system. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0089345

Norrander, J., Kempe, T., & Messing, J. (1983). Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene*, *26*(1), 101–106. https://doi.org/10.1016/0378-1119(83)90040-9

Oldfield, E. H., Culver, K. W., Blaese, R. M., DeVroom, H. L., & Anderson, W. F. (1993). Gene therapy for the treatment of brain tumors using intra-tumoral transduction with the thymidine kinase gene and intravenous ganciclovir. *Hum Gene Ther . , 4*(1), 39–69. https://doi.org/DOI: 10.1089/hum.1993.4.1-39

Palmer, K. L., & Gilmore, M. S. (2010). Multidrug-Resistant Enterococci Lack CRISPR- *cas*. *MBio*, *1*(4). https://doi.org/10.1128/mBio.00227-10

Parato, K. A., Senger, D., Forsyth, P. A. J., & Bell, J. C. (2005). Recent progress in the battle between oncolytic viruses and tumours. In *Nature Reviews Cancer* (Vol. 5, Issue 12, pp. 965–976). https://doi.org/10.1038/nrc1750

Perler, F. B., Davis1, E. O., Dean2, G. E., Gimble3, F. S., Jack, W. E., Neff4, N., Noren, C. J., Thorner5, J., & Belfort6, M. (1994). Protein splicing elements: inteins and exteins a definition of terms and recommended nomenclature. In *Nucleic Acids Research* (Vol. 22, Issue 7).

Ponder, K. P. (2003). Vectors in Gene Therapy. In *An Introduction to Molecular Medicine and Gene Therapy* (pp. 77–112). John Wiley & Sons, Inc. https://doi.org/10.1002/0471223875.ch4

Pul, Ü., Wurm, R., Arslan, Z., Geißen, R., Hofmann, N., & Wagner, R. (2010). Identification and characterization of E. coli CRISPR- cas promoters and their silencing by H-NS. *Molecular Microbiology*, *75*(6), 1495–1512. https://doi.org/10.1111/j.1365-2958.2010.07073.x

Ramamoorth, M., & Narvekar, A. (2015). Non viral vectors in gene therapy - An overview. In *Journal of Clinical and Diagnostic Research* (Vol. 9, Issue 1, pp. GE01–GE06). Journal of Clinical and Diagnostic Research. https://doi.org/10.7860/JCDR/2015/10443.5394

Rao, R. C., & Zacks, D. N. (2014). Cell and gene therapy. *Developments in Ophthalmology*, *53*, 167–177. https://doi.org/10.1159/000357376

Raper, S. E., Chirmule, N., Lee, F. S., Wivel, N. A., Bagg, A., Gao, G. P., Wilson, J. M., & Batshaw, M. L. (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Molecular Genetics and Metabolism*, *80*(1–2), 148–158. https://doi.org/10.1016/j.ymgme.2003.08.016

Rehman, H., Silk, A. W., Kane, M. P., & Kaufman, H. L. (2016). Into the clinic: Talimogene laherparepvec (T-VEC), a first-in-class intratumoral oncolytic viral therapy. *Journal for ImmunoTherapy of Cancer*, *4*(1), 53. https://doi.org/10.1186/s40425-016-0158-5

Rodríguez, E. G. (2004). Nonviral DNA vectors for immunization and therapy: Design and methods for their obtention. In *Journal of Molecular Medicine* (Vol. 82, Issue 8, pp. 500–509). https://doi.org/10.1007/s00109-004-0548-x

Rybchin, V. N., & Svarchevsky, A. N. (1999). *The plasmid prophage N15: a linear DNA with covalently closed ends*.

Salima, H.-B.-A., Kalle Christof, von, Manfred, S., Deist Françoise, L., Nicolas, W., Elisabeth, M., Isabelle, R., Jean-Luc, V., Christopher, F. C., Marina, C.-C., & Alain, F. (2003). *A Serious Adverse Event after Successful Gene Therapy for X-Linked Severe Combined Immunodeficiency*. www.nejm.org

Samson, S. L., Gonzalez, E. V, Yechoor, V., Bajaj, M., Oka, K., & Chan, L. (2008). Gene Therapy for Diabetes: Metabolic Effects of Helper-dependent Adenoviral Exendin 4 Expression in a Diet-induced Obesity Mouse Model. *Molecular Therapy*, *16*(11), 1805–1812. https://doi.org/10.1038/mt.2008.198

Saragliadis, A., Trunk, T., & Leo, J. C. (2018a). Producing Gene Deletions in *Escherichia coli* by P1 Transduction with Excisable Antibiotic Resistance Cassettes. *Journal of Visualized Experiments*, *2018*(139). https://doi.org/10.3791/58267

Saragliadis, A., Trunk, T., & Leo, J. C. (2018b). Producing Gene Deletions in *Escherichia coli* by P1 Transduction with Excisable Antibiotic Resistance Cassettes. *Journal of Visualized Experiments*, *139*. https://doi.org/10.3791/58267

Schakowski, F., Gorschlüter, M., Junghans, C., Schroff, M., Buttgereit, P., Ziske, C., Schöttker, B., König-Merediz, S. A., Sauerbruch, T., Wittig, B., & Schmidt-Wolf, I. G. H. (2001). A novel minimal-size vector (MIDGE) improves transgene expression in colon carcinoma

cells and avoids transfection of undesired DNA. *Molecular Therapy*, *3*(5 I), 793–800. https://doi.org/10.1006/mthe.2001.0322

Shiriaeva, A. A., Kuznedelov, K., Fedorov, I., Musharova, O., Khvostikov, T., Tsoy, Y., Kurilovich, E., Smith, G. R., Semenova, E., & Severinov, K. (2022). Host nucleases generate prespacers for primed adaptation in the *E. coli* type I-E CRISPR-Cas system. *Science Advances*, *8*(47). https://doi.org/10.1126/sciadv.abn8650

SOB Medium. (2018). *Cold Spring Harbor Protocols*, *2018*(3), pdb.rec102723. https://doi.org/10.1101/pdb.rec102723

Spies, B., Hochrein, H., Vabulas, M., Huster, K., Busch, D. H., Schmitz, F., Heit, A., & Wagner, H. (2003). Vaccination with Plasmid DNA Activates Dendritic Cells via Toll-Like Receptor 9 (TLR9) but Functions in TLR9-Deficient Mice. *The Journal of Immunology*, *171*(11), 5908–5912. https://doi.org/10.4049/jimmunol.171.11.5908

Steuer, S., Pingoud, V., Pingoud, A., & Wende, W. (2004). Chimeras of the Homing Endonuclease PI-SceI and the Homologous *Candida tropicalis* Intein: A Study to Explore the Possibility of Exchanging DNA-Binding Modules to Obtain Highly Specific Endonucleases with Altered Specificity. *ChemBioChem*, *5*(2), 206–213. https://doi.org/10.1002/cbic.200300718

Stratagene Technical Services. (2003). *Manual: JM109 Competent Cell*.

Sum, C. H., Chong, J. Y., Wettig, S., & Slavcev, R. A. (2014). Separation and purification of linear covalently closed deoxyribonucleic acid by Q-anion exchange membrane chromatography. *Journal of Chromatography A*, *1339*, 214–218. https://doi.org/10.1016/j.chroma.2014.03.016

Talebnia, F., Pushparajah, D., Chandrasekaran, S., Hersch, S. J., Nafissi, N., & Slavcev, R. (2023a). Application of an electro elution system for direct purification of linear covalently closed DNA fragments. *Journal of Chromatography B*, *1218*, 123622. https://doi.org/10.1016/j.jchromb.2023.123622

Talebnia, F., Pushparajah, D., Chandrasekaran, S., Hersch, S. J., Nafissi, N., & Slavcev, R. (2023b). Application of an electro elution system for direct purification of linear covalently closed DNA fragments. *Journal of Chromatography B*, *1218*, 123622. https://doi.org/10.1016/J.JCHROMB.2023.123622

van Houte, S., Ekroth, A. K. E., Broniewski, J. M., Chabas, H., Ashby, B., Bondy-Denomy, J., Gandon, S., Boots, M., Paterson, S., Buckling, A., & Westra, E. R. (2016). The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature*, *532*(7599), 385–388. https://doi.org/10.1038/nature17436

Wang, H. S., Chen, Z. J., Zhang, G., Ou, X. L., Yang, X. L., Wong, C. K. C., Giesy, J. P., Du, J., & Chen, S. Y. (2012). A Novel Micro-Linear Vector for In Vitro and In Vivo Gene Delivery and Its Application for EBV Positive Tumors. *PLoS ONE*, *7*(10). https://doi.org/10.1371/journal.pone.0047159

Wende, W., Grindl, W., Christ, F., Pingoud, A., & Pingoud, V. (1996). Binding, bending and cleavage of DNA substrates by the homing endonuclease PI-SceI. In *Nucleic Acids Research* (Vol. 24, Issue 21).

Wolfgang Wende, W. G. F. C. A. P. V. P. (1996). Binding, bending and cleavage of DNA substrates by the homing endonuclease Pl-SceI. *Nucleic Acids Research*, *24*(21), 4123–4132. https://doi.org/10.1093/nar/24.21.4123

Wong, S., Lam, P., Nafissi, N., Denniss, S., & Slavcev, R. (2016). Production of double-stranded DNA ministrings. *Journal of Visualized Experiments*, *2016*(108). https://doi.org/10.3791/53177

Xu, Z.-L., Mizuguchi, H., Ishii-Watabe, A., Uchida, E., Mayumi, T., & Hayakawa, T. (2001). *Optimization of transcriptional regulatory elements for constructing plasmid vectors*. www.elsevier.com/locate/gene

Xue, C., & Sashital, D. G. (2019a). Mechanisms of Type I-E and I-F CRISPR-Cas Systems in *Enterobacteriaceae*. *EcoSal Plus*, *8*(2). https://doi.org/10.1128/ecosalplus.ESP-0008-2018

Xue, C., & Sashital, D. G. (2019b). Mechanisms of Type I-E and I-F CRISPR-Cas Systems in *Enterobacteriaceae*. *EcoSal Plus*, *8*(2). https://doi.org/10.1128/ecosalplus.ESP-0008-2018

Yin, H., Kanasty, R. L., Eltoukhy, A. A., Vegas, A. J., Dorkin, J. R., & Anderson, D. G. (2014). Non-viral vectors for gene-based therapy. In *Nature Reviews Genetics* (Vol. 15, Issue 8, pp. 541–555). Nature Publishing Group. https://doi.org/10.1038/nrg3763

Zucca, S., Pasotti, L., Politi, N., Gabriella, M., De Angelis, C., & Magni, P. (2013). *A standard vector for the chromosomal integration and characterization of BioBrick™ parts in Escherichia coli*. http://www.jbioleng.org/content/7/1/12

Appendices

Appendix A

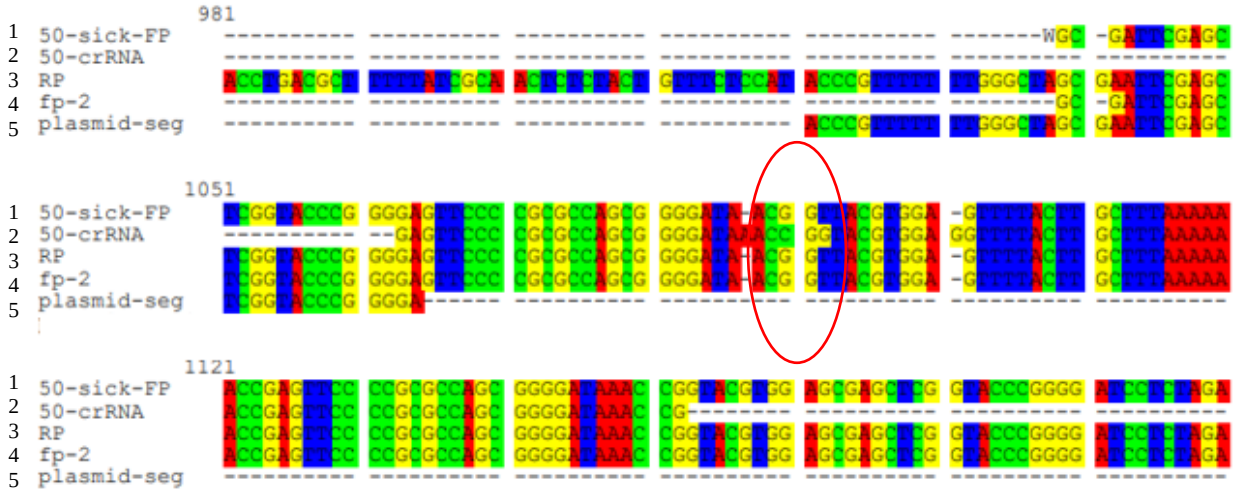


Figure A.1: Confirmation of the cloning of pre-crRNA (targeting the ampicillin resistant gene of the msDNA synthesizing precursor plasmid) encoding crRNA90 gene sequence. The presence of the 90 bp pre-crRNA encoding gene sequence targeting the ampicillin-resistant gene of the msDNA-synthesizing precursor plasmid was confirmed through sequencing. Alignment of Sanger sequencing data obtained with the expected crRNA90 target sequence was performed using Seaview online sequence alignment software. The following sequences were aligned to confirm the presence of the intact crRNA90 sequence:**1.** Obtained Sanger sequence data;**2.** crRNA90 target sequence;**3.** Obtained Sanger sequence data with RP;**4.** Obtained Sanger sequence data with FP2;**5.** Gene sequence of the plasmid upstream of the crRNA90 target sequence. The circled region indicates the locations of mutations in the crRNA90 target sequence of the CRISPR plasmid.