T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

COMPARISON OF SCREENING METHODS FOR CRISPR/CAS9-INDUCED MUTATIONS

ASENA CANTÜRK

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

GEBZE

2019

T.R.

GEBZE TECHNICAL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

COMPARISON OF SCREENING METHODS FOR CRISPR/CAS9-INDUCED MUTATIONS

ASENA CANTÜRK

A THESIS SUBMITTED FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

THESIS SUPERVISOR

ASSOC. PROF. DR. NURİ ÖZTÜRK

GEBZE

2019

T.C.

GEBZE TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

CRISPR/CAS9-İNDÜKLENMİŞ MUTASYONLAR İÇİN GÖRÜNTÜLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

ASENA CANTÜRK

YÜKSEK LİSANS TEZİ

MOLEKÜLER BİYOLOJİ VE GENETİK ANABİLİM DALI

DANIŞMANI

DOÇ. DR. NURİ ÖZTÜRK

GEBZE

2019

GEBZE TEKNİK ÜNİVERSİTESİ	YÜKSEK LİSANS JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 29/01/2019 tarih ve 2019/07 sayılı kararıyla oluşturulan jüri tarafından 19/02/2019 tarihinde tez savunma sınavı yapılan Asena CANTÜRK' ün tez çalışması Moleküler Biyoloji ve Genetik Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : DOÇ. DR. Nuri ÖZTÜRK

ÜYE : PROF. DR. Tamer YAĞCI

ÜYE : PROF. DR. Halil KAVAKLI

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

...../...../..... tarih ve ...../..... sayılı kararı.

SUMMARY

CRISPR/Cas9-induced mutations vary from point mutation to large insertion/deletions. It is difficult to detect the type of mutation after applying CRISPR/Cas9 system for gene knockout purpose because after formation of double-strand breaks, the repair system does not use a template for recovering the lesions and induces a random type of mutations. Basically, the available methods are based on MutS protein which has an affinity for mismatches, T7 Endonuclease Cleavage Assay or High-Resolution Melting Analysis. The aim of this thesis is to compare and/or optimize the available mutation screening methods for CRISPR/Cas9-induced mutations and to test a new method for mutation screening.

Genomic DNA from *BMAL1* knockout MCF10A cell line and a plasmid containing a point mutation in *Drosophila* Cryptochrome coding sequence alongside with their wild type partners were used for mutation detection analysis. The Cas9 protein can be reengineered as a split enzyme, providing a platform for controlled use of Cas9 for genome-engineering applications in the cells. In this study, by utilizing from the research in the CRISPR/Cas9 editing area, MutS was replaced with PAM-interacting domain and reengineered MutS-Cas9 protein was generated. In addition to a controlled comparison of assays was performed for detecting CRISPR/Cas9-induced mutations, as a novel method, a chimeric enzyme for detection of mutations was tried to be generated. Reengineered MutS-Cas9 enzyme presents an alternative to available mutation detection methods and it gives a point of view in the current Cas9 reengineering studies.

Key Words: CRISPR/Cas9, mutation screening, MutS, T7 Endonuclease Cleavage Assay, High-Resolution Melting Analysis (HRMA).

ÖZET

CRISPR/Cas9 ile indüklenen mutasyonlar nokta mutasyonundan büyük ekleme/delesyonlara kadar değişiklik göstermektedir. CRISPR/Cas9 sistemini gen susturma amacıyla uyguladıktan sonra, oluşan mutasyon türünü tespit etmek zordur; çünkü çift iplik kırıklarının oluşumundan sonra, onarım sistemi lezyonları kurtarmak için bir kalıp kullanmaz ve rastgele mutasyonlar meydana getirir. Temel olarak, mevcut yöntemler, yanlış eşleşmeler için afiniteye sahip olan MutS proteinine, T7 Endonükleaz Kesim Analizi'ne ya da Yüksek Çözünürlüklü Erime Analizi'ne dayanmaktadır. Bu tezin amacı, CRISPR/Cas9-indüklenmiş mutasyonlar için mevcut mutasyon tarama yöntemlerini karşılaştırmak ve/veya optimize etmek ve mutasyon taraması için yeni bir yöntem test etmektir.

BMAL1 knockout MCF10A hücre hattı genomik DNA'sı ve kodladığı *Drosophila* Kriptokrom dizisinde nokta mutasyonu içeren bir plazmit DNA'sı onların yabanıl karşılıklarıyla birlikte mutasyon tespit analizi için kullanılmıştır. Cas9 proteini, hücrelerde genom mühendisliği uygulamalarında Cas9'un kontrollü kullanımı için bir platform sağlayan bölünmüş bir enzim olarak yeniden yapılandırılabilir. Bu çalışmada, CRISPR/Cas9 düzenleme alanındaki araştırmadan yararlanılmış ve MutS, PAM etkileşimli alan ile değiştirilerek yeniden yapılandırılmış Cas9 proteini üretilmiştir. CRISPR/Cas9 ile indüklenen mutasyonların saptanması için analizlerin kontrollü bir karşılaştırması gerçekleştirilirken, aynı zamanda mutasyonların saptanması için kimerik bir enzim üretmek üzere yeni bir yöntem denenmiştir. Yeniden yapılandırılmış MutS-Cas9 enzimi, mevcut mutasyon tespit yöntemlerine bir alternatif sunmaktadır ve mevcut Cas9 yeniden yapılandırma çalışmalarına bir bakış açısı kazandırmaktadır.

Anahtar Kelimeler: CRISPR/Cas9, mutasyon görüntüleme, MutS, T7 Endonükleaz Kesim Analizi, Yüksek Çözünürlüklü Erime Analizi (HRMA).

ACKNOWLEDGEMENTS

First of all, I would like to express my greatest and sincere gratitude to my thesis advisor Assoc. Prof. Dr. Nuri ÖZTÜRK for his thoughtful guidance. I would also like to thank my jury members Prof. Dr. Halil KAVAKLI and Prof. Dr. Tamer YAĞCI who spent their valuable times to evaluate this thesis.

I am especially grateful to Gözde ÖZÇELİK and Hüseyin GÜL for their endless support. Special thanks go to my dear laboratory mates Bihter MURATOĞLU, Tuba KORKMAZ, Handan EMİŞOĞLU, Feraye Hatice CANBAZ, Fatih AYGENLİ, Büşra ACAR for their great help and precious friendship.

Finally, I would like to express my appreciation to my dear parents Fatma CANTÜRK and Veli CANTÜRK, my sisters, and Burak HAMALİ for the optimistic supports and loves.

The study was funded by the Scientific and Technical Research Council of Turkey (TUBITAK) grant 114S446 (to N.O.).

TABLE of CONTENTS

	Page
SUMMARY	v
ÖZET	vi
ACKNOWLEDGEMENTS	vii
TABLE of CONTENTS	viii
LIST of ABBREVIATIONS and ACRONYMS	xi
LIST of FIGURES	xii
LIST of TABLES	xv
1. INTRODUCTION	1
1.1. The Purpose, Contribution, and Content of Thesis	1
2. LITERATURE REVIEW	2
2.1. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9	2
2.1.1. History of Genome Engineering and CRISPR/Cas System	2
2.1.2. Mechanism of CRISPR/Cas9 System	5
2.1.3. Applications and Challenges of CRISPR/Cas9 System	7
2.1.4. Methods and Tools of CRISPR/Cas9 System	8
2.1.5. Engineering of Cas9	9
2.2. Screening Methods of CRISPR/Cas9-induced Mutations	10
2.2.1. Analysis of Cleaved Amplified Polymorphic Sequences	11
2.2.2. Loss of a Primer Binding Site	12
2.2.3. Sequencing	12
2.2.4. Amplified Fragment Length Polymorphisms	13
2.2.5. Fluorescent PCR Capillary Gel Electrophoresis	13
2.2.6. Mismatch Detection Assays	13
2.2.6.1. The Mismatch Cleavage Assay	14
2.2.6.2. High-Resolution Melting (HRM) Analysis	15
2.2.6.3. Electrophoretic Mobility Shift Assay	16
2.3. MutS Based Methods for Mutation Detection	17

2.4. Biological Function of Prokaryotic Argonaute Proteins 19
3. MATERIALS 21
3.1. General Kits and Reagents 21
3.2. Buffers and Solutions 22
3.3. Antibodies 23
3.4. Equipment 24
4. METHODS 25
4.1. Molecular Cloning 25
4.1.1. Digestion of DNA with Restriction Enzymes 25
4.1.2. Agarose Gel Electrophoresis 25
4.1.3. DNA Isolation from Agarose Gel 25
4.1.4. Ligation of Inserts into Plasmids 26
4.1.5. Transformation to Competent Bacteria 26
4.2. Protein Expression 27
4.3. Protein Purification 27
4.4. Western Blotting 28
4.4.1. Sample Preparation 28
4.4.2. SDS-PAGE Gel Preparation 28
4.4.3. SDS-PAGE Gel Electrophoresis 28
4.4.4. Electrophoretically Transfer the Proteins onto Membrane 29
4.4.5. Blocking and Antibody Incubation 29
4.4.6. Chemiluminescence Visualization of the Membrane 29
4.5. Coomassie Staining 29
4.6. Silver Staining 30
4.7. Cell Culture 30
4.8. Genome Editing 31
4.9. Genomic DNA Isolation 32
4.10. Site-Directed Mutagenesis 32
4.11. Polymerase Chain Reaction (PCR) 33
4.11.1.PCR of MCF10A Wild Type and MCF10A *BMAL1* Knockout Set1 33
4.11.2. PCR of pAc5.1 HisV5 dCRY Wild Type and pAc5.1 HisV5 dCRY Q311E 34
4.11.3. Agarose Gel Electrophoresis 35

4.11.4. DNA Isolation from PCR Products 35
4.12. Annealing of DNA 35
4.12.1. Agarose Gel Electrophoresis 36
4.13. T7 Endonuclease 1 (T7E1) Cleavage Assay 37
4.14. High Resolution Melting (HRM) Analysis 37
4.15. Rapid Agarose Gel Electrophoretic Mobility Shift Assay 38
5. RESULTS 39
5.1. Purification of MutS, TtAgo-MutS and Reengineered MutS-Cas9 Proteins 39
5.2. Analysis of PCR-amplified Wild Type DNA and Mutant DNA Products 42
5.3. Formation of Heteroduplex DNA 43
5.4. T7 Endonuclease 1 (T7E1) Cleavage Assay 45
5.5. High Resolution Melting (HRM) Analysis 46
5.6. Rapid Agarose Gel Electrophoretic Mobility Shift Assay 54
5.7. The Cleavage Activity of TtAgo-MutS Protein and MutS-Cas9 Protein 56
6. DISCUSSION 65

REFERENCES 69
BIOGRAPHY 74
APPENDICES 75

LIST of ABBREVIATIONS and ACRONYMS

Abbreviations and Acronyms		**Explanations**
°C	:	Celsius
µl	:	Microliter
µg	:	Microgram
bp	:	Base pair
APS	:	Ammonium persulfate
BMAL1	:	Brain and muscle Arnt-like protein-1 gene
CRISPR/Cas9	:	Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
dCRY	:	*Drosophila* Cryptochrome
ddH_2O	:	Double distilled water
EDTA	:	Ethylenediaminetetraacetic acid
HRM	:	High resolution melting
IPTG	:	Isopropylthio-β-galactoside
LB	:	Luria Bertani
MCF10A	:	Non-tumorigenic Human Epithelial Breast Cell Line
Ng	:	Nanogram
PBS	:	Phosphate Buffered Saline
pH	:	Power of Hydrogen
PCR	:	Polymerase chain reaction
RPM	:	Revolution per minute
RT-qPCR	:	Real-time quantative polymerase chain reaction
SDS	:	Sodium dodecyl sulfate
SDS-PAGE	:	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TB	:	Tris Boric acid
TBE	:	Tris Boric acid EDTA
Tm	:	Melting temperature

LIST of FIGURES

Figure No:		Page
2.1:	Overview of bacterial CRISPR/Cas system.	4
2.2:	The history of important developments in CRISPR/Cas system and genome engineering areas.	5
2.3:	Native and engineered CRISPR/Cas9 systems.	6
2.4:	The mechanism of DSB repair.	7
2.5:	Domain structure of Cas9.	10
2.6:	The general process of three mismatch detection assays.	14
2.7:	DNA mismatch repair mechanism of *E. coli*.	18
5.1:	The schematic representation of reengineered MutS-Cas9 protein.	39
5.2:	Coomassie Staining (shown in A) and Western Blot (shown in B) analysis of BL21 pET21b MutS C1.	40
5.3:	Coomassie Staining analysis of purified His-tag protein BL21 pET21b MutS C1.	40
5.4:	Coomassie Staining (shown in A) and Western Blot (shown in B) analysis of BL21 pET21b TtAgo-MutS C1 and BL21 pET21b MutS-Cas9 C1.1.	41
5.5:	Silver Staining analysis of purified His-tag proteins BL21 pET21b TtAgo-MutS C1 and BL21 pET21b MutS-Cas9 C1.1.	41
5.6:	2% agarose gel electrophoresis of PCR products amplified from MCF10A set.	42
5.7:	2% agarose gel electrophoresis of PCR products amplified from pAc5.1 HisV5 dCRY set.	43
5.8:	2% agarose gel electrophoresis of annealed MCF10A set.	44
5.9:	2% agarose gel electrophoresis of annealed pAc5.1 HisV5 dCRY set.	44
5.10:	2% agarose gel electrophoresis of T7E1 digestion products (MCF10A set).	45
5.11:	2% agarose gel electrophoresis of T7E1 digestion products (pAc5.1 HisV5 dCRY set).	46

5.12: High resolution melting curve of MCF10A set (derivative). 47
5.13: High resolution melting curve of MCF10A set (normalized). 47
5.14: High resolution melting curve of pAc5.1 HisV5 dCRY set (derivative). 48
5.15: High resolution melting curve of pAc5.1 HisV5 dCRY set (normalized). 48
5.16: High resolution melting curve of MCF10A set including dialysis buffer. 50
5.17: High resolution melting curve of MCF10A set including MutS. 51
5.18: High resolution melting curve of pAc5.1 HisV5 dCRY set including dialysis buffer. 52
5.19: High resolution melting curve of pAc5.1 HisV5 dCRY set including MutS. 53
5.20: 1% agarose gel electrophoresis of MutS binding activity for MCF10A set. 54
5.21: Quantification of MutS binding activity for MCF10A set. 55
5.22: 1% agarose gel electrophoresis of MutS binding activity for pAc5.1 HisV5 dCRY set. 55
5.23: Quantification of MutS binding activity for pAc5.1 HisV5 dCRY set. 56
5.24: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 2 and 6 hours. 57
5.25: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 2 hours. 58
5.26: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 6 hours. 58
5.27: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 2, 6, and 16 hours. 59
5.28: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 2, 6, and 16 hours. 60
5.29: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 6 hours (with internal control). 61

5.30: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 6 hours (with internal control). 61

5.31: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours (with internal control). 62

5.32: Quantification of cleavage activity of TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours (with internal control). 63

5.33: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 72 °C for 3 hours. 64

LIST of TABLES

Table No:		Page
3.1:	List of kits and reagents.	21
3.2:	List of buffers and solutions.	22
3.3:	List of SDS-PAGE gel recipe.	23
3.4:	List of antibodies.	23
3.5:	List of equipment.	24
4.1:	The cycling conditions of site-directed mutagenesis.	33
4.2:	The thermocycling conditions used in PCR of MCF10A.	34
4.3:	The thermocycling conditions used in PCR of pAc5.1 HisV5 dCRY.	34
4.4:	The annealing protocol used in this study.	36
4.5:	The formula is used for indel (%) calculation.	37
4.6:	The melt curve protocol used in this study.	38

1. INTRODUCTION

The subject of this thesis is to present an outline of the efficiencies of different methods for detecting Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9-induced mutations.

While optimization and/or comparison of the available methods were used to detect mutations induced by CRISPR/Cas9, a new method was also tested to generate a chimeric enzyme to detect these mutations.

1.1. The Purpose, Contribution, and Content of Thesis

We aimed to compare and/or optimize the available mutation screening methods for CRISPR/Cas9-induced mutations and to test a new method for mutation screening. Therefore, we tested the available methods; MutS protein which has an affinity for mismatches, T7 Endonuclease which cuts heteroduplexes, High-Resolution Melting Analysis which uses differential binding of DNA dyes to DNA. We replaced MutS with PAM-interacting domain utilizing from the research in the CRISPR/Cas9 editing area and generated reengineered Cas9 protein. Mutant MCF10A *BMAL1* knockout Set1 DNA and pAc5.1 HisV5 dCRY Q311E DNA samples were used for mutation detection analysis.

In recent years, different groups which utilized from CRISPR/Cas9 gene editing method, used different methods to detect mutations. The random formation of mutations and their heterogeneity require some optimizations. Therefore, we expect that this thesis will give a table for selecting the most efficient and affordable method(s). In addition, we tested a challenging idea for reprogramming of Cas9 for PAM independence for detecting mismatches by replacing DNA binding region with MutS.

Chapter 2 gives detailed information about CRISPR/Cas9 System and previously used CRISPR/Cas9-induced mutation detection methods. In chapter 3-4, materials and all experimental methods are explained. In chapter 5, the results are presented. The thesis is concluded with the discussion section in chapter 6.

2. LITERATURE REVIEW

2.1. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9

2.1.1. History of Genome Engineering and CRISPR/Cas System

The development of technologies to manipulate DNA has provided improvements in the new era of biology which began with the discovery of DNA double helix. The application of recombinant DNA technology has facilitated to create mutant genes in cells and in model organisms. The recent improvements in genomic sequencing technologies and the fast creation of whole-genome sequencing data for various types of organisms and humans, has sparked a new revolution in biological research. Ever since the development of technologies in biology, researchers have focused on recognizing the site-specific DNA sequences and aimed to target the site-specific regions of genomes of organisms.

In 1979, chromosome segments were replaced with modified DNA sequences in yeast [1]. In yeast and bacteria, the research of DNA recombination mechanisms and DNA repair pathways revealed that cells have the ability to repair double-strand DNA breaks (DSBs) [2]. In this way, methods for creating breaks in the DNA at desired sites was approved as a precious strategy for targeted genome engineering. Then oligonucleotides or small molecules which recognize DNA base pairs were used for targeted DNA cleavage. An alternative approach based on the base pairing of nucleic acid was the utilization of self-splicing introns to alter DNA or RNA sequences [3]. The utilization of self-splicing introns and site-specific DNA breaks for genome editing revealed the intron-encoded nucleases (homing endonucleases) can be used for incorporation of the selected genetic sequence into a genome [4].

Then the targeted site-specific endonucleases were created by using zinc finger proteins with the nuclease domain of the Fok I endonuclease. This zinc finger nucleases (ZFNs) were used for altering the genomic sequence in cells [5]. However, ZFNs were not largely approved due to some difficulties in designing. Therefore, TAL

effector nucleases (TALENs) were created by the integration of transcription activator-like effectors (TALEs) with the catalytic domain of the FokI endonuclease on the purpose of targeted genome editing [6]. Although the utilization of TAL effector nucleases (TALENs) were better than ZFNs, difficulties of protein synthesis and validation caused a problem for use of TALENs.

In a parallel but different area of genome editing research, a few laboratories in the 2000s began to research the CRISPRs (**C**lustered **R**egularly **I**nterspaced **P**alindromic **R**epeats), which had been defined as a sequence of short direct repeats in the genome of *Escherichia coli* by Japanese researchers in 1987 [7]. To defense themselves from viruses and plasmids, bacteria and archaea have evolved adaptive immune mechanism called CRISPR/Cas (CRISPR-associated) system [8], [9].

This defense system uses small RNAs to detect and silence the specific foreign nucleic acid sequence. CRISPR/Cas systems include a CRISPR array consisting of repetitive unique genome-targeted sequences (spacers) and Cas genes arranged in operon(s). CRISPR/Cas mediated immune system has three steps. Firstly, in the adaptive phase, the microbial invasion is responded with a combination of the proximal site of the CRISPR array and short fragments of foreign sequence (protospacers) into the chromosome of the bacterial and archaeal host. In the expression phase, the repeat-spacer element is transcribed into precursor CRISPR RNA (pre-crRNA) molecules. Then in the response phase, enzymatic cleavage of pre-crRNA forms the short crRNAs which can match with the protospacer sequence of microbial targets. After target recognition by crRNAs, the foreign sequences are silenced with coactivation of crRNA and Cas proteins (Figure 2.1) [10].

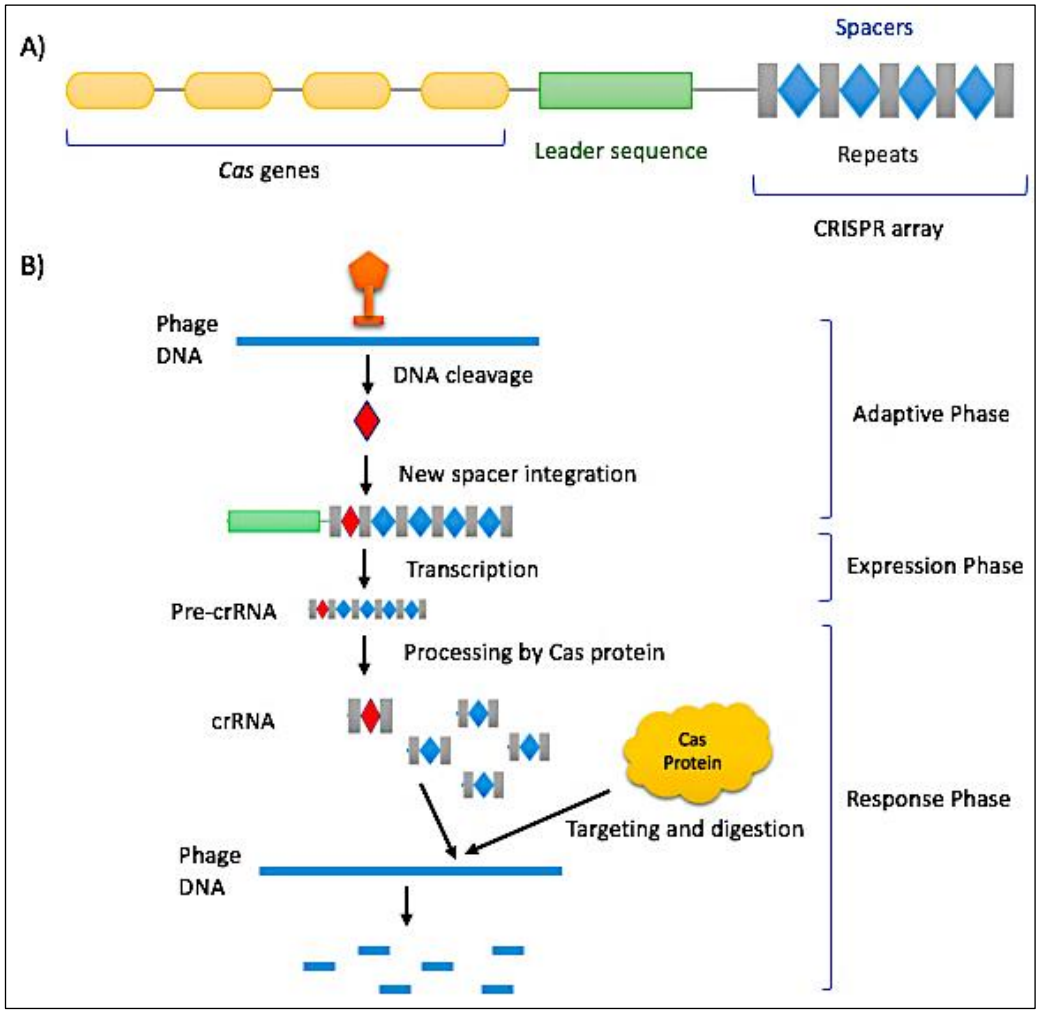

Figure 2.1: Overview of bacterial CRISPR/Cas system. A) The structure of CRISPR locus. B) Adaptive, expression, response phases of bacterial CRISPR/Cas system provide the digestion of invading phage DNA.

There are three CRISPR/Cas system types (I, II, and III) which use different molecular mechanisms to recognize and cleave the nucleic acid. A short sequence motif contiguous to the crRNA-targeted sequence on the invader DNA, the protospacer adjacent motif (PAM), has an important role in the steps of adaptation and response in type I and type II systems. For crRNA-guided targeting, a large complex of Cas proteins is used by the type I and type III systems [11]. However, in the type II system, only a single protein (Cas9) is required for RNA-guided DNA recognition and cleavage of foreign DNA [10]. These important advancements in genome engineering and in CRISPR biology areas were associated in 2012 with the discovery that Cas9 is an RNA-programmable DNA endonuclease in *Streptococcus Pyogenes* (Figure 2.2) [10], [12].

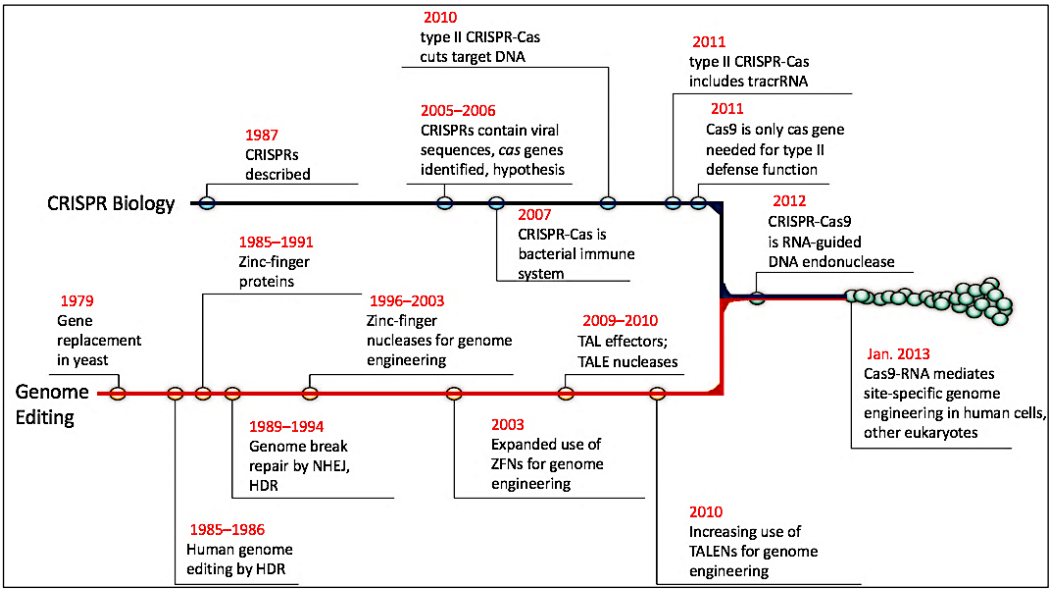

Figure 2.2: The history of important developments in CRISPR/Cas system and genome engineering areas. Important advancements in these fields were merged and shown in 2012.

2.1.2. Mechanism of CRISPR/Cas9 System

The type II CRISPR/Cas system uses Cas9 protein, which includes a RuvC-like nuclease domain and an HNH nuclease domain. CRISPR/Cas9-mediated genome editing relies on the formation of double-strand break (DSB) and thereafter DNA repair. In the CRISPR/Cas9 system, mature crRNA is joined with transactivating crRNA (tracr-RNA) to create a tracrRNA:crRNA complex that leads Cas9 to a target site. A PAM (located immediately downstream of the target sequence) and a DNA sequence protospacer matching crRNA are used for CRISPR/Cas9-mediated sequence recognition at the target site. Subsequent binding to the target site, the HNH and RuvC nuclease domains of Cas9 cut the DNA single-strand matching crRNA and complementary strand, respectively to form a DSB at the target site. Then the researchers create a single guide RNA (sgRNA) with a combination of crRNA and tracrRNA (Figure 2.3). This chimeric gRNA comprises of 22 nucleotide sequences at the 5' end that detect target site of DNA and double-stranded structure at the 3’ end that links to Cas9 [10]. A few CRISPR/Cas9 systems have been created which have 2-4 nucleotide PAM sequences of target sites and programmable gRNA. Thus,

specific 22-29 nucleotide sequences are targeted and changed with CRISPR/Cas9 system (Figure 2.3) [13].

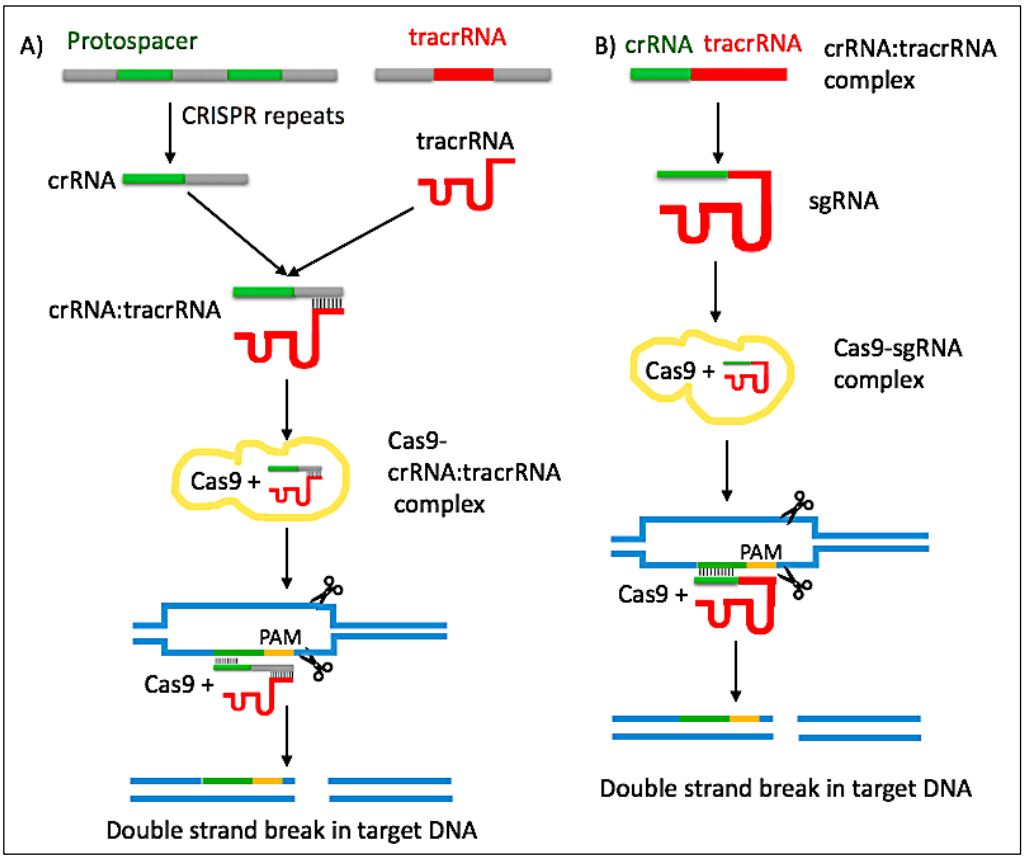

Figure 2.3: Native and engineered CRISPR/Cas9 systems. A) Native CRISPR/Cas9 systems insert foreign DNA into CRISPR arrays, which then form crRNAs including protospacer sequences that are the complement of the foreign DNA. crRNAs combine with tracrRNAs and this RNAs complex guides the Cas9 to degrade the foreign DNA. B) The sgRNA, which is a researcher-engineered hybrid of the crRNA and the tracrRNA sequences, is used in engineered CRISPR/Cas9 system. sgRNAs combine with Cas9 to create DSB in target DNA site.

The DNA repair processes are stimulated with DSB created by RNA guided Cas9. The DSB is repaired via DNA repair mechanisms, containing nonhomologous end joining (NHEJ) mediated DNA repair and homology-directed repair (HDR) mediated DNA repair. NHEJ mediated repair is an error-prone mechanism, ligates the DSB simply and results with deletion or insertion mutations. NHEJ mechanism is used to disrupt the function of target genes. HDR mediated repair is an error-free mechanism which uses a homology including donor DNA sequence as a repair

template. HDR mechanism is used for adding the desired gene into the target location (Figure 2.4) [14].

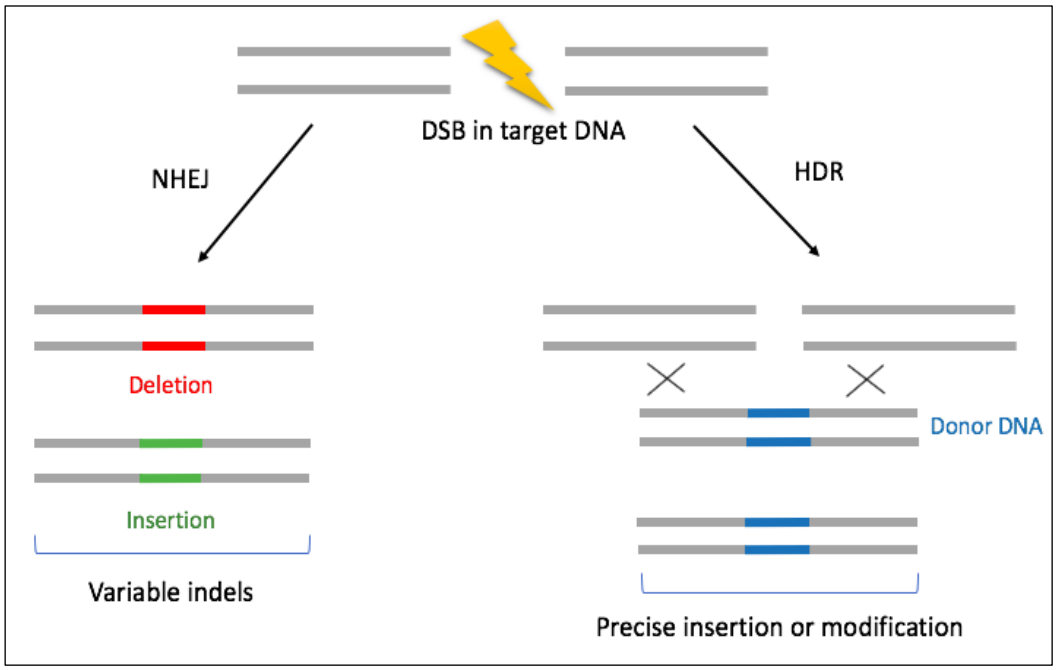

Figure 2.4: The mechanism of DSB repair. The NHEJ mechanism, which creates random deletions or insertions, is used for knockout mutations. The HDR mechanism, which provides the specific gene editing from donor DNA, is used for knock-in mutations.

2.1.3. Applications and Challenges of CRISPR/Cas9 System

In contrast to ZFNs and TALENs, which have difficult protein engineering step for each DNA target site, the CRISPR/Cas9 system uses only an alteration in the gRNA sequence. CRISPR/Cas9 system is an advantageous genome engineering tool for researchers to change target genomic elements. The co-delivery of plasmids including Cas9 and crRNA has been used to create targeted modifications in *Drosophila, Zebrafish, Caenorhabditis elegans,* and mammalian cells. In addition, multiple mutations in genes can synchronically be induced by integrating various gRNA with Cas9. CRISPR/Cas9 is a novel robust genome editing tool used to increase or decrease expression of the desired gene for gene therapy studies [15]. In transcription regulation studies, the transcriptional activity of specific genes can be understood with disrupting transcription related sites via the CRISPR/Cas9 system.

Recently, a catalytically defective Cas9 (dCas9) lacking nuclease activity is developed to create altered CRISPR/Cas9 system known as CRISPR inference (CRISPRi). dCas9 was used for studies of transcription regulation [16].

In spite of the perfect feasibility of CRISPR/Cas9 in genome editing, some problems which are related to PAM dependence, off-target mutations, construction of gRNA and delivery systems of CRISPR/Cas9 should be developed. CRISPR/Cas9 is applied to any DNA sequence via customizable gRNA. In addition to complementarity of gRNA-target sequence, 2–5 nucleotides PAM sequence positioning at next to downstream of the target site is important for specificity of CRISPR/Cas9. The PAM recognition sequence varies according to the properties of bacteria from which the Cas9 is obtained. The defined PAM sequences differ between varied Cas9 orthologs, such as 5'NGGNG and 5'NNAGAA PAM from *Streptococcus thermophilus*, 5'NNNNGATT PAM from *Neisseria meningitidis* and 5'NGG PAM from *Streptococcus pyogenes*. In genome editing, off-target mutations are an important problem for CRISPR/Cas9 system. Large genomes generally include multiple DNA sequences which are homologous sequences of target DNA. The Cas9 cleaves same homologous DNA sequences and this cleavage causes mutations at untargeted locations that named off-target mutations. Off-target mutations can cause cell death or transformation. To decrease the negative effect of CRISPR/Cas9, these undesired mutations should be removed and the target sites including minimum off-target sites should be selected. In addition, Cas9 nickase can be used to decrease off-target mutations [15].

2.1.4. Methods and Tools of CRISPR/Cas9 System

To edit the genome successfully with CRISPR Cas9 system, the gRNA, and the Cas9 should be delivered into cells efficiently. Several methods are used for gene delivery, containing viral and non-viral systems. The appropriate method should be selected according to some parameters. These parameters are; type of research (*in vitro* or *in vivo*), size of interest gene, target cell type, type of expression (transient or permanent). In *in vivo* applications, the vector should cause minimal immune

response after delivery and target desired tissue or cell types. The viral vector can deliver small genes. However, non-viral delivery method can be used for expressing larger genes. Target cell type is important for determining the vector type. Many viral vectors can join into the host's genome to elicit permanent gene expression in target cells [17]. To plan genome editing with CRISPR/Cas9 system according to preferred gene mutation, design of gRNA for a target gene, choosing of the appropriate delivery system and evaluation of genome editing efficiency are crucial steps to be followed.

2.1.5. Engineering of Cas9

Streptococcus pyogenes Cas9 (SpCas9) is a subtype II-A Cas9 protein contains an HNH domain that cuts the DNA strand which is complementary to the gRNA and a RuvC domain that cut the other strand to create DSBs. To understand the crystal structure of spCas9, electron microscopy (EM) and X-ray screening methods were used. This research revealed that spCas9 has a bilobed structure including the catalytic nuclease lobe (contain a C-terminal domain which is related to PAM recognition called PAM-interacting (PI) domain) and α-helical lobe. The two lobes are conformationally rearranged relative to each other after the binding of DNA. Binding of RNA stimulates the nuclease domain to turn 100° relative to the helical domain, creating a nucleic acid binding groove that can hold DNA [18]. The RuvC domain contains three different motifs: Motifs II and III are divided by the HNH domain, and motifs I and II are divided by a wide α-helical lobe and bridge helix. This α-helical lobe called the recognition (REC) lobe creates a large groove via rotation to contact with the gRNA and target DNA [19].

Using studies of Cas9 crystal architectures as a guide, a split-Cas9 was designed (Figure 2.5). The enzymatic activity of the split-Cas9 was similar of natural Cas9. These results exhibited that the Cas9 might be split into two distinct polypeptides, which regenerate a functional Cas9 when returned together by inducing [20]. Additionally, the PI domain can be also partitioned into C-terminal and Topo-homology domains [18].

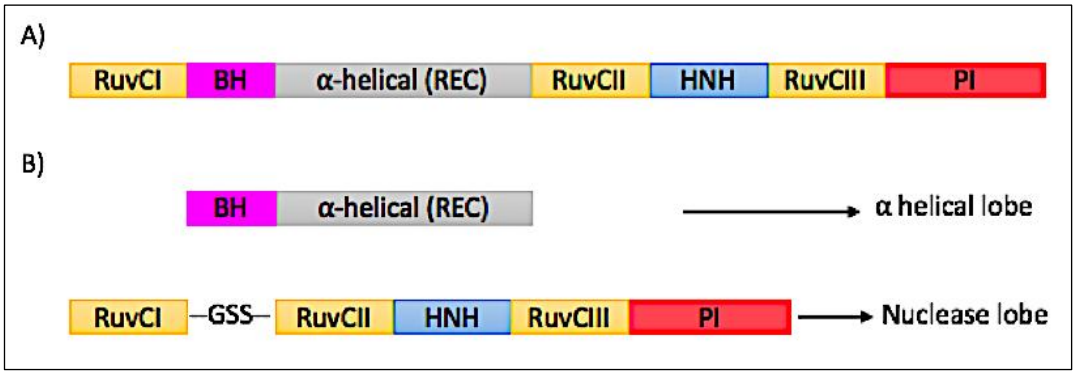

Figure 2.5: Domain structure of Cas9. Domain structure of wild type Cas9 (shown in A) and designed split-Cas9 (shown in B), comprised of the nuclease lobe and the α-helical lobe. BH: bridge helix; REC, recognition lobe; PI: PAM-interacting; GSS: Gly-Ser-Ser.

The architecture of split Cas9 provides to understand the structure of engineered Cas9 and to control the mechanism of engineered Cas9. However, the use of split Cas9 will be beneficial for genome editing applications. The split Cas9 can ease CRISPR technology applications to create Cas9 containing chemically inducible domains. For instance, light and temperature switchable Cas9 variants were engineered [21]. In addition, a split dCas9 containing mutations in the PI domain, which recognizes the NGCG PAM instead of the NGG PAM was developed [22].

2.2. Screening Methods of CRISPR/Cas9-induced Mutations

The CRISPR/Cas9 system is applied to form DSBs at the targeted region of the genome. In higher eukaryotes, DSBs are generally repaired by the NHEJ mechanism which is error-prone and causes small insertions and/or deletions (indels) at the site of the breakage. The indels form a frameshift mutation and they can knock out the gene function by the reason of mRNA decay. The target sequence of the CRISPR/Cas9 system can be altered with changing the gRNA, thus the production and testing of multiple targeting constructs gets easy. However, after delivery of the desired gene into the host organism, created mutations should be confirmed and characterized.

When targeting a single gene of a diploid cell, there are four possible results: no mutation, heterozygous mutation (one allele is edited), biallelic mutation (both alleles are edited but the each allele has different sequence) or homozygous mutation (both alleles have same mutation) [23]. All methods used for analyzing CRISPR/Cas9-induced mutations have advantages and disadvantages. The best method should be selected according to some factors, containing the expected length and rate of mutations, the type of mutation, the sensitivity, cost and throughput of the method. Detection methods for on-target mutations are based on the polymerase chain reaction (PCR).

Detection methods of on target mutations consist of the analysis of cleaved amplified polymorphic sequences, the loss of a primer binding site, sequencing, the amplified fragment length polymorphisms, the fluorescent PCR capillary gel electrophoresis, and mismatch detection assays. These detection methods of on target mutations depend on the polymerase chain reaction (PCR).

2.2.1. Analysis of Cleaved Amplified Polymorphic Sequences

The location of CRISPR/Cas9-induced mutations is usually predictable due to Cas9 induces DSBs at 3 bp upstream of the PAM. This method that includes the use of a combination of PCR and restriction enzymes which detect indels is called cleaved amplified polymorphic sequences (CAPS). The cleaved products are analyzed via HPLC or agarose gel electrophoresis [24]. The frequency of indels is evaluated according to measured densities of cleaved DNA fragments and PCR products [25]. CAPS analysis is able to detect homozygous mutants and all types of mutations (both small and large indels, SNPs) and is thereby highly efficient. Although, the presence of restriction sites including the mismatch limits CAPS analysis [26].

2.2.2. Loss of a Primer Binding Site

When genomic DNA is amplified with two pairs of primers, one including the target site but annealing outside it, and another that contains a primer overlapping the indel region, mutations at the target region will prohibit the next primer annealing and only the bigger amplicon including the whole target region will be generated. For characterization of the mutations, the frequency of mutations can be detected by qPCR and the bigger amplicon including the whole target region can be sequenced. Then the PCR products are analyzed with electrophoresis. This method is fast and cheap. However, the major problem is that point mutations can not be detected because primers can bind to mismatched fragments, and extension continues on condition that terminal nucleotide is matched accurately. The sensitivity of this method is only ~10%, therefore, it is not convenient [27].

2.2.3. Sequencing

In Sanger sequencing, the targeted site is amplified by PCR and the amplicons are cloned into a vector. Thus, each vector has only one gene copy which creates a single trace when sequenced. To identify the sequence of all gene copies, lots of colonies required to be scanned. This method is suitable for indel detection but it may be time-consuming, effortful and costly. Alternatively, PCR products are sequenced directly to save labor and time. Nevertheless, in the case of non homozygous mutations, copy number variations or polyploid organisms, direct Sanger sequencing causes multiple overlapping traces which are troublesome to select. To solve this problem, some web-based programs can be used [28].

Another method of indel detection is Next Generation Sequencing. To perform efficiently NGS screening, the edited site is PCR amplified with barcoded primers, PCR amplicons are pooled, the library is prepared and sequenced. Subsequently, sequencing data is analyzed using software tools and desired mutant clones are identified. NGS method is powerful by the reason of it can detect the

frequency of mutation accurately with 0.01% sensitivity [29]. However, the NGS method is expensive and complicated.

2.2.4. Amplified Fragment Length Polymorphisms

To detect big deletions, the PCR is performed and the different sizes of amplicons can be visualized on an agarose gel. These differences are called as amplified fragment length polymorphisms (AFLPs) [30]. This method is fast and inexpensive. However, this method is limited for detection of length differences with the stability of some base pairs and different methods can be required.

2.2.5. Fluorescent PCR Capillary Gel Electrophoresis

The amplicon is generated by PCR using fluorophore primers and resolved by capillary gel electrophoresis. The mobility of the amplicons containing indel mutation and wild type are different. However, this technique can not report large indels correctly due to overestimation of mutations longer than 30 bp [31]. Additionally, the equipment and software tool are expensive.

2.2.6. Mismatch Detection Assays

Mismatch detection assays are the mismatch cleavage assay, high resolution melting assay and heteroduplex mobility assay. To detect mismatches using these methods, three steps are followed: (1) the target region and its flanking site are amplified by PCR; (2) to create heteroduplex DNA, wild type and mutant DNA fragments are hybridized; (3) the heteroduplex DNA is detected and analyzed by convenient method which can distinguish the difference in structure or melting temperature (Figure 2.6) [23]. These methods are avantegous due to they are fast, simple and cheap. They can genotype single clones and analyze pooled samples but can not identify any properties of the mutation structure. Moreover, if the targeted

region is mostly polymorphic, the results can be hard to comment due to different WT alleles can also create heteroduplex DNA [26]. These methods are semi-quantitative and generally used to compare the efficiency of gRNAs, evaluation of experimental conditions which effect genome editing, initial screening and identifying of lines for the next analysis.

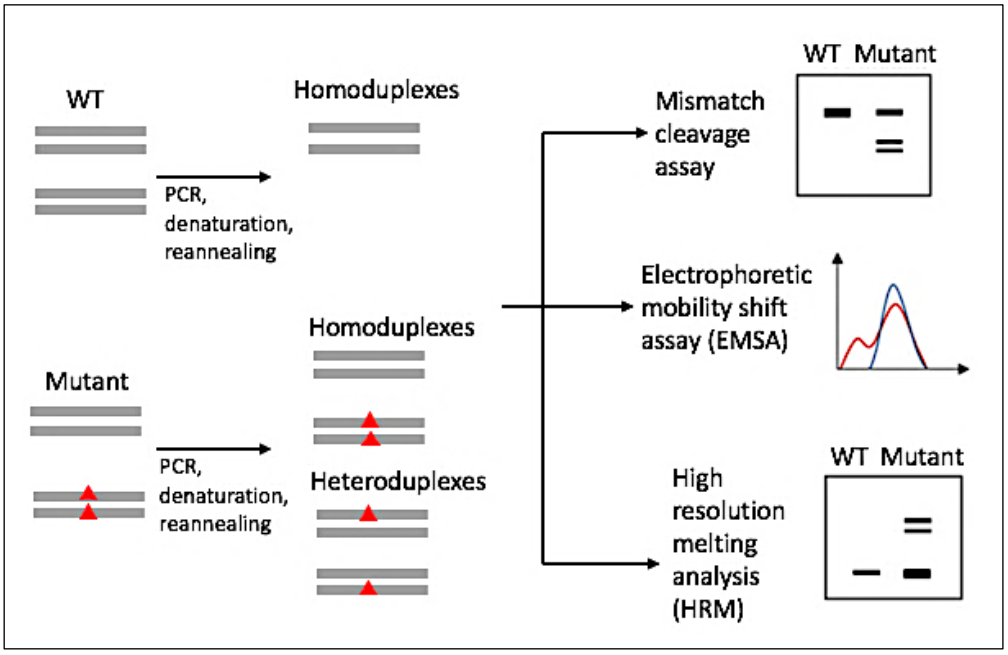

Figure 2.6: The general process of three mismatch detection assays. The preliminary steps of the mismatch detection methods are identical and based on amplification of the target DNA, hybridization of PCR products to create homoduplexes and heteroduplexes. Mismatch cleavage assay with Surveyor or T7 endonuclease is based on the cleavage of DNA protrusions in heteroduplexes, and the cleaved fragments are analyzed in an agarose gel. HRMA detects homoduplexes and heteroduplexes using the Tm difference measured with fluorescence-mediated readout. In EMSA, homoduplexes and heteroduplexes are detected according to their movement in polyacrylamide gel by PAGE.

2.2.6.1. The Mismatch Cleavage Assay

The mismatch cleavage assay is the most common method for detecting CRISPR/Cas9 induced mutations due to it is simple and inexpensive. The protocol of mismatch cleavage assay contains four steps: (1) PCR amplification of the target region from both mutant and wild type DNA; (2) denaturation and re-annealing of mutant and wild type DNA to create heteroduplexes; (3) incubation of both

heteroduplex and reference homoduplex DNA with nuclease to digest heteroduplexes; and (4) detection and analysis of cleaved DNA fragments using agarose gel electrophoresis or high performance liquid chromatography (HPLC). The quality of the PCR product used as a substrate is important for successfully performing of the assay. Therefore PCR amplification must be optimized to form a single band quality product in enough amounts. The PCR product which is low yield causes background after digestion [24]. To estimate the frequency of indels, densities of the PCR product and cleaved DNA fragments at the locus are measured [32].

Generally two enzymes are used in this assay. These enzymes are Surveyor nuclease and T7 endonuclease 1 (T7E1). The Surveyor nuclease is a member of CEL I nuclease family, isolated from celery, recognizes DNA mismatches (single nucleotide polymorphisms (SNPs) or small indels) and cuts both DNA fragments at the 3' site of mismatches [24].

T7E1 is a 149 amino acid protein encoded by gene 3 of *E.coli* bacteriophage T7 [33]. The conformational changes in T7E1 are formed by binding of heteroduplex DNA. The N-terminus of T7E1 creates a central opening for cleavage activity [34]. Then T7E1 recognizes and cuts heteroduplex DNA at the first, second or third phosphodiester bond at 5' site of the mismatches. T7E1 is more sensitive (0.5–5%) and convenient for the detection of indels. However, Surveyor nuclease is better to detect SNPs [35]. Therefore the enzyme should be chosen according to types of mutations which required to be identified.

The using of Surveyor kit is more robust due to coming with a systematized protocol, but it is more expensive. The performance of T7E1 depends on reaction conditions (temperature, salt concentration, DNA/enzyme ratio, incubation time) so optimization can be required.

2.2.6.2. High-Resolution Melting (HRM) Analysis

High resolution melting analysis (HRMA) includes the amplification of target DNA sequence (90–200 bp) by realtime PCR using the intercalating fluorescent dye, subsequently melt curve analysis of the PCR products. The dsDNA intercalating and

saturating dyes used in HRMA are SYTO9, SYBR GreenER, LCGreen/LCGreen Plus, ResoLight Dye and EvaGreen. EvaGreen and LCGreen Plus have high efficiency and sensitivity [36]. The fluorescent dye fluoresces while the DNA is annealed in a double strand structure. When the temperature increases narrowly (0.1-0.3 °C), the dye is released during dsDNA denaturation. Loss of florescence provides to analyze denaturation profile which reveals the Tm. While a homoduplex amplicon created from a homozygous DNA sample will have a specific Tm, a heteroduplex amplicon created from a heterozygous DNA will have an extra Tm, usually a much lower Tm signal. This temperature shift provides the identification of mutant alleles containing SNPs and small indels [37].

In HRMA, to detect mutations at any target site, PCR conditions and primers should be optimized for obtaining smooth melt curve with a single peak. Use of short amplicon increases the resolution of HRM because it provides an increasement in difference of Tm caused by mutation. But use of very small amplicon can not detect large indels. Therfore, to maximize the resolution of HRM, ~100 bp amplicon should be used, although to detect SNPs, 50 bp amplicon should be preferred [37].

HRMA is an extremely sensitive and simple method which is also presents a high throughput detection format (96 well microtiter plates). In addition, amplicons are not damaged by HRMA, so amplicons can be sequenced to identify the precise sequence of the edited region. The mutation type and the amplicon size effects the sensitivity of HRM. In detection of 100 bp amplicon, the sensitivity of HRM was determined as 2% for indels bigger than 4 bp [38]. However, HRMA can not detect larger indels. The setup costs can be much, although it can be decreased by combination of an available qPCR machine and free HRMA software.

2.2.6.3. Electrophoretic Mobility Shift Assay

Typical electrophoretic mobility shift assay (EMSA) experiments include separation of nucleic acids using polyacrylamide gel electrophoresis (PAGE). CRISPR/Cas9 induced mutations can be detected by PAGE of rehybridized PCR products. This method utilizes from the differential movement of homoduplex and

heteroduplex DNA in polyacrylamide gel. The structural difference between matched and mismatched DNA strands containing indels, causes the heteroduplex DNA move slower than homoduplex DNA. The migration difference provides to distinguish the heteroduplex DNA. PCR products of 130–160 bp can be analyzed in a polyacrylamide gel. When reaction conditions are optimized, the sensitivity of such assays can be increased [39]. This method is advantageous because it does not contain slow enzyme reactions and removes the false negative scores arised by the unfinished cleavage of mismatched DNA strands. Additionaly, this method is simple, fast and low cost. Nevertheless, solely small fragments can be identified so this method detects SNPs and small indels [40].

Because of its rapidity, effortlessness and low cost, EMSA is the most widely used method for analyzing connections between proteins and DNA or RNA. The use of polyacrylamide gels for traditional EMSA contains some disadvantages. For instance, PAGE requires preparation of unpolymerized acrylamide, a powerful neurotoxin. In addition, preparation of polyacrylamide gels is more difficult than the preparation of agarose gels. Methods which avoid using of PAGE increase safety and decrease cost. Therefore, agarose gels have been used in some past studies required EMSA [41].

2.3. MutS Based Methods for Mutation Detection

All organisms have mismatch repair systems containing enzymes which repair damaged and mispaired DNA to prevent the formation of mutations. One of the most well-understood mismatch repair systems is the MutHLS protein complex from *Escherichia coli*. The MutS protein is the 97 kDa polypeptide that recognizes and binds to the mismatched DNA. The MutL protein joins MutS-mismatched DNA complex with the help of ATP. Then, the formation of MutS-MutL-DNA complex stimulates the endonuclease activity of MutH protein. MutH cuts the DNA at hemimethylated GATC sites. The region of DNA between the MutH slit site and the mismatch (up to 1000 bases long) is later eliminated and resynthesized changing the mismatch with the correct base (Figure 2.7) [42].

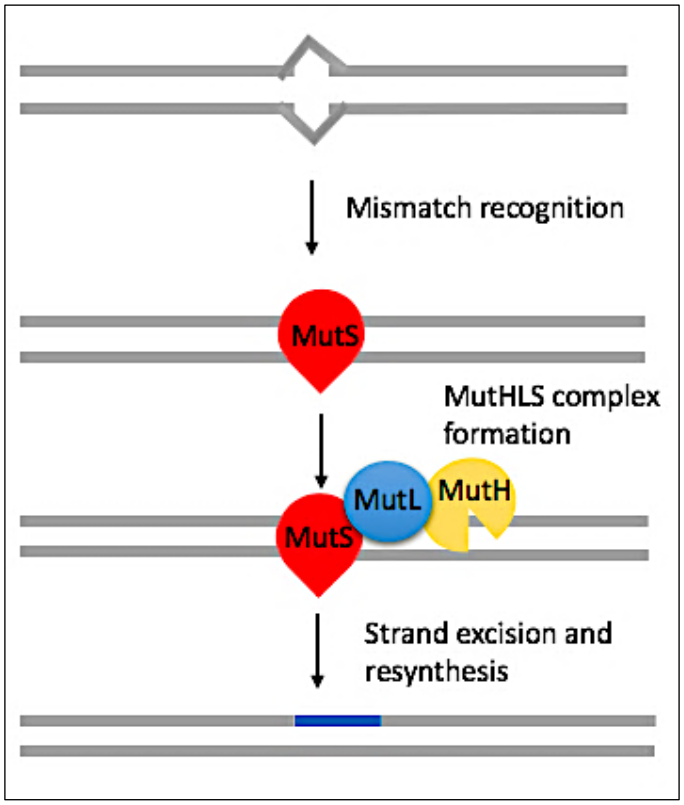

Figure 2.7: DNA mismatch repair mechanism of *E. coli*. The MutS recognizes and binds to mispaired or unpaired bases. After mismatch binding, MutS creates a complex with the MutL ATP dependently. MutS–MutL complex induces the MutH, which cuts the hemimethylated dGATC sequences. The removed mismatch region is resynthesized changing the mismatch with the correct base.

In vitro, the MutS can recognize and bind to single base mismatches and 1-5 base indels, independently of other proteins and cofactors [43]. The MutS origin and mismatch type effect the yield of mismatch binding. *Escherichia coli* MutS can efficiently bind to A:A, A:C, G:G, G:T mismatches, but binding of A:G, T:C, T:T mismatches depend on nucleotide sequence context. In addition, the efficiency of C:C mismatch binding is very low level [44], [45]. The N-terminal part of MutS recognizes mismatch and the C-terminal part, including an ATPase domain, is responsible for oligomerization [43].

MutS is efficiently used for mutation detection. The MutS mediated mutation detection method contains four steps: (1) PCR amplification of the target site from both mutant and wild type DNA; (2) hybridization of mutant and wild type DNA to create heteroduplex; (3) generation of MutS-DNA complexes; and (4) detection and

analysis of MutS-DNA complexes [43]. Many methods are used for detection and analyzing of MutS-DNA complex.

These used methods are: microscopic screening of homoduplex and heteroduplex with atomic force microscopy (AFM), the analysis of MutS using GFP or fluorescent MutS (cyanine labeled) or biotinylated tag, the analyzing of MutS-DNA binding on a solid phase, containing filter tests with biotin or radioactive marked DNA, and on-chip detection [43].

The electrophoretic mobility shift assay (EMSA) is used to determine MutS-DNA interaction [46]. The PAGE of MutS-DNA complex, provides to detect the mutation. The MutS-heteroduplex DNA complex moves slowly in polyacrylamide gel and this movement results in the formation of retarded DNA fragments. Another method, named SYBR Gold-MutEx assay is used for MutS mediated mutation detection. This method is based on DNA protection by MutS from nuclease digestion [47].

2.4. Biological Function of Prokaryotic Argonaute Proteins

Argonaute protein has RNA or DNA-guided endonuclease activity and contains two groups including eukaryotic argonaute (eAgo) and prokaryotic argonaute (pAgo). Argonaute protein has bilobed structure which consists of the C-terminal lobe including the PIWI (P element-induced wimpy testis) domain and the MID (middle) domain, and the N-terminal lobe including the PAZ (PIWI-Argonaute-Zwille) domain and the N-terminal. The C-terminal lobe of Ago has endonucleolytic cleavage activity [48]. The eAgo is a part of the RNA-induced silencing complex (RISC) which has role in RNA interference (RNAi) pathways. The pAgo has similar gene silencing activity [49].

The Argonaute protein of *Thermus thermophilus* (TtAgo) was studied to understand the biological function of pAgos. The structure of TtAgo reveals that pAgos have some different properties. TtAgo cuts ssDNA and dsDNA in addition to ssRNA using DNA guide. Proposed three models for activity of TtAgo are DNA chopping, DNA guide loading and canonical target slicing. The unguided TtAgo has

chopping activity which is cleavage of dsDNA at AT-rich sites. This activity produces small duplex DNA. To perform the DNA guide loading activity, TtAgo binds to small duplex DNA and cuts the target strand. Then cleaved strands are released from TtAgo and guide strand is held. To perform canonical target slicing activity, TtAgo binds to guide ssDNA and target DNA. Then target DNA is cleaved by TtAgo. Additionally, one model which is cleavage of ssDNA by chopping products-guided, was proposed by association of these three activity [50].

These properties and roles of TtAgo provide that pAgos can be used in gene editing areas. In mammalian cells, the requirement of high temperature for cleavage of dsDNA limits the TtAgo-mediated gene editing. However, TtAgo can be used as a programmable DNA-binding platform due to its ssDNA cleavage activity at 37 ˚C. The recent studies shows pAgo based platform can be used to create artificial restriction enzymes and gene editing enzymes [51].

3. MATERIALS

3.1. General Kits and Reagents

Cell Culture: The MCF10A (Cat. No. CRL-1031) cell line was supplied by American Type Culture Collection (ATCC; Rockville, MD).

Genome Editing: The ultracompetent Stbl3 strain of the *Escherichia coli* cells was used in cloning studies. pX462 (#48141), and LentiCRISPR.V2 (#52961) were bought from Addgene.

Table 3.1: List of kits and reagents.

10X NEBuffer 2	B7002S, New England Biolabs (NEB), USA
Chemiluminescent HRP Substrate	K-12043-D10, Advansta, USA
Cholera toxin	Lonza
DNA Ladder	N3231S, New England Biolabs (NEB), USA
DNA Loading Dye	B7021S, New England Biolabs (NEB), USA
Dulbecco's Modified Eagle Medium (DMEM)	41965-039, Gibco Invitrogen, USA
Epidermal Growth Factor (EGF)	Sigma, USA
EvaGreen Dye, 20X in water	#31000, Biotium, USA
Fetal Bovine Serum (FBS)	Thermo Fisher Scientific
Genomic DNA Mini Kit	K1820-02, PureLink Genomic DNA Mini Kit, Thermo Scientific, USA
Hydrocortisone	Sigma
Insulin	Sigma
Maxima SYBR Green/ROX qPCR Master Mix	K0221, Thermo Scientific, USA
MEGM™ Mammary Epithelial Cell Growth Medium, MEGM Bulletkit	Cat No: CC-3150, Lonza, Switzerland
Nonessential amino acids	Thermo Fisher Scientific
Nucleospin Plasmid Kit	740588.50, Macherey Nagel, Germany
PCR clean-up Gel Extraction Kit	740609.50, Macherey Nagel, Germany
Penicillin	Thermo Fisher Scientific
Polybrene	Santa Cruz, TX, USA

Table 3.1: Continued.

Pronosafe Nucleic Acid Staining	CK130, CONDA, S. A
Protein Molecular Weight Marker	P7712S, New England Biolabs, USA
QuikChange II Site Directed Mutagenesis Kit	200523, Agilent, USA
Streptomycin	Thermo Fisher Scientific
T7 Endonuclease 1	M0302S, New England Biolabs, USA
Taq DNA with Standard Taq Buffer	M0273S, New England Biolabs, USA

3.2. Buffers and Solutions

Table 3.2: List of buffers and solutions.

5% Blocking Solution	5% non-fat dry milk in 1X TBS-T
1X Lysis Buffer	1X Laemmli Buffer 1% Beta-Mercaptoethanol
10X PBS	BE17-517Q, Lonza, Belgium
4X Protein Loading Dye (Laemmli Buffer)	20% β-mercaptoethanol 0.008% Bromophenol Blue 40% Glycerol and 8% SDS 250mM Tris-HCL (pH: 6.8)
0.5X TB	45 mM Tris, 45 mM boric acid
10X TBE	55g Boric acid 40 ml 0.5M EDTA (pH 8.0) in 1 L ddH2O 108g Tris base
10X TBS	1.5 M NaCl 200 mM Tris-Cl (pH: 7.6)
1X TBST	1X TBS 0.1% Tween-20
1X Transfer Buffer	20% Methanol 10% 10X SDS Running Buffer
10X SDS Running Buffer	1.90 M Glycine 1% SDS 250 mM Tris-Base (pH:8.3)
Coomassie Destain Solution	700 ml ddH2O 100 ml Glacial acetic acid 200 ml Methanol
Cooomassie Stain Solution	1 g Coomassie R250 500 ml ddH2O 100 ml Glacial acetic acid 400 ml Methanol
LB Agar and LB Broth Medium	LBP04, Caisson Labs, USA and LBP03, Caisson Labs, USA

Table 3.3: List of SDS-PAGE gel recipe.

6% Separating gel	5 ml	10 ml	15 ml
ddH2O	2.9	5.8	8.7
40% Acryl:Bis	0.75	1.5	2.25
1 M Tris pH 8.8	1.25	2.5	3.75
10% SDS	0.05	0.1	0.15
10% APS	0.05	0.1	0.15
TEMED	0.004	0.008	0.012
7% Separating gel	5 ml	10 ml	15 ml
ddH2O	2.775	5.55	8.325
40% Acryl:Bis	0.875	1.75	2.625
1 M Tris pH 8.8	1.25	2.5	3.75
10% SDS	0.05	0.1	0.15
10% APS	0.05	0.1	0.15
TEMED	0.004	0.008	0.012
8% Separating gel	5 ml	10 ml	15 ml
ddH2O	2.65	5.3	7.95
40% Acryl:Bis	1	2	3
1 M Tris pH 8.8	1.25	2.5	3.75
10% SDS	0.05	0.1	0.15
10% APS	0.05	0.1	0.15
TEMED	0.003	0.006	0.009
5% Stacking Gel	2 ml	4ml	6 ml
ddH2O	1.46	2.92	4.38
40% Acryl:Bis	0.25	0.5	0.75
1 M Tris pH 6.8	0.25	0.5	0.75
10% SDS	0.02	0.04	0.06
10% APS	0.02	0.04	0.06
TEMED	0.002	0.004	0.006

3.3. Antibodies

Table 3.4: List of antibodies.

Name	Dilution	Source
His Tag Mouse IgG	1:2000	#2366, Cell Signaling Technology
Anti-mouse, HRP	1:3000	#7076, Cell Signaling Technology

3.4. Equipment

Table 3.5: List of equipment.

Agarose Gel Electrophoresis	1704406, Bio-Rad, USA
Centrifuges	521-1647 refrigerated, Micro Star 17R, VWR, USA 521-1646 Micro Star 17, VWR, USA 521-2844 Mini Star, VWR, USA NF400-Nuve, 02-1827
Chemiluminescence Imaging System	ChemiDoc™ XRS+ System, Bio-Rad, USA
Electrophoresis System	1658004, Mini-PROTEAN® Bio-Rad, USA
Incubator (CO2)	Thermo-Scientific, USA
Inverted Microscope	Nicon, Eclipse, E100
Laminar Flow Cabinet	Hera-Safe Class II Cabinet, Thermo Scientific, USA
NanoDrop	06-26300-48, Shimadzu Biotech
PCR Thermal Cycler	Biometra Thermal Cycler, Analytik Jena, Germany
Power Supply	Bio-Rad, USA
RT-qPCR system	StepOne Real-Time PCR Systems, Thermo Scientific, USA
Shaker	Standart Analog Shaker, VWR, USA
TransBlot Transfer System	1704155, Bio-Rad, USA
Vortex	Scientific Industries, USA
Water Bath	Wisd, Germany

4. METHODS

4.1. Molecular Cloning

To create reengineered MutS-Cas9 protein and recombinant TtAgo-MutS protein and MutS protein, molecular cloning method was used.

4.1.1. Digestion of DNA with Restriction Enzymes

The DNA and pET21b plasmid vector were digested with appropriate restriction enzymes. EcoRI and XhoI restriction enzymes are used for MutS. HindIII and XhoI restriction enzymes are used for MutS-Cas9. To set digestion reaction in 50 μl reaction volume, 1 μl each restriction enzyme, 2 μl DNA, 2 μl 10X restriction enzyme buffer and ddH_2O (until 50 μl final volume) were added into tube. The digestion reaction was incubated at 37 °C for 2 hours.

4.1.2. Agarose Gel Electrophoresis

1% agarose gel was prepared with agarose powder with 1X TBE buffer and pronosafe then it was placed in horizontal electrophoresis tank loaded with 1X TBE buffer. To extract double digested vectors and DNA from agarose gel, digestion products and the 1 kb marker were run on 1% agarose gel. Then the gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.1.3. DNA Isolation from Agarose Gel

To obtain the double digested vector from the agarose gel, gel slice containing DNA fragment was cut with scalpel under the UV light. PCR Clean-up kit (Macherey-Nagel) was used. Threefold of gel slice volume NT1 was added into gel slice in

eppendorf tube. Then tube was incubated at 55 °C at 10 min until the gel was dissolved. The sample was loaded into the column. The column was centrifuged at 10000 rpm for 10 sec and 13000 rpm for 10 sec. The supernatant was discarded and 650 µl NT3 was added into the column. The centrifugation of column was carried out at 10000 rpm for 10 sec and 13000 rpm for 10 sec. The supernatant was discarded and the column was centrifuged 13000 rpm for 1 min. The column was placed to new tube and 50 µl ddH_2O was added into the column. For elution of the DNA, column was centrifuged at 13000 rpm for 1 min. The column was discarded and DNA was placed into new tube.

4.1.4. Ligation of Inserts into Plasmids

T4 DNA ligase was used to ligate DNA fragments into vectors. 2 µl T4 DNA Ligase Buffer (5X), 3 µl (50 ng) vector DNA, 4 µl (150 ng) insert DNA (1:3 M vector-insert ratio) and 1 µl T4 DNA Ligase were added into microcentrifuge tube. 10 µl reaction mix was incubated at 4 °C overnight.

4.1.5. Transformation to Competent Bacteria

Competent *Escherichia coli* BL21 cells were taken from -80 °C deep freeze and dissolved on ice. 1-2 µl of ligated plasmids were added into competent cells in eppendorfs. The samples were incubated on ice for 20 min before heat shock at 42 °C for 1 min. Samples were placed on ice for 1,5 min. 900 µl LB was added into tubes and they were incubated at 37 °C for 1 hour on shaker. 100 µl of the samples were spread on LB agar plates including ampicillin. Then plates were incubated at 37 °C overnight.

4.2. Protein Expression

After transformation, single colonies were selected by micropipette tip and added into 10 ml liquid LB medium containing 10 µl ampicillin to produce a starter culture. The tubes were incubated at 30 °C on shaker overnight. For large scale, 500 ml of liquid LB medium including 500 µl ampicillin was inoculated with 1 ml of 16 hours grown culture. Samples were incubated at 37 °C until OD_{600} reaches 0.4. Then samples were incubated at 16 °C for 30 min until OD_{600} reaches 0.7. To induce cultures for protein expression, 0.5 M IPTG was added into erlenmeyer flask. Samples were incubated at 16 °C on shaker overnight. After induction, bacteria were centrifuged at 4000 rpm for 15 min. Pellets were resuspended with 20 ml PBS and centrifugation was performed at 5000 rpm for 10 min. Supernatant was poured off and pellets were frozen at -80 °C.

4.3. Protein Purification

The supporting information of article [52] was followed for protein purification. To purify recombinant proteins from bacterial cells, pellets of BL21 cells were treated with 30 ml of lysis buffer containing 0.5% Triton-X100, 50 mM Tris pH 7.5, 0.1% NP40, 150 mM NaCl with 10 times of sonication for 10 sec each time on ice. The cell lysates were cleared by centrifugation at 17000 xg for 1 hour. AntiFLAG M2-agarose beads (Sigma) were used for purifying of flag tagged protein. To remove unbound proteins, 500 µl of flag-agarose beads were treated with 30 ml of cleaned cell lysate for 2 hours and then treated three times with 15 ml of 1X TBS. Recombinant proteins were extracted from Flag-agarose beads in 5 ml of 1X TBS containing 100 µg/ml Flag peptide. Eluted proteins were dialyzed towards Storage Buffer containing 50% (v/v) glycerol, 100 mM NaCl, 50 mM Tris-HCl, pH 7.5, 5 mM dithiothreitol.

To purify recombinant proteins containing the V5His tags, nickel affinity chromatography was used. The cell pellets from 1 L cultures were lysed in 30 mL lysis buffer with 10 times of sonication for 10 sec each time on ice. The cell lysates

were cleared by centrifugation at 17000 xg for 1 hour. To avoid nonspecific binding to Ni-agarose beads, 30 mM imidazol was added into cleared cell lysates. 500 µl of Ni-agarose beads was incubated with each cell lysate at 4 °C for 2 hours. Then unbound proteins were treated with lysis buffer containing 30 mM imidazole three times. Recombinant proteins were seperated from the beads with 1 X TBS containing 250 mM imidazole for 1 hour and then dialyzed against Storage Buffer.

4.4. Western Blotting

4.4.1. Sample Preparation

500 µl of induced and uninduced bacterial cells were centrifuged at 13000 rpm for 1 min. Pellets were lysed in 100 µl 1X Protein Loading Dye (Laemmli Buffer). 2X Protein Loading Dye (Laemmli Buffer) was added into purified protein samples.

4.4.2. SDS-PAGE Gel Preparation

The separating SDS-PAGE gel was prepared using materials in Table 3.3 and poured between the glasses in gel casting system. 2-propanol was added top of the gel to prevent bubbles. After separating gel was solidified, stacking gel which was prepared according to Table 3.3 and comb were placed on the separating gel. After the polymerization of the gel, comb was discarded and gel casting system was placed in the vertical electrophoresis tank.

4.4.3. SDS-PAGE Gel Electrophoresis

After preparation of the SDS-PAGE gel electrophoresis system, 1X SDS running buffer was added into the system. Protein samples which were incubated at 95 °C for 10 min and protein marker were loaded in wells of the gel respectively. The gel was run at 80 V for 15 min, then at 130 V until loading dye exited from the gel.

4.4.4. Electrophoretically Transfer the Proteins onto Membrane

After electrophoresis, the gel was placed into 1X transfer buffer. 1 nitrocellulose membrane and 4 filter papers were soaked with 1X transfer buffer. Then to create electrotransfer sandwich, 2 filter papers (+), nitrocellulose membrane, the gel, 2 filter papers (-) were placed into trans-blot turbo transfer system. The transfer was performed at 12V for 60 min.

4.4.5. Blocking and Antibody Incubation

The membrane was blocked with 5% blocking solution for 1 h on shaker after transferring. Then the blocked membrane was incubated with 5% blocking solution containing primary anti-His antibody at 4 °C overnight. After overnight incubation, the membrane was washed with 1X TBS-T for 5-15-5-5 min respectively. The membrane was incubated with 5% blocking solution containing secondary anti-mouse HRP antibody at 25 °C for 1 h. Then the membrane was treated with 1X TBS-T for 5-15-5-5 min respectively.

4.4.6. Chemiluminescence Visualization of the Membrane

After the antibody incubation, the membrane was treated with the chemiluminescent substrate using the manufacturer's recommendation (In our laboratory, Western Blotting Detection Kit Advansta was used) for 2 min. Then the membrane was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.5. Coomassie Staining

SDS-PAGE gel was displaced from glass and placed in a suitable container with a lid. Coomassie stain solution was added into container to cover the gel. The gel was incubated in Coomassie stain for 1 hour on shaker. After incubation, Coomassie

stain was removed. The gel was treated with Coomassie destaining solution until the Coomassie blue background removed from the gel. Then the gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.6. Silver Staining

The purified TtAgo-MutS-His and Cas9-MutS-His proteins were run at 7% SDS gel. After running, this gel was washed with ddH_2O. The gel was washed with Fixation Buffer (20 ml acetic acid, 180 ml ddH_2O) for 1 hour until the stain of the paint disappears for fixation. Then the gel was rinsed with ddH_2O 2 min for three times. The gel was incubated in shaker with AgNO3 Buffer (0.2 g AgNO3 in 200 ml ddH2O, 300 μl 37% Formaldehyde) in the dark. Then the gel was washed with ddH_20 for 5 sec. The gel was incubated in shaker with Developing Buffer (6 g Sodium Carbonate in 200 ml ddH_2O, 300 μl 37% Formaldehyde, 10 mg/ml Sodium Thiosulphate (before using, 0.1 g Sodium Thiosulphate in 10 ml ddH2O) until the bands appear. When the bands were appeared, the gel was washed with Fixation Buffer for 2-3 min. Then the gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.7. Cell Culture

The MCF10A cell line is a non-tumorigenic mammary epithelial cell line and grows as a monolayer and can be differentiated in a convenient culture medium to grow into similar cells in breast epithelial cell. The MCF10A cell line was used to create *BMAL1* mutant gene in this study. This cell line maintained in a humidified incubator at 37 °C under 5% CO_2.

Firstly, for initial growing and stocking, the MCF10A cell line was grown in Mammary Epithelial Basal Medium (MEBM), containing SingleQuots (MEGM) and 100 ng/mL cholera toxin. This cell line was grown and cultured in a DMEM-based medium including 100 μg/mL streptomycin, 100 U/mL penicillin, 1% nonessential amino acids, 10 μg/mL insulin, 0.5 mg/mL hydrocortisone, 10% fetal bovine serum (FBS) and 20 ng/mL EGF.

4.8. Genome Editing

The CRISPR Design Tool sgRNA design program was used for selection of the most efficient sgRNA to target gene in the MCF10A cell genome-edition with CRISPR [32]. 1× NEB buffer 2 solution was added into double stranded oligos in microcentrifuge tubes. The reaction was annealed by heating to 95 °C for 5 min, then cooling to 25 °C. 100 ng of annealed oligos were ligated into 50 ng of LentiCRISPRv2 digested with BsmBI.

Ligation product was transformed into Stbl3 strain of *Escherichia coli* cells. 50% of the ligation product was added into 100 μL competent cells, and kept on ice for 30 min. The cells were incubated at 42 °C for 60 s to heat shock and placed back into ice for 2 min. 900 μL of LB medium was added into the cells. These cells were incubated at 37 °C for 1 hour with rotating. Then, 100 μL of the transformation products was spread on LB agar plates including 100 μg/mL ampicillin for selection of single colonies. The hU6-F oligo (5′-GAGGGCCTATTTCCCATGATT-3′) was used with DNA sequencing to confirm selected colonies.

The co-transfection of LentiCRISPRv2-based CRISPR construct was performed with pCMV-dR8.2 dvpr and pCMV-VSVG packing plasmids into HEK293T cells for production of lentiviral particles. After 48 hours of transfection, the collected media which contains lentiviral particles was filtered with a 0.2 μm filter and it was protected at −80 °C. To infect MCF10A cells with lentiviral particles, 0.5 mL fresh culture medium, 0.5 mL medium including viral particles, 8 ng/mL polybrene were added into MCF10A cells in 12-well plates.

To select infected MCF10A cells, 0.5 mg/mL puromycin was used for 3–4 days. The single cell-derived clones were grown and selected by serial dilution. The selected cells were grown in well plates. Two genomic locations of *BMAL1* gene were targeted separately for decreasing of nonspecific off-target effects of coincidental joining or picking up. To use in next experiments for each targeting sgRNA, two single cell-derived cell lines were selected.

The human *BMAL1* gene was targeted by guide sequence: Forward primer of *BMAL1* T2: 5′ CACCGTAGATAAACTTACTGTGCTA 3′; Reverse primer of *BMAL1* T2: 5′ AAACTAGCACAGTAAGTTTATCTAC 3′.

4.9. Genomic DNA Isolation

PureLink Genomic DNA Mini Kit protocol was followed for isolation of genomic DNA (gDNA). The lysate of MCF10A *BMAL1* wild type and MCF10A *BMAL1* knockout Set1 cells pellets were suspended in 200 μl PBS and 20 μl Proteinase K and 20 μl RNase A was added. The samples were vortexed and incubated at 25 °C for 2 min. The Lysis/Binding Buffer was added into the samples. To promote protein digestion, the samples were incubated at 55 °C for 10 min. 200 μl 100% ethanol was added into the lysate. The lysate was vortexed and loaded into the PureLink Spin Column. For binding of the DNA, the column was centrifuged at 10000 xg for 1 min at 25 °C. The column was centrifuged with 500 μl Wash Buffer 1 at 10000 xg for 1 min and 500 μl Wash Buffer 2 at 17000 xg for 3 min. To elute the DNA, 50 μl Elution Buffer was added into the column. The column was centrifuged at 17000 xg for 1 min. The column was discarded and DNA was placed into new tube. NanoDrop spectrophotometer was used to measure concentration and OD260 and OD280 values were utilized for quantifying concentration and purity of the genomic MCF10A *BMAL1* wild type DNA and MCF10A *BMAL1* knockout Set1 DNA.

4.10. Site-Directed Mutagenesis

To form point mutation (CAG-GAG) Q311E in pAc5.1 HV5 dCRY cell, QuikChange II Site-Directed Mutagenesis Kit was used. To prepare the reaction, 5 μl of 10X reaction buffer, 125 ng of forward primer, 125 ng of reverse primer, template DNA (5-50 ng), 1 μl of dNTP mix, 1 μl of PfuUltra HF DNA polymerase (2.5 U/μl), and nuclease-free water were added into microcentrifuge tube. This 50 μl reaction was amplified by below cycling conditions:

Table 4.1: The cycling conditions of site-directed mutagenesis.

Step	**Temperature**	**Time**
Initial Denaturation	95 °C	30 sec
12 cycles	95 °C	30 sec
	55 °C	1 min
	68 °C	1 min/kb of plasmid length

After temperature cycling, samples were placed on ice for 2 min. To digest the amplified products, 1 µl of Dpn I restriction enzyme was added into reaction and reaction was incubated at 37 °C for 1 hour. After this step, 1 µl of reaction was transformed into 50 µl of XL1-Blue competent cells. This reaction was incubated on ice for 30 min. Heat shock was applied at 42 °C for 45 sec and this reaction was incubated on ice for 2 min. 500 µl LB broth was added into this reaction and it was shaken at 37 °C for 1 hour. 40-100 µl of reaction was spread onto LB ampicillin agar plate, and plate was incubated at 37 °C overnight.

The dCRY Q311E gene was targeted by guide sequence: Forward primer of dCRY Q311E: 5’ CACATCACGGGAGAGTTGATCTGGCG 3’; Reverse primer of dCRY Q311E: 5’ CGCCAGATCAACTCTCCCGTGATGTG 3’.

4.11. Polymerase Chain Reaction (PCR)

4.11.1.PCR of MCF10A Wild Type and MCF10A *BMAL1* Knockout Set1

To set 50 µl PCR, 5 µl 10X standard Taq reaction buffer, 1 µl 10mM dNTPs, 1 µl 10 µM *BMAL1* T2 forward primer, 1 µl 10 µM *BMAL1* T2 reverse primer, 200 ng MCF10A WT DNA / MCF10A *BMAL1* KO Set1 DNA, 0.25 µl Taq DNA polymerase, water were added into microfuge tube. The microfuge tube was gently mixed and placed in thermal cycler. Sequences of forward and reverse primers for PCR were

BMAL1 T2 Forward: 5’ CCTGGCAGTGAGACCATTTT 3’; *BMAL1* T2 Reverse: 5’ TCACTTCATTTGGCATCACG 3’.

Table 4.2: The thermocycling conditions used in PCR of MCF10A.

Step	**Temperature**	**Time**
Initial Denaturation	95 °C	30 sec
35 Cycles	95 °C	20 sec
	60 °C	20 sec
	68 °C	30 sec
Final Extension	68 °C	2 min
Hold	4-10 °C	

4.11.2. PCR of pAc5.1 HisV5 dCRY Wild Type and pAc5.1 HisV5 dCRY Q311E

To set 50 µl PCR, 5 µl 10X standard Taq reaction buffer, 1 µl 10mM dNTPs, 1 µl 10 µM Q311E forward primer, 1 µl 10 µM Q311E reverse primer, 200 ng pAc5.1 HisV5 dCRY WT DNA / pAc5.1 HisV5 dCRY Q311E DNA, 0.25 µl Taq DNA polymerase, water were added into microfuge tube. The microfuge tube was gently mixed and placed in a thermal cycler. Sequences of forward and reverse primers for PCR were dCRY Q311E Forward: 5’ CACATCACGGGAGAGTTGATCTGGCG 3’; dCRY Q311E Reverse: 5’ CGCCAGATCAACTCTCCCGTGATGTG 3’;

Table 4.3: The thermocycling conditions used in PCR of pAc5.1 HisV5 dCRY.

Step	**Temperature**	**Time**
Initial Denaturation	95 °C	30 sec
35 Cycles	95 °C	20 sec
	60 °C	20 sec
	68 °C	40 sec
Final Extension	68 °C	2 min
Hold	4-10 °C	

4.11.3. Agarose Gel Electrophoresis

2% agarose gel was prepared with agarose powder using 1X TBE buffer and Pronosafe, then it was placed in horizontal electrophoresis tank loaded with 1X TBE buffer. To check the single-band of PCR products, the 100 bp DNA ladder and PCR products (MCF10A Wild Type and MCF10A *BMAL1* Knockout Set1, pAc5.1 HisV5 dCRY Wild Type and pAc5.1 HisV5 dCRY Q311E) were run on 2% agarose gel for control of amplification. The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.11.4. DNA Isolation from PCR Products

To extract DNA from PCR products (MCF10A Wild Type and MCF10A *BMAL1* Knockout Set1, pAc5.1 HisV5 dCRY Wild Type and pAc5.1 HisV5 dCRY Q311E), Macherey-Nagel Nucleospin Gel and PCR Clean-up kit was used. 150 μl NT1 was added 50 μl PCR product and 200 μl mixture was loaded into column. The column was centrifuged at 10000 rpm for 10 sec and 13000 rpm for 10 sec. The supernatant was discarded and 650 μl NT3 was added into column. The column was centrifuged at 10000 rpm for 10 sec and 13000 rpm for 10 sec. The supernatant was discarded and the column was centrifuged 13000 rpm for 1 min. The column was placed into new tube and 50 μl ddH_2O was added into the column. To elute the DNA column was centrifuged at 13000 rpm for 1 min. The column was discarded and DNA was placed into new tube. NanoDrop spectrophotometer was used to measure concentration and OD260 and OD280 values were utilized for quantify concentration and purity of the DNA.

4.12. Annealing of DNA

To form the heteroduplex DNA, MCF10A Wild Type and MCF10A *BMAL1* Knockout Set1, pAc5.1 HisV5 dCRY Wild Type and pAc5.1 HisV5 dCRY Q311E DNA were denatured and reannealed by annealing protocol [32]. 5 μl 10X NEBuffer 2 was added into 87,5 ng WT DNA and 87,5 ng mutant DNA sample and this reaction

was completed with ddH_2O to 50 µl volume. Likewise, this reaction was prepared with 175ng WT DNA. To anneal these reactions, optimized annealing conditions were followed. These optimized annealing conditions are below.

Table 4.4: The annealing protocol used in this study.

Temperature	**Time**
95 °C	10 min
95-85 °C	-2 °C/sec
85 °C	1 min
85-75 °C	-0.1 °C/sec
75 °C	1 min
75-65 °C	-0.1 °C/sec
65 °C	1 min
65-55 °C	-0.1 °C/sec
55 °C	1 min
55-45 °C	-0.1 °C/sec
45 °C	1 min
45-35 °C	-0.1 °C/sec
35 °C	1 min
35-25 °C	-0.1 °C/sec
25 °C	1 min
25-4 °C	-0.1 °C/sec
4 °C	hold

4.12.1. Agarose Gel Electrophoresis

To check for equal amount of annealed WT and mutant DNA, the 100 bp DNA ladder and annealed WT and mutant DNA products were run on 2% agarose gel for control of annealing. The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

4.13. T7 Endonuclease 1 (T7E1) Cleavage Assay

To determine the efficient T7 Endonuclease 1 activity, amount of used enzyme and annealing conditions were optimized. According to these optimizations, 1,5 µl T7E1 was added into 20 µl annealed product. This reaction was incubated at 37 °C for 2 hours. Annealed products were digested with T7E1.

After T7 Endonuclease incubation, to control the digestion activity of annealed WT and mutant DNA, the 100 bp DNA ladder and annealed products were run on 2% agarose gel. The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System. To estimate the frequency of indels, densities of the PCR product and cleaved DNA fragments in gel were measured. Indel (%) was calculated according to below formula [32]:

Table 4.5: The formula is used for indel (%) calculation.

$$\text{Indel (\%)} = 100 \times \left(1 - \sqrt{1 - \frac{b+c}{a+b+c}}\right)$$

4.14. High Resolution Melting (HRM) Analysis

0.5 µl MutS-His was added into 15 µl annealed product. To create control group, 0.5 µl dialysis buffer was added into 15 µl annealed product separately. Then 5 µl (5X diluted) EvaGreen Dye was added into annealed product and MutS enzyme/dialysis buffer mixture. Amount of EvaGreen dye was optimized for efficient fluorescent signal. Conditions of melt curve stage was optimized for high resolution signal in melt curve analysis.

Table 4.6: The melt curve protocol used in this study.

Temperature	Time
37°C	15 min
95°C	15 sec
60°C	1 min
95°C	15 sec

4.15. Rapid Agarose Gel Electrophoretic Mobility Shift Assay

1-2-4 μl MutS-His was added into 20 μl heteroduplex products. To create a control group, 1-2-4 μl dialysis buffer was added into 20 μl heteroduplex products separately. For binding activity, this reaction was incubated at 37 °C for 15 minutes. 1% agarose gel was prepared with agarose powder using 0.5X TB buffer. The gel was poured into tray approximately in 6-8 mm thick. The gel was placed in a horizontal electrophoresis tank including 0.5X TB buffer. The samples were run in this gel at 175V for 10 min. After electrophoresis, the gel was stained for 40 min with shaking in 40 ml Pronosafe, which was prepared by diluting 2 μl Pronosafe stock into 40 ml 0.5X TB buffer. The gel was destained in 40 ml 0.5X TB buffer for 15 min with shaking to reduce background. The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System.

5. RESULTS

5.1. Purification of MutS, TtAgo-MutS and Reengineered MutS-Cas9 Proteins

MutS and Cas9 protein were reengineered using molecular cloning techniques to generate a chimeric enzyme for mutation detection. The design method of MutS-Cas9 protein was shown in the methods chapter and reengineered MutS-Cas9 protein was represented schematically in figure 5.1.

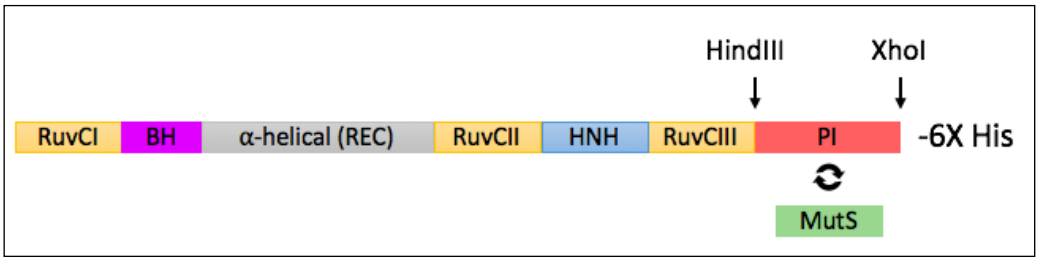

Figure 5.1: The schematic representation of reengineered MutS-Cas9 protein.

MutS and TtAgo-MutS proteins which will be used in later experiments were designed according to a previously published study [20]. Coomassie staining and Western Blot analysis of MutS, TtAgo-MutS and MutS-Cas9 proteins were shown (Figure 5.2 and Figure 5.4). MutS, TtAgo-MutS and MutS-Cas9 proteins were expressed in *E. coli* strain BL21. Expressed and purified proteins were shown (Figure 5.3 and Figure 5.5).

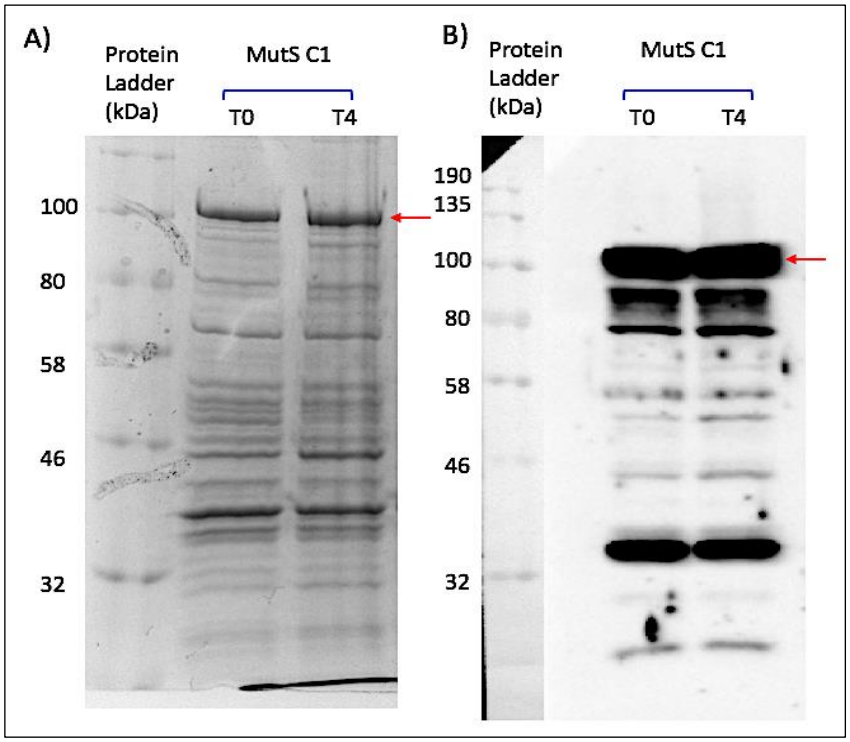

Figure 5.2: Coomassie Staining (shown in A) and Western Blot (shown in B) analysis of BL21 pET21b MutS C1 at 8% SDS-PAGE gel.

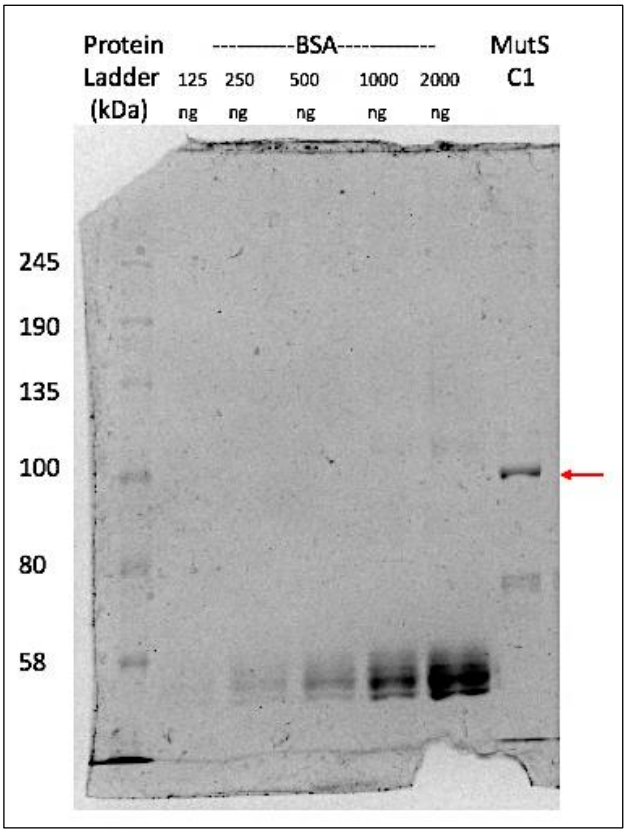

Figure 5.3: Coomassie Staining analysis of purified His-tag protein BL21 pET21b MutS C1 at 6% SDS-PAGE gel.

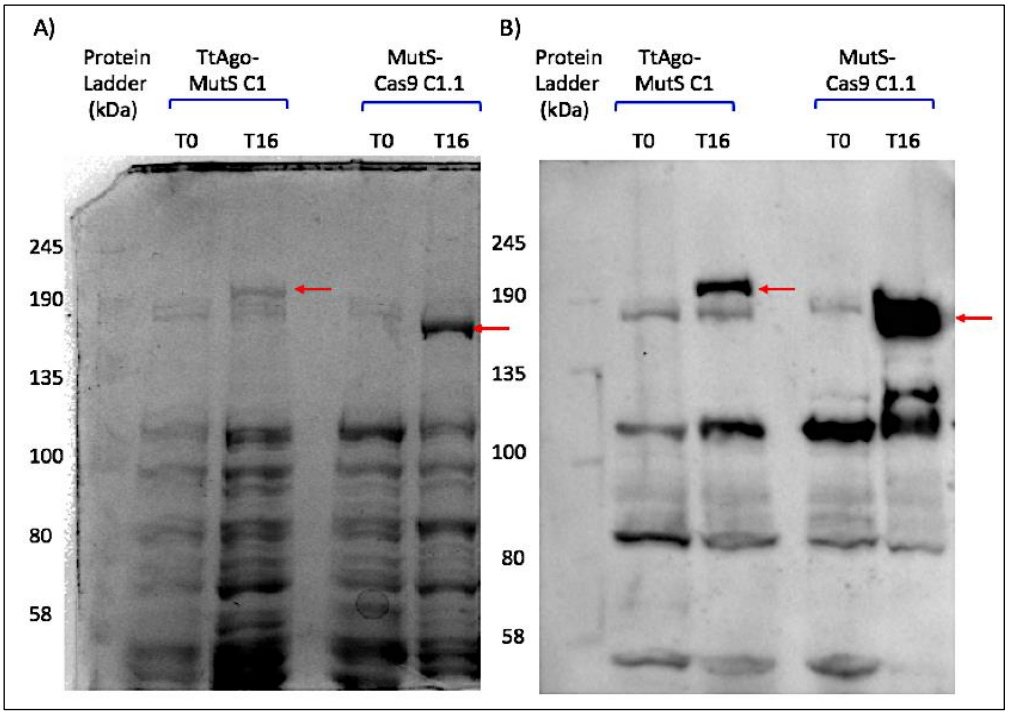

Figure 5.4: Coomassie Staining (shown in A) and Western Blot (shown in B) analysis of BL21 pET21b TtAgo-MutS C1 and BL21 pET21b MutS-Cas9 C1.1 at 6% SDS-PAGE gel.

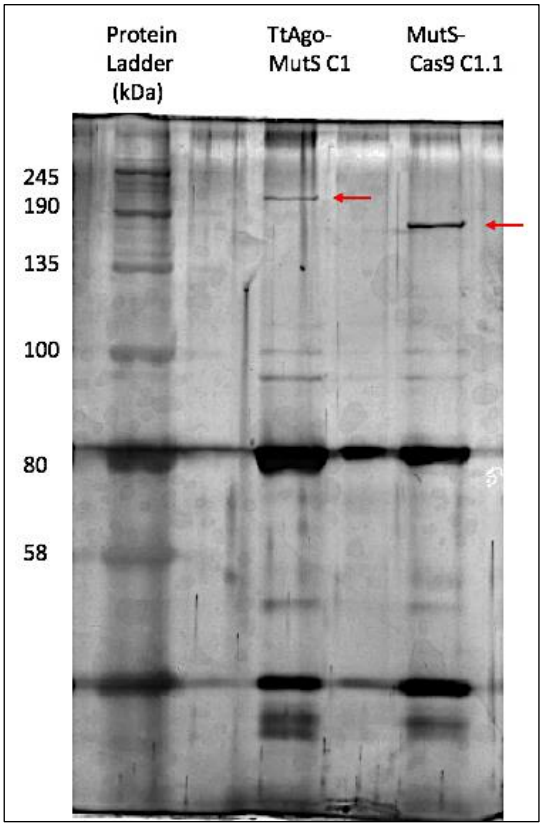

Figure 5.5: Silver Staining analysis of purified His-tag proteins BL21 pET21b TtAgo-MutS C1 and BL21 pET21b MutS-Cas9 C1.1 at 7% SDS-PAGE gel.

5.2. Analysis of PCR-amplified Wild Type DNA and Mutant DNA Products

The first step of process is identical for mutation detection methods used in this study. In order to perform this first step, 200 ng wild type and mutant genomic DNA sequences were PCR-amplified by Taq DNA Polymerase using specific forward and reverse primers. These primers were designed to cover a genomic region which is expected to be mutated by CRISPR-Cas9 system in our edited cell lines (see Methods section for the details). PCR products were run on 2% agarose gels and images were captured and analyzed using at Bio-Rad ChemiDoc™ XRS+ System. Agarose gel image of PCR products from genomic DNAs isolated from wild type (MCF10A WT DNA) or its *BMAL1* knock clone (MCF10A *BMAL1* KO Set1) was shown (Figure 5.6).

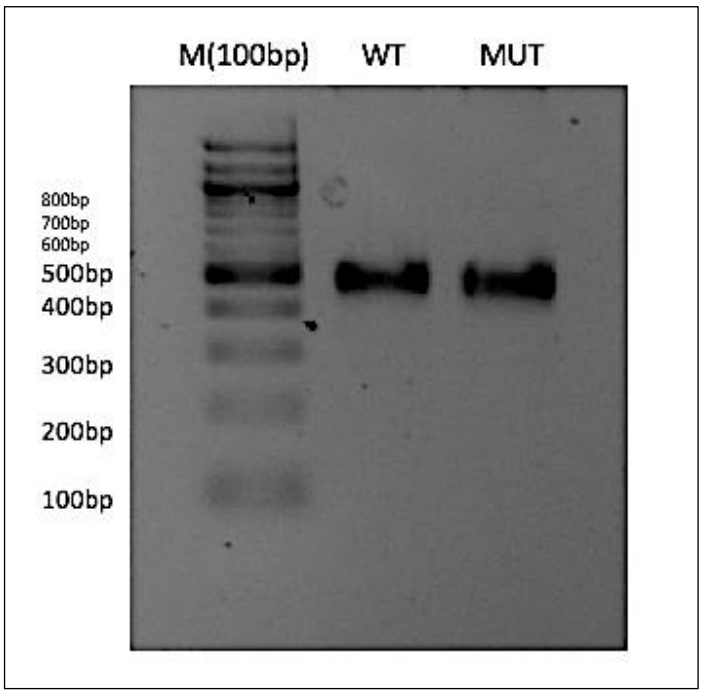

Figure 5.6: 2% agarose gel electrophoresis of PCR products amplified from MCF10A set. MCF10A WT DNA (abbreviated as WT) and MCF10A *BMAL1* KO Set1 DNA (abbreviated as MUT).

Agarose gel image of PCR-amplified products using oligos specific for dCRY (*Drosophila* Cryptochrome) from plasmids was shown (Figure 5.7). pAc5.1 HisV5

dCRY plasmid harbors wild-type dCRY coding sequence while pAc5.1 HisV5 dCRY Q311E DNA harbors a mutant dCRY coding squence (Q311E).

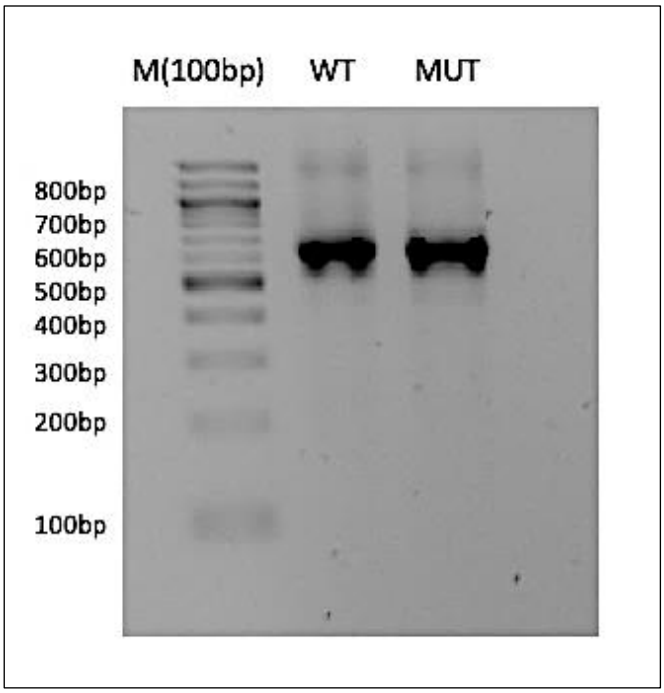

Figure 5.7: 2% agarose gel electrophoresis of PCR products amplified from pAc5.1 HisV5 dCRY set. pAc5.1 HisV5 dCRY WT DNA (abbreviated as WT) and pAc5.1 HisV5 dCRY Q311E DNA (abbreviated as MUT).

5.3. Formation of Heteroduplex DNA

The homoduplexes and heteroduplexes were formed to use in further mutation detection assays. For this purpose, PCR-amplified wild type and mutant DNA were purified from agarose gel. Equal amounts of purified wild type (175 ng) and mutant DNA (175 ng) were denatured and reannealed to form heteroduplexes. Control homoduplexes were formed by annealing equal amounts of purified wild type DNA (350 ng) only. Annealing conditions were shown in methods chapter. Then annealed products were analyzed by agarose gel electrophoresis to control annealing. The homoduplex wild type and heteroduplex mutant annealed products containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA were shown (Figure 5.8). Another annealed products containing pAc5.1 HisV5 dCRY WT DNA and pAc5.1 HisV5 dCRY Q311E DNA were shown (Figure 5.9).

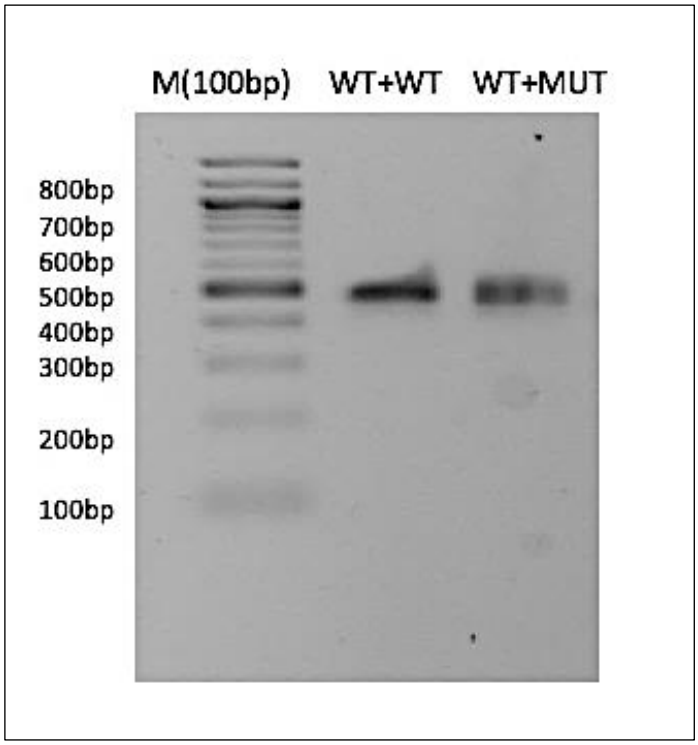

Figure 5.8: 2% agarose gel electrophoresis of annealed MCF10A set. Homoduplex MCF10A WT DNA (shown as WT+WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (shown as WT+MUT).

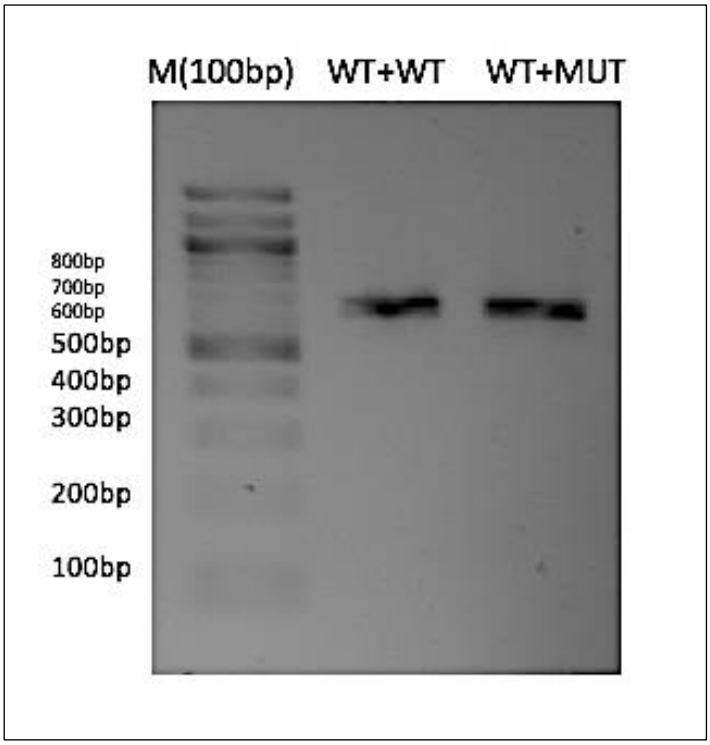

Figure 5.9: 2% agarose gel electrophoresis of annealed pAc5.1 HisV5 dCRY set. Homoduplex pAc5.1 HisV5 dCRY WT DNA (shown as WT+WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (shown as WT+MUT).

5.4. T7 Endonuclease 1 (T7E1) Cleavage Assay

In order to detect mutations, 20 µl of homoduplex and heteroduplex samples were treated with 1.5 µl of T7 Endonuclease 1 (T7E1) at 37 °C for 2 hours. The digested and undigested products were analyzed by agarose gel electrophoresis (Figure 5.10 and Figure 5.11) The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System. To evaluate mutation efficiency, densities of the PCR products and cleaved DNA fragments in gel were measured. Indel (%) was calculated according to shown formula in methods chapter. These results shows T7E1 has less efficiency for single base mutation detection.

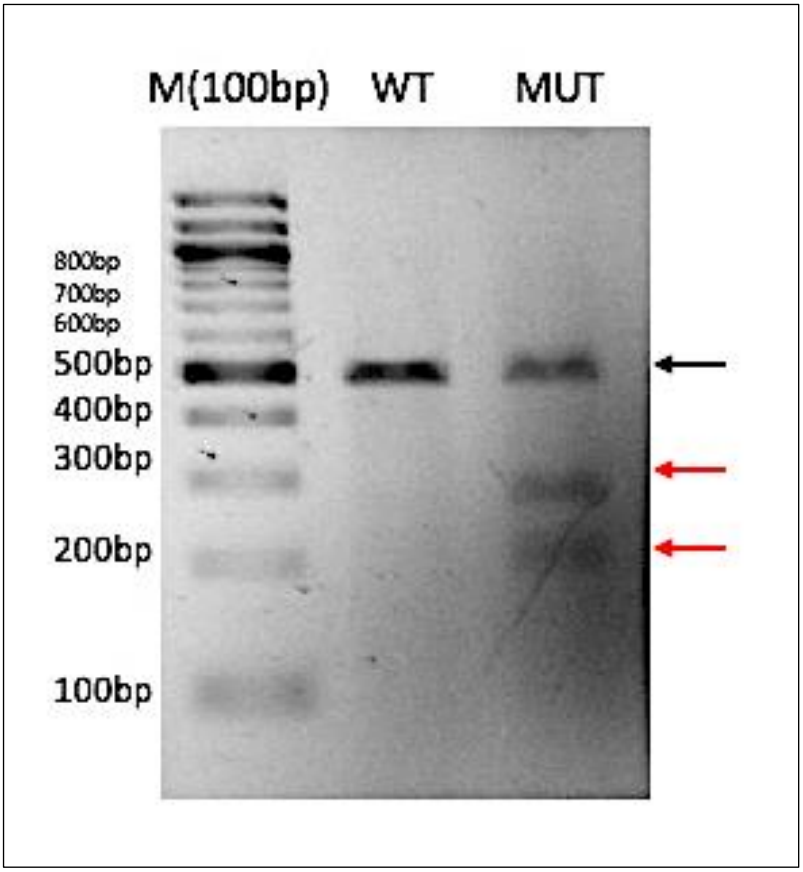

Figure 5.10: 2% agarose gel electrophoresis of T7E1 digestion products (MCF10A set). Homoduplex MCF10A WT DNA (shown as WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (shown as MUT) were incubated with T7E1. Red arrows show digested DNA fragments and black arrow shows undigested annealed product. Indel value is 48.63%.

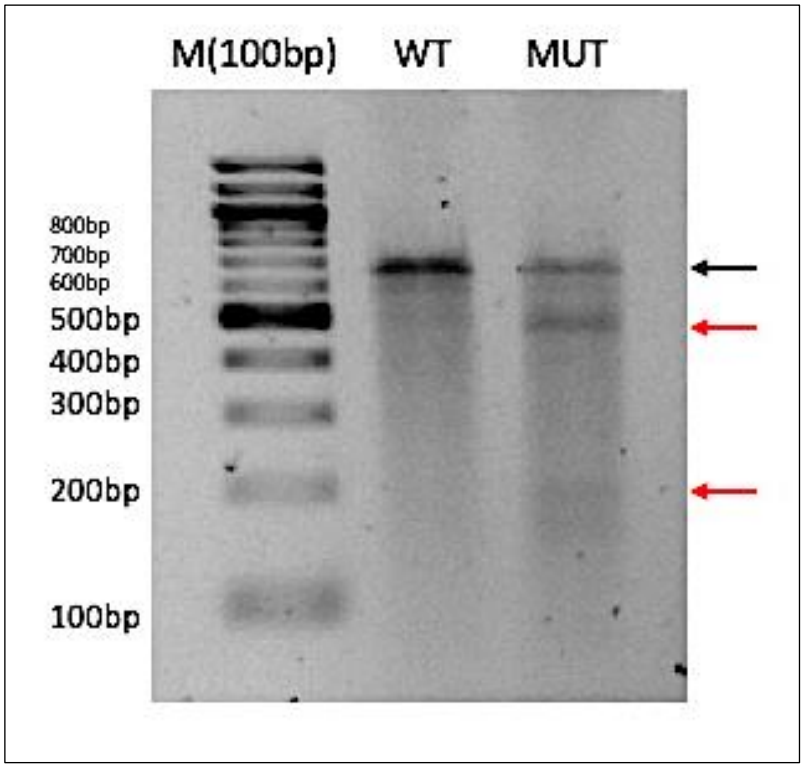

Figure 5.11: 2% agarose gel electrophoresis of T7E1 digestion products (pAc5.1 HisV5 dCRY set). Homoduplex pAc5.1 HisV5 dCRY WT DNA (shown as WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (shown as MUT) were incubated with T7E1. Red arrows show digested DNA fragments and black arrow shows undigested annealed product. Indel value is 9%.

5.5. High Resolution Melting (HRM) Analysis

HRM analysis is based on the evaluation of high-resolution denaturation, which can reflect the existence of heteroduplex and homoduplex DNA in a sample. In this way, DNA mutations were attempted to be identified by comparing unknown DNA samples with those known to be normal. Denaturation of DNA samples was shown with normalized and derivative curves. Different genotypes exhibit different denaturation behaviors. Heteroduplexes have lower Tm than homoduplexes due to their instability. To perform HRM, experiment conditions (amount of EvaGreen dye and annealed product, increase rate of temperature) were optimized and this optimization was explained in methods chapter. In normalized melt curve, Tm (midway in the melting curve) indicates the temperature at which 50% of the DNA is denatured (Figure 5.13 and Figure 5.15).

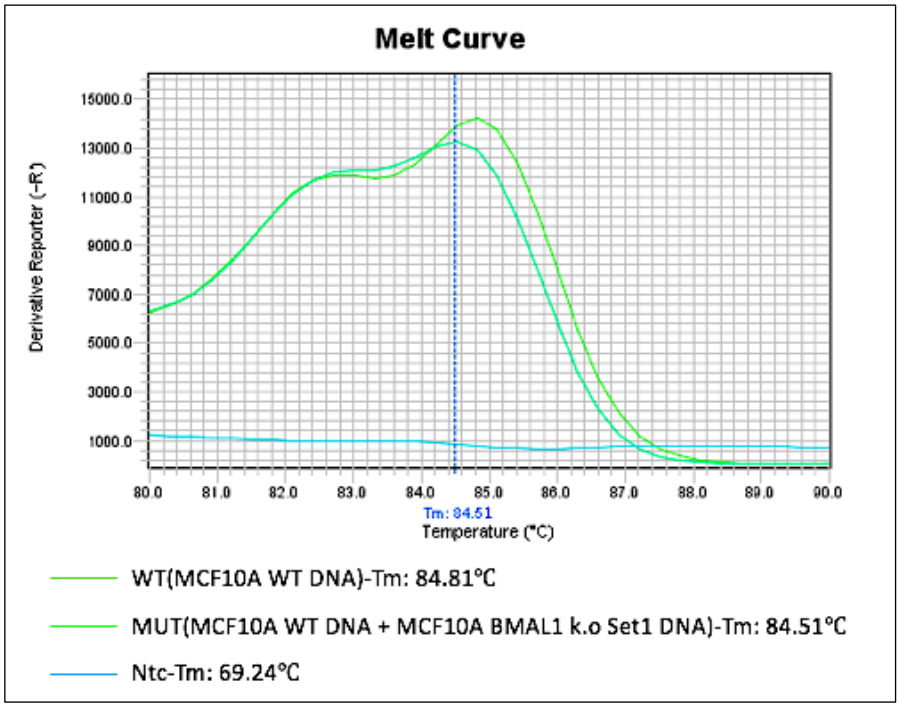

Figure 5.12: High resolution melting curve of MCF10A set (derivative). Derivative melting curve plots of homoduplex MCF10A WT DNA (abbreviated as WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (abbreviated as MUT). The Tm of the WT product is 84.81°C and the MUT product is 84.51°C.

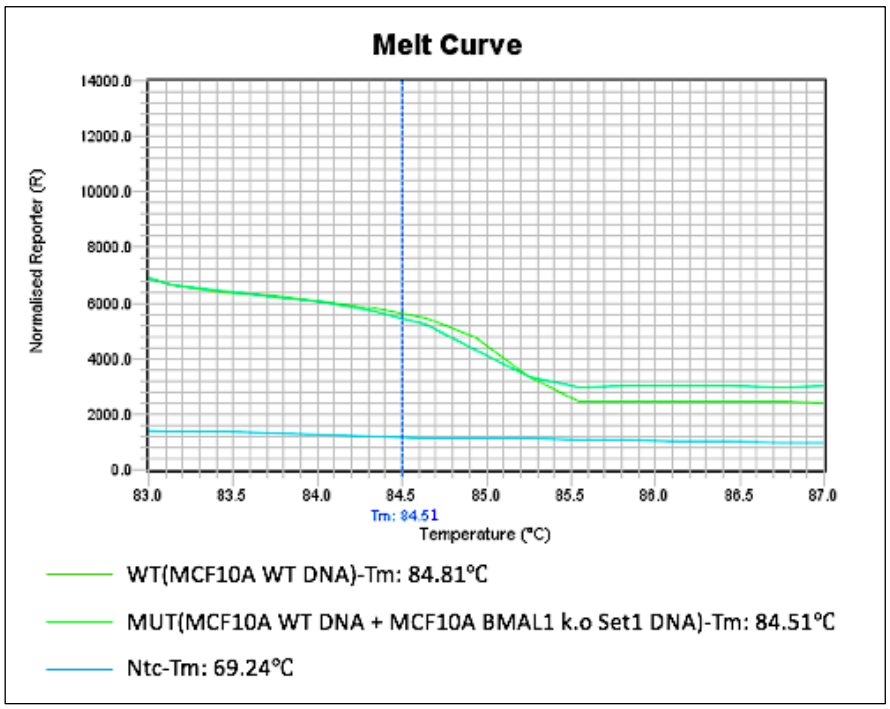

Figure 5.13: High resolution melting curve of MCF10A set (normalized). Normalized melting curve plots of homoduplex MCF10A WT DNA (abbreviated as WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (abbreviated as MUT). The Tm of the WT product is 84.81°C and the MUT product is 84.51°C.

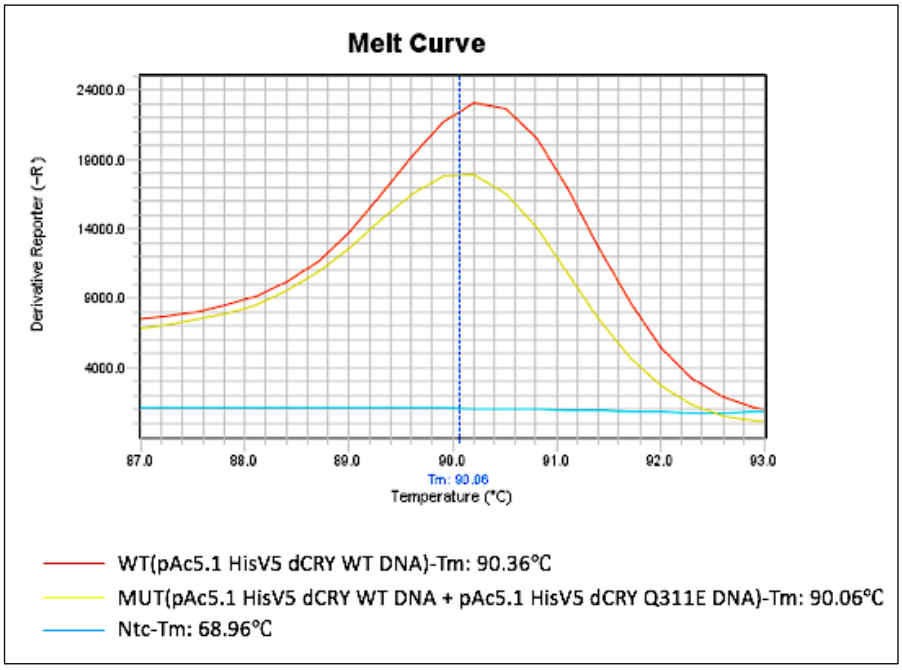

Figure 5.14: High resolution melting curve of pAc5.1 HisV5 dCRY set (derivative). Derivative melting curve plots of homoduplex pAc5.1 HisV5 dCRY WT DNA (abbreviated as WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (abbreviated as MUT). The Tm of the WT product is 90.36°C and the MUT product is 90.06°C.

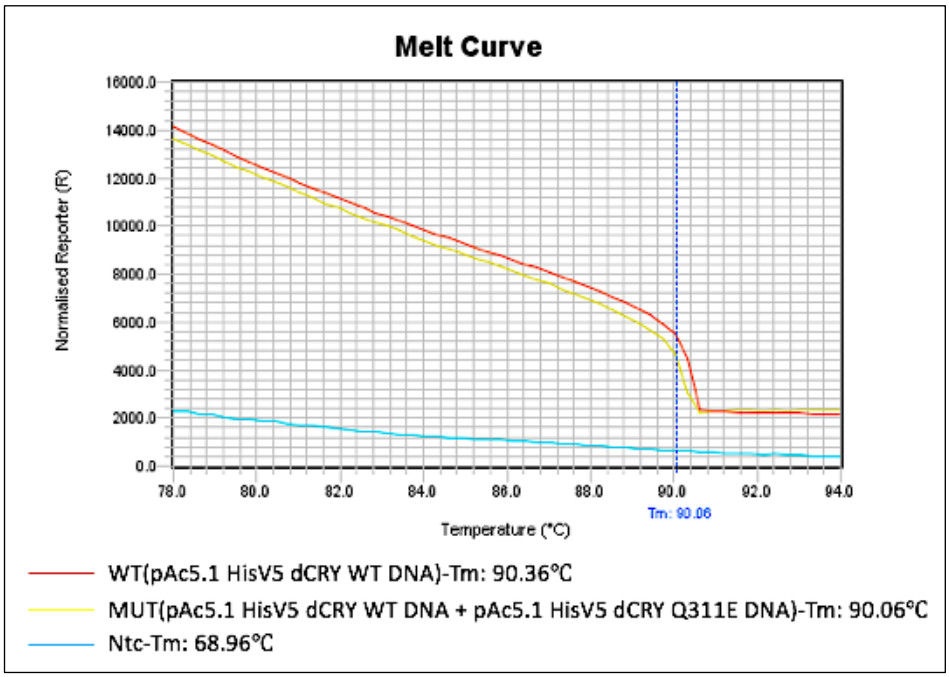

Figure 5.15: High resolution melting curve of pAc5.1 HisV5 dCRY set (normalized). Normalized melting curve plots of homoduplex pAc5.1 HisV5 dCRY WT DNA (abbreviated as WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (abbreviated as MUT). The Tm of the WT product is 90.36°C and the MUT product is 90.06°C.

The pAc5.1 HisV5 dCRY Q311E DNA has point mutation which is occured by C/G base change. This results 0.3°C Tm curve shift (Figure 5.14). The high resolution melting curve of MCF10A set (derivative) was shown (Figure 5.12).

Additionally, to observe binding of MutS-DNA, HRM analysis was performed adding MutS, and dialysis buffer (as a control group). MutS binds to mismatched DNA and this binding causes more difficult denaturation during the heating. Therefore, different Tm shift profile can be observed.

The high resolution melting curves of MCF10A set which include MutS (Figure 5.17) and dialysis buffer (Figure 5.16) were shown. The high resolution melting curves of pAc5.1 HisV5 dCRY set which include MutS (Figure 5.19) and dialysis buffer (Figure 5.18) were shown.

Control group (contains dialysis buffer) of both two sets has similar Tm shift profile with normal HRM curve. MutS added group of pAc5.1 HisV5 dCRY set has similar Tm shift profile with normal HRM curve. However, MutS added group of MCF10A set has different Tm shift profile with normal HRM curve. The binding of MutS-heteroduplex MCF10A DNA causes high Tm compared with normal HRM curve (not including MutS), thus Tm of heteroduplex MCF10A DNA is not lower than wild type MCF10A DNA.

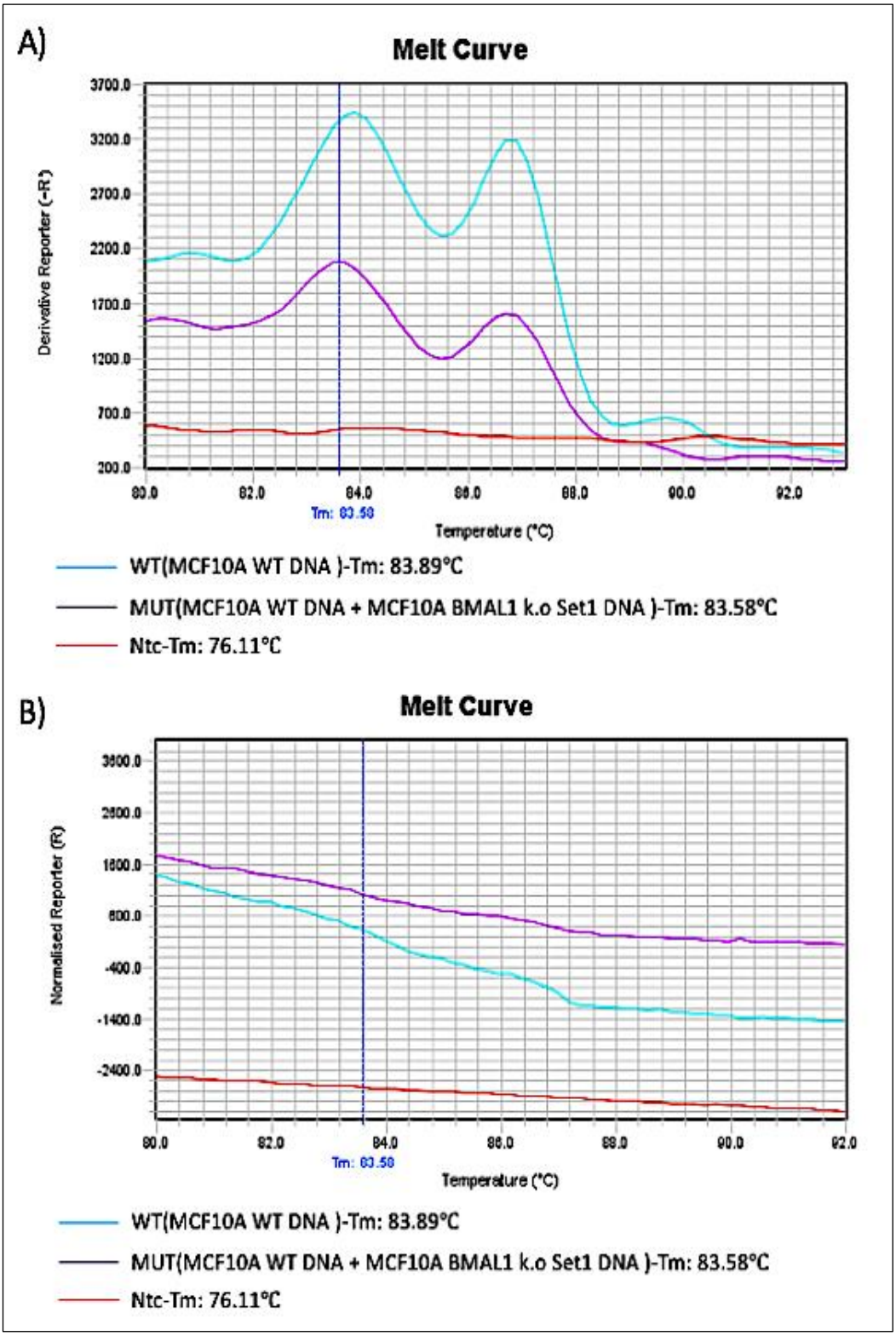

Figure 5.16: High resolution melting curve of MCF10A set including dialysis buffer. Derivative (shown in A) and normalized (shown in B) melting curve plots of homoduplex MCF10A WT DNA (abbreviated as WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (abbreviated as MUT). The Tm of the WT product is 88.89°C and the MUT product is 88.58°C.

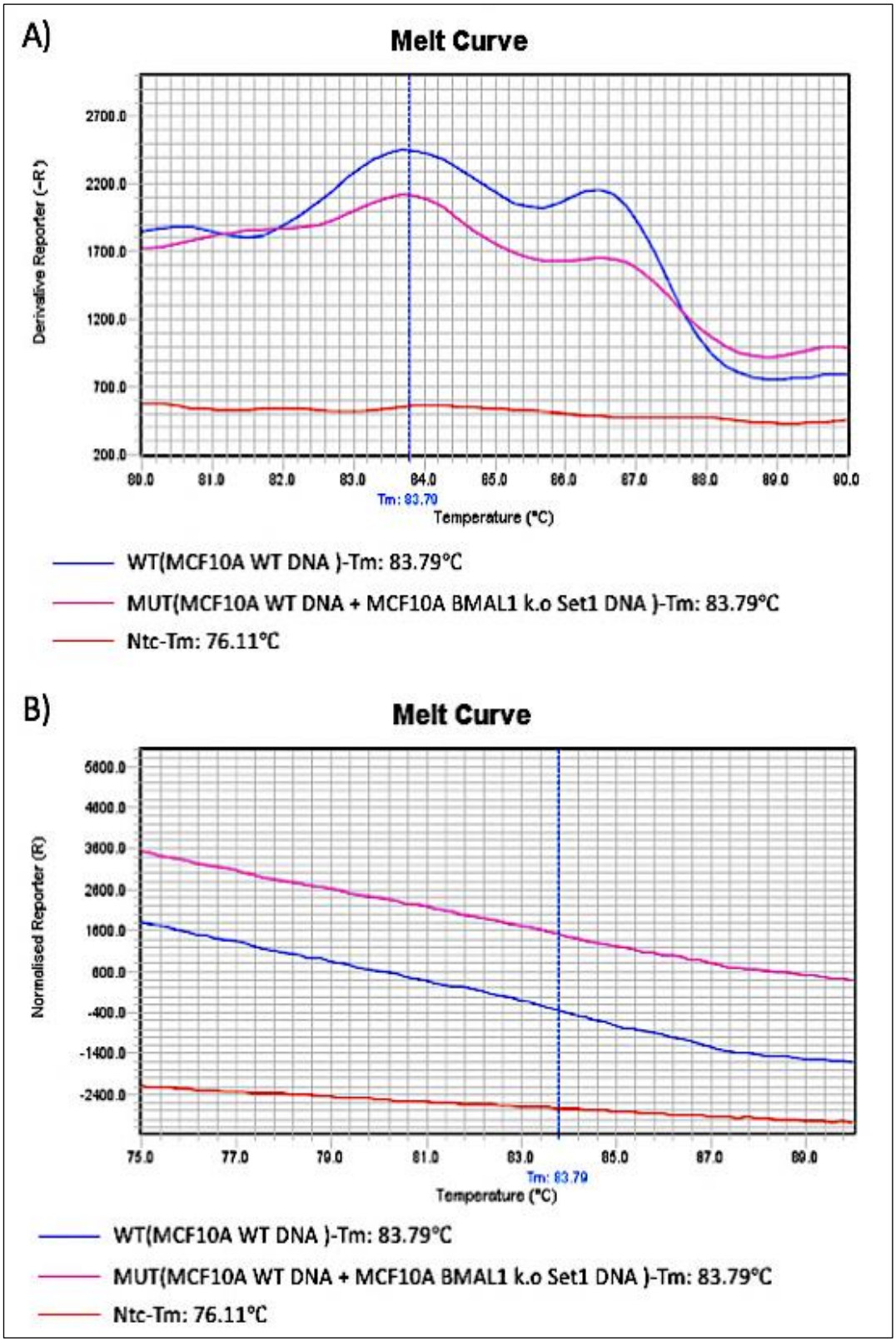

Figure 5.17: High resolution melting curve of MCF10A set including MutS. Derivative (shown in A) and normalized (shown in B) melting curve plots of homoduplex MCF10A WT DNA (abbreviated as WT) and heteroduplex MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA (abbreviated as MUT). The Tm of the WT product is 88.79°C and the MUT product is 88.79°C.

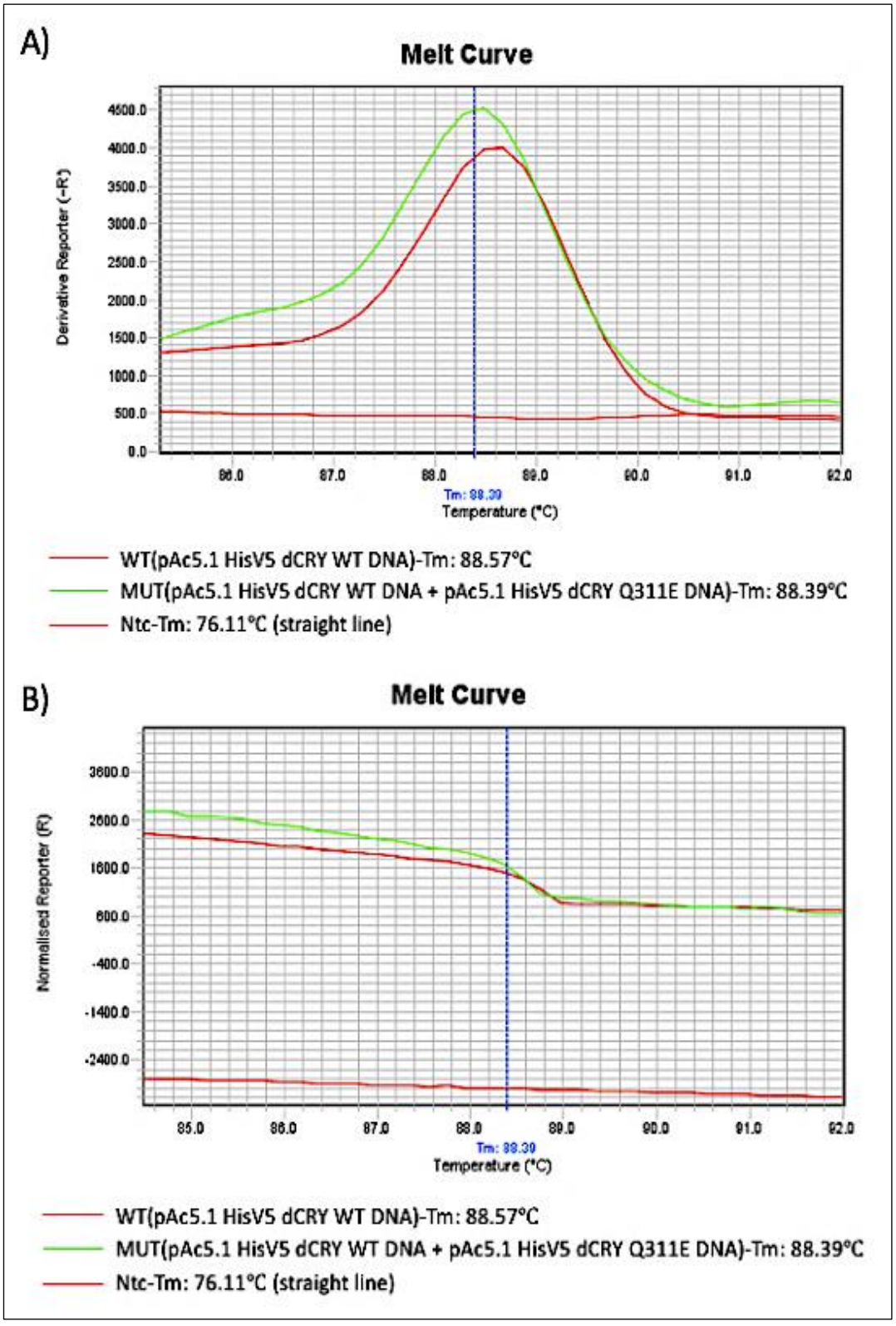

Figure 5.18: High resolution melting curve of pAc5.1 HisV5 dCRY set including dialysis buffer. Derivative (shown in A) and normalized (shown in B) melting curve plots of homoduplex pAc5.1 HisV5 dCRY WT DNA (abbreviated as WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (abbreviated as MUT). The Tm of the WT product is 88.57°C and the MUT product is 88.39°C.

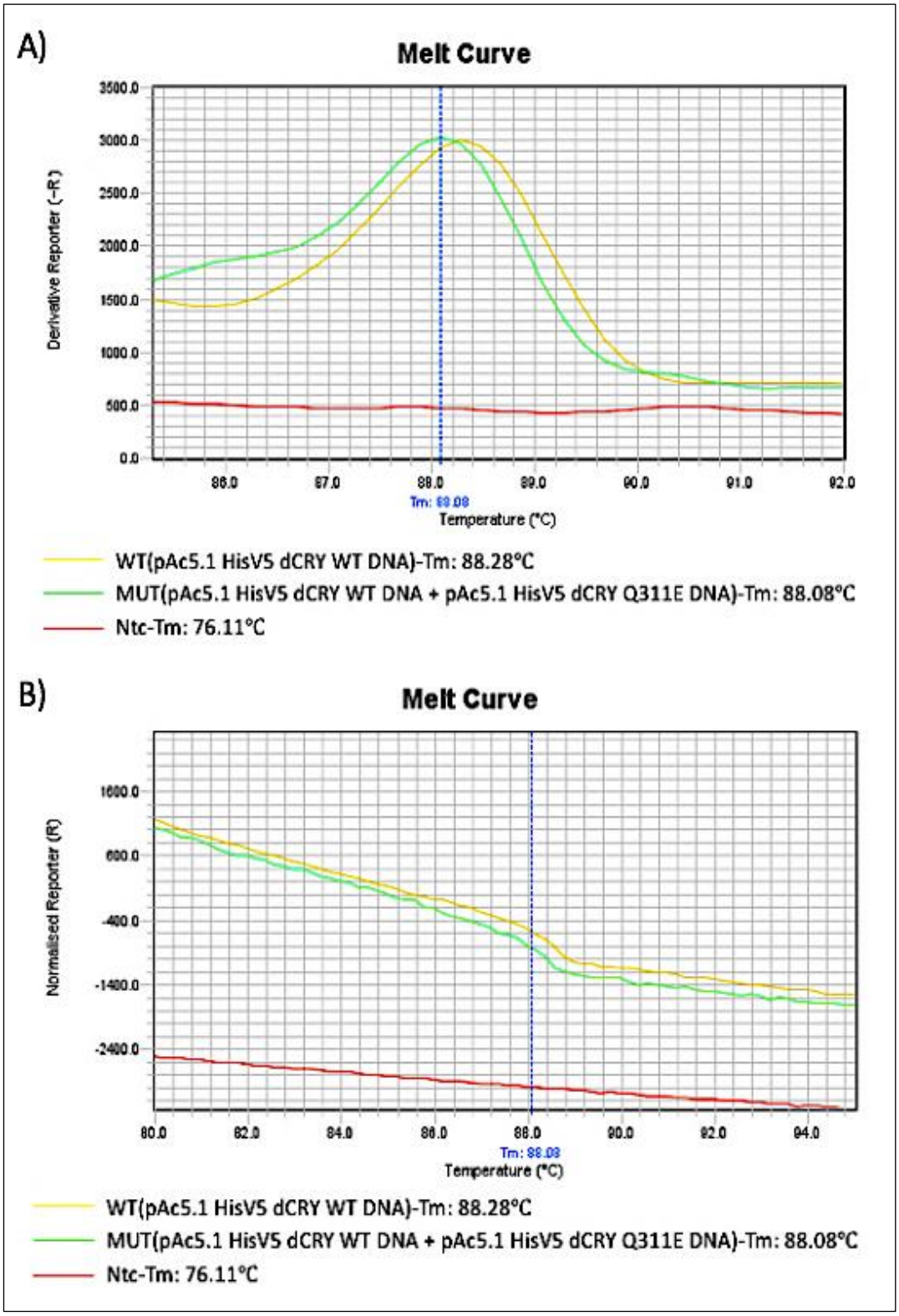

Figure 5.19: High resolution melting curve of pAc5.1 HisV5 dCRY set including MutS. Derivative (shown in A) and normalized (shown in B) melting curve plots of homoduplex pAc5.1 HisV5 dCRY WT DNA (abbreviated as WT) and heteroduplex pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA (abbreviated as MUT). The Tm of the WT product is 88.28°C and the MUT product is 88.08°C.

5.6. Rapid Agarose Gel Electrophoretic Mobility Shift Assay

To detect mutation using MutS, mutant DNA-MutS interaction was analyzed by rapid agarose gel electrophoretic mobility shift assay. For this purpose, 20 µl of heteroduplex products were incubated with 1-2-4 µl MutS-His. 20 µl of heteroduplex products were incubated with 1-2-4 µl dialysis buffer to create control group. For MutS binding activity, these samples was incubated at 37 °C for 15 min. Samples were analyzed by agarose gel electrophoresis using RAEMSA protocol, which is shown in methods chapter. The gel was imaged at Bio-Rad ChemiDoc™ XRS+ System. According to these results, MutS has mutant DNA binding activity and this activity causes retardation of MutS-mutant DNA complex in gel (Figure 5.20 and Figure 5.22). This assay was repeated three times. The MutS binding activity was measured and graphics were created according to this activity. The MutS binding activity increases depend on amount of enzyme (Figure 5.21 and Figure 5.23).

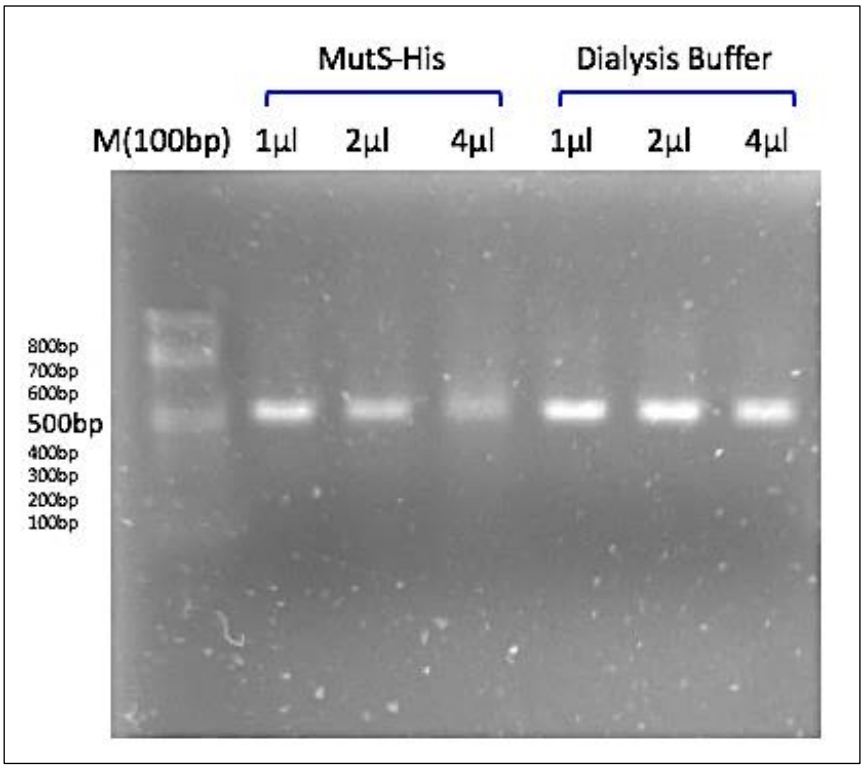

Figure 5.20: 1% agarose gel electrophoresis of MutS binding activity for MCF10A set. MutS binds to heteroduplex product (containing MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA) and causes formation of weak bands due to retardation. Three bands on right is control group.

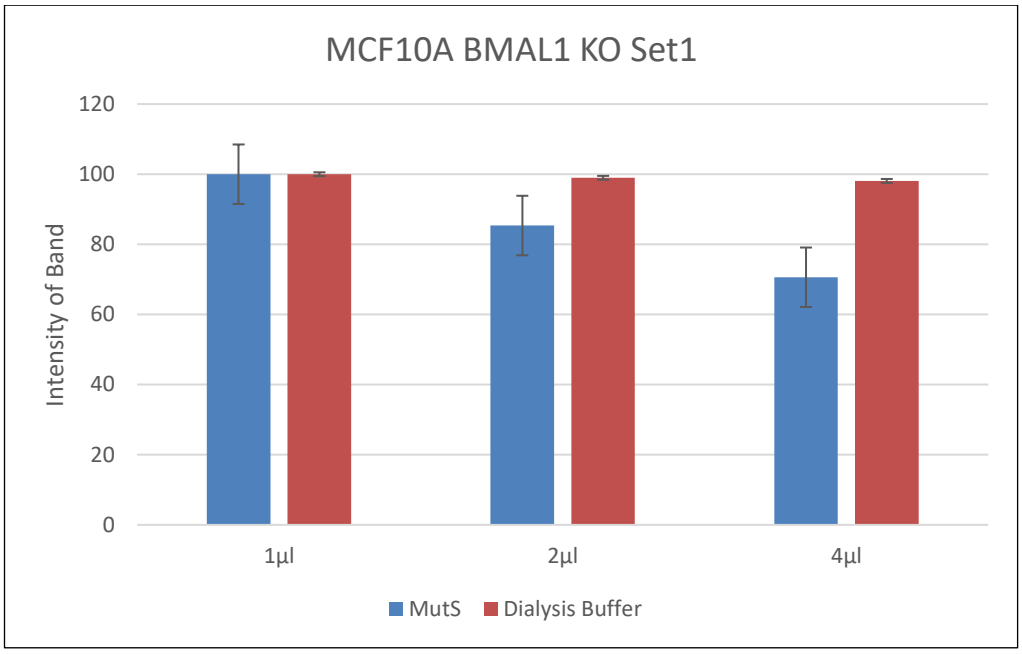

Figure 5.21: Quantification of MutS binding activity for MCF10A set. Graph shows the decreased band intensity due to MutS-heteroduplex product (containing MCF10A WT DNA + MCF10A *BMAL1* k.o Set1 DNA) binding. Mutation detection efficiency of MutS increases depend on amount of enzyme.

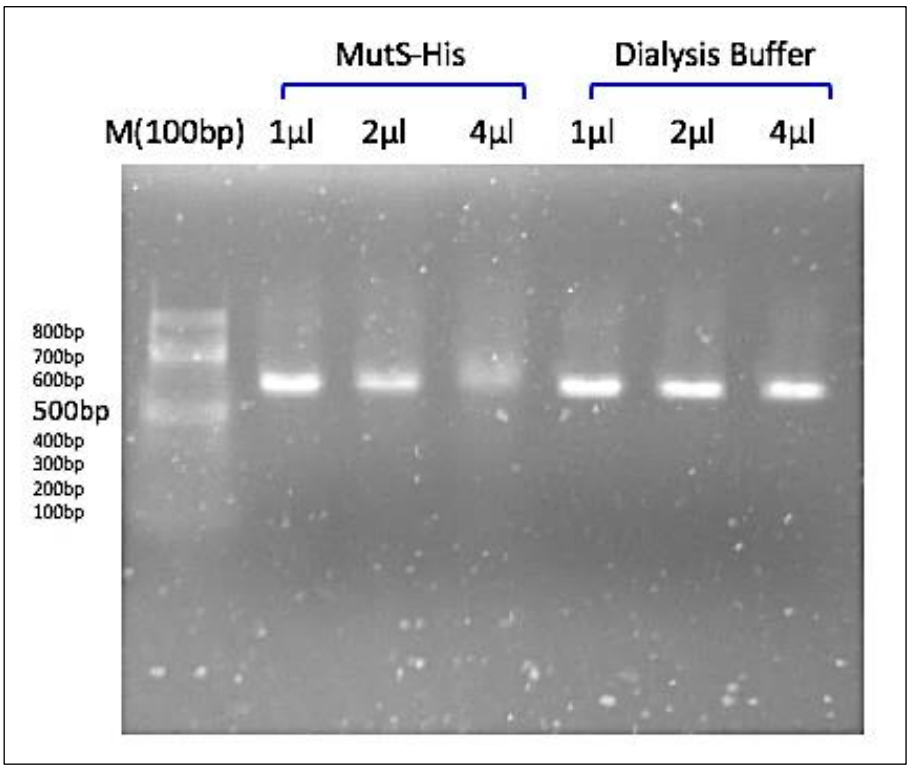

Figure 5.22: 1% agarose gel electrophoresis of MutS binding activity for pAc5.1 HisV5 dCRY set. MutS binds to heteroduplex product (containing pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA) and causes formation of weak bands due to retardation. Three bands on right is control group.

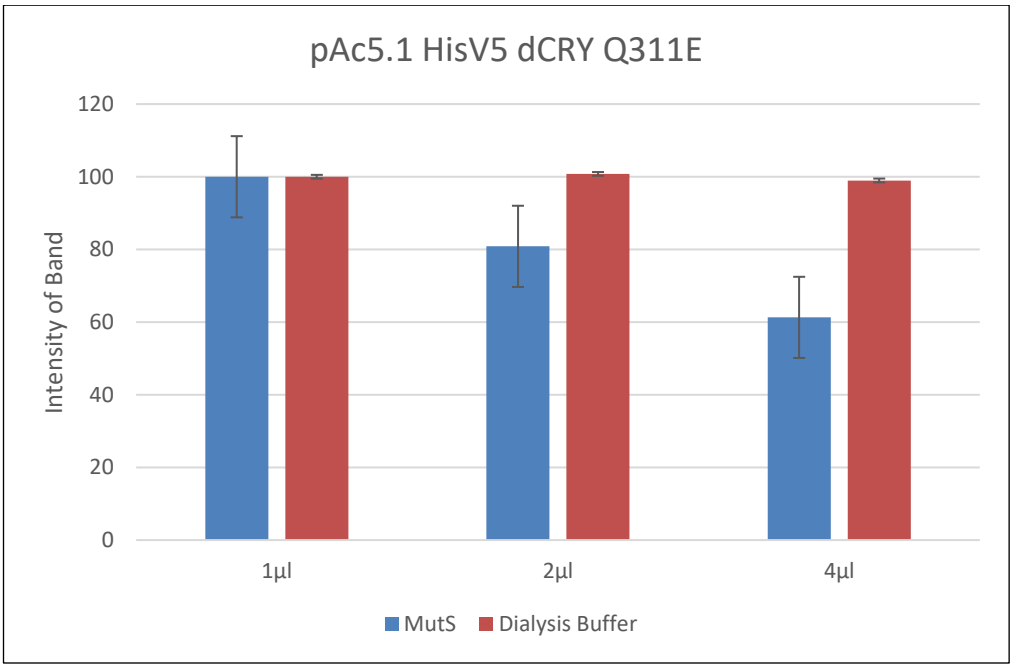

Figure 5.23: Quantification of MutS binding activity for pAc5.1 HisV5 dCRY set. Graph shows the decreased band intensity due to MutS-heteroduplex product (containing pAc5.1 HisV5 dCRY WT DNA + pAc5.1 HisV5 dCRY Q311E DNA) binding. Mutation detection efficiency of MutS increases depend on amount of enzyme.

5.7. The Cleavage Activity of TtAgo-MutS Protein and MutS-Cas9 Protein

To understand the cleavage activity of reengineered MutS-Cas9 protein, 15 µl of homoduplex and heteroduplex products were incubated with 5 µl of MutS-Cas9 protein at 37 °C for 2 and 6 hours. The control groups were prepared by incubation with 5 µl of TtAgo-MutS protein to compare cleavage activity. Then MutS-Cas9 protein and TtAgo-MutS protein treated products were analyzed by 2% agarose gel electrophoresis. The cleavage activity of MutS-Cas9 protein and TtAgo-MutS protein causes degradation of heteroduplex product and results in formation of weak bands. The agarose gel image of the cleavage activity of MutS-Cas9 protein and TtAgo-MutS protein was shown (Figure 5.24). The quantification of the cleavage activity of MutS-Cas9 protein and TtAgo-MutS protein was shown (Figure 5.25 and Figure 5.26).

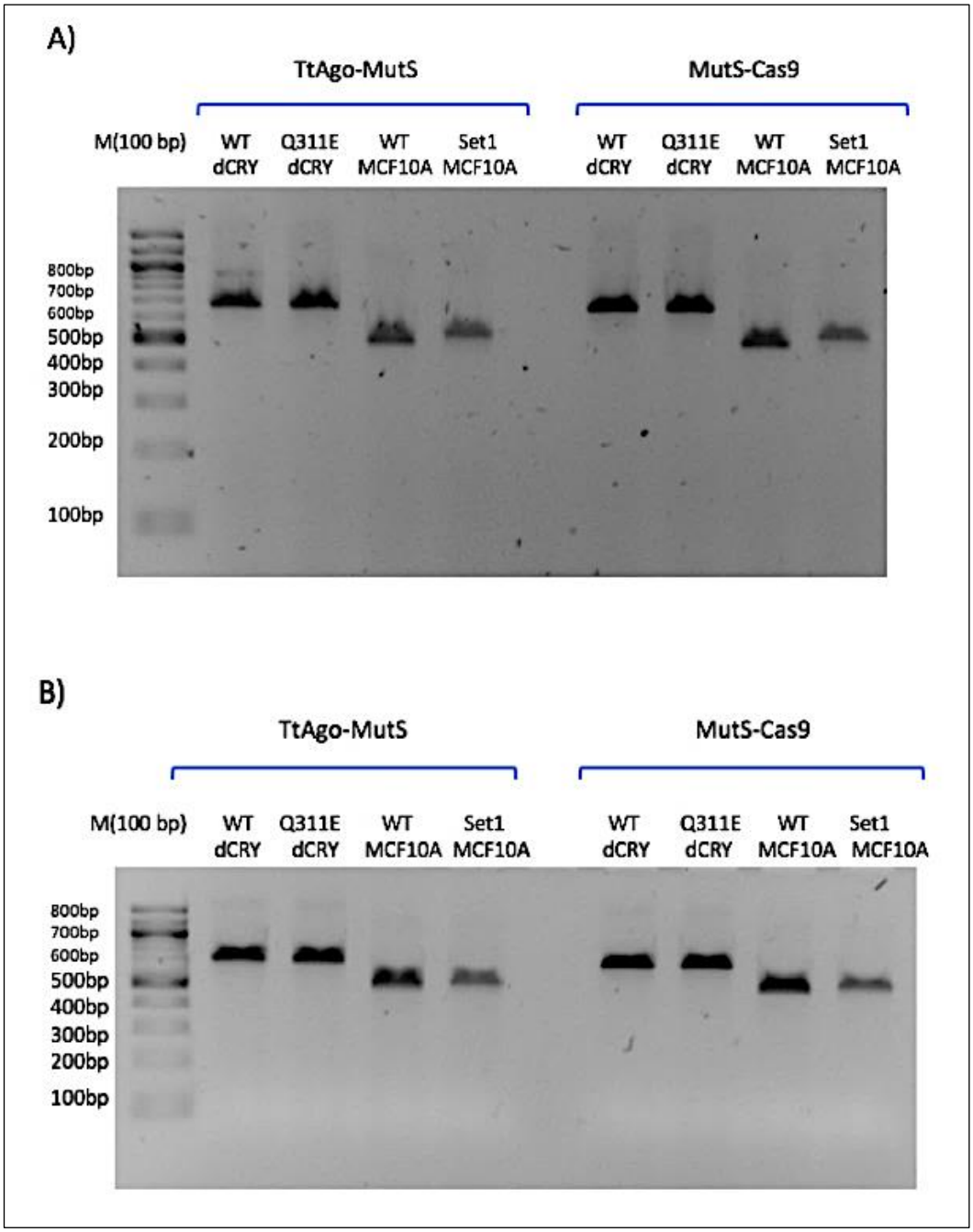

Figure 5.24: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 2 and 6 hours. The heteroduplex and homoduplex products were incubated with TtAgo-MutS and MutS-Cas9 at 37 °C for 2 hours (shown in A) and 6 hours (shown in B). WT dCRY is homoduplex pAc5.1 HisV5 dCRY WT DNA, Q311E dCRY is heteroduplex that contains pAc5.1 HisV5 dCRY WT DNA and pAc5.1 HisV5 dCRY Q311E DNA. WT MCF10A is homoduplex MCF10A WT DNA, Set1 MCF10A is heteroduplex that contains MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA. The heteroduplex product which is shown as Set1 MCF10A was degraded by TtAgo-MutS and MutS-Cas9.

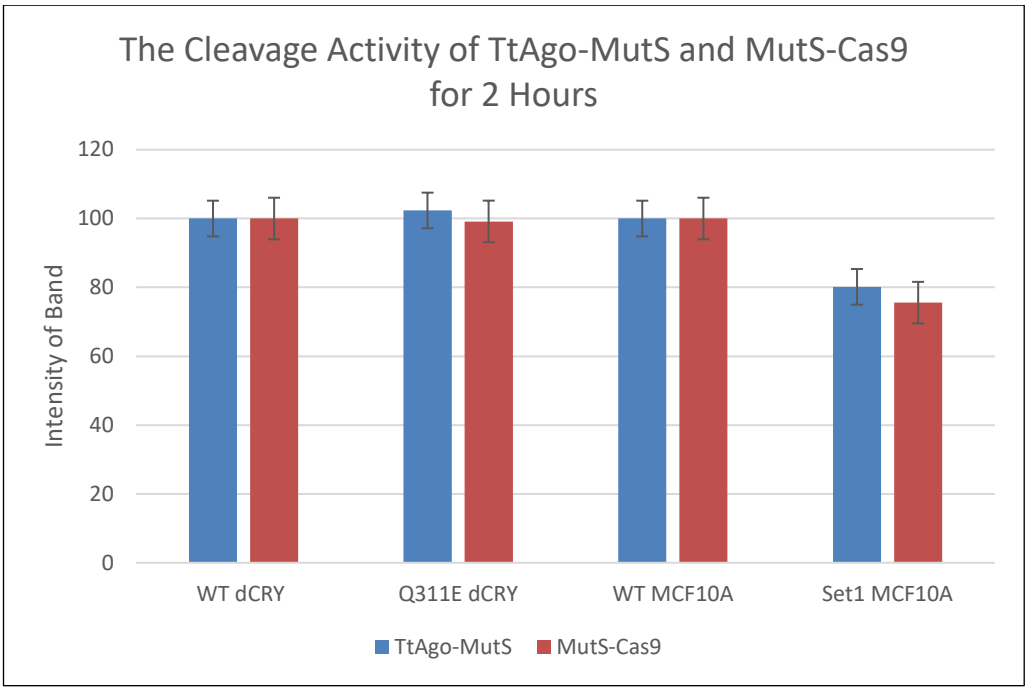

Figure 5.25: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 2 hours. Graph shows the decreased band intensity due to cleavage activity of TtAgo-MutS and MutS-Cas9 on Set1 MCF10A heteroduplex product (containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA).

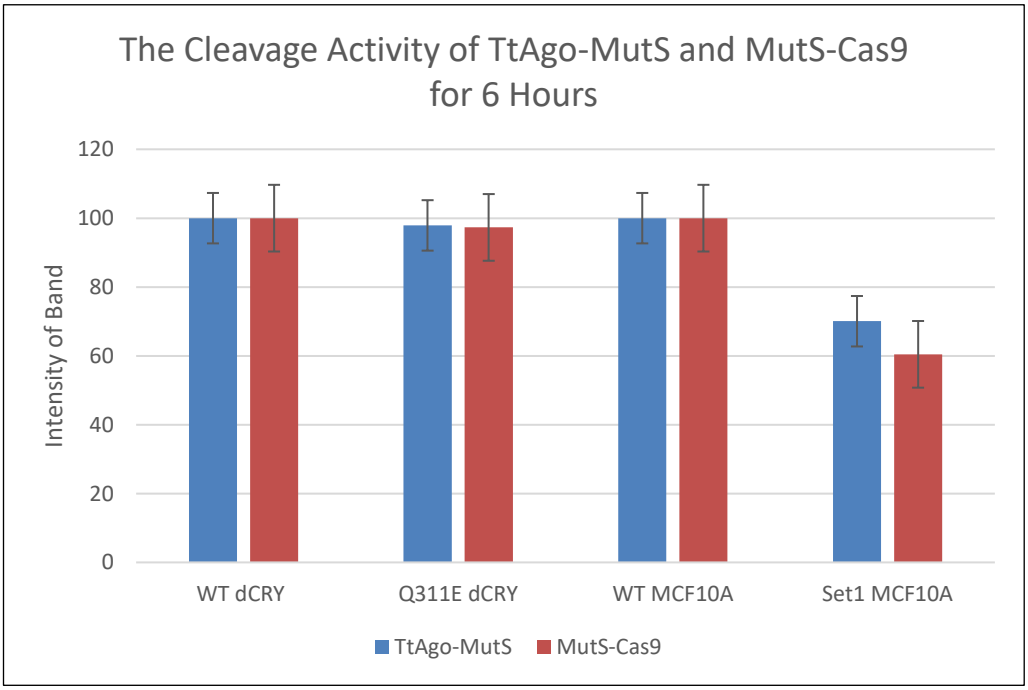

Figure 5.26: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 6 hours. Graph shows the decreased band intensity due to cleavage activity of TtAgo-MutS and MutS-Cas9 on Set1 MCF10A heteroduplex product (containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA).

These results show TtAgo-MutS and MutS-Cas9 has less efficiency for single base mutation detection. Therefore, the cleavage activity of TtAgo-MutS and MutS-Cas9 were tested on MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA samples. To optimize incubation time, products were incubated with TtAgo-MutS and MutS-Cas9 for 2, 6, and 16 hours at 37 °C (Figure 5.27).

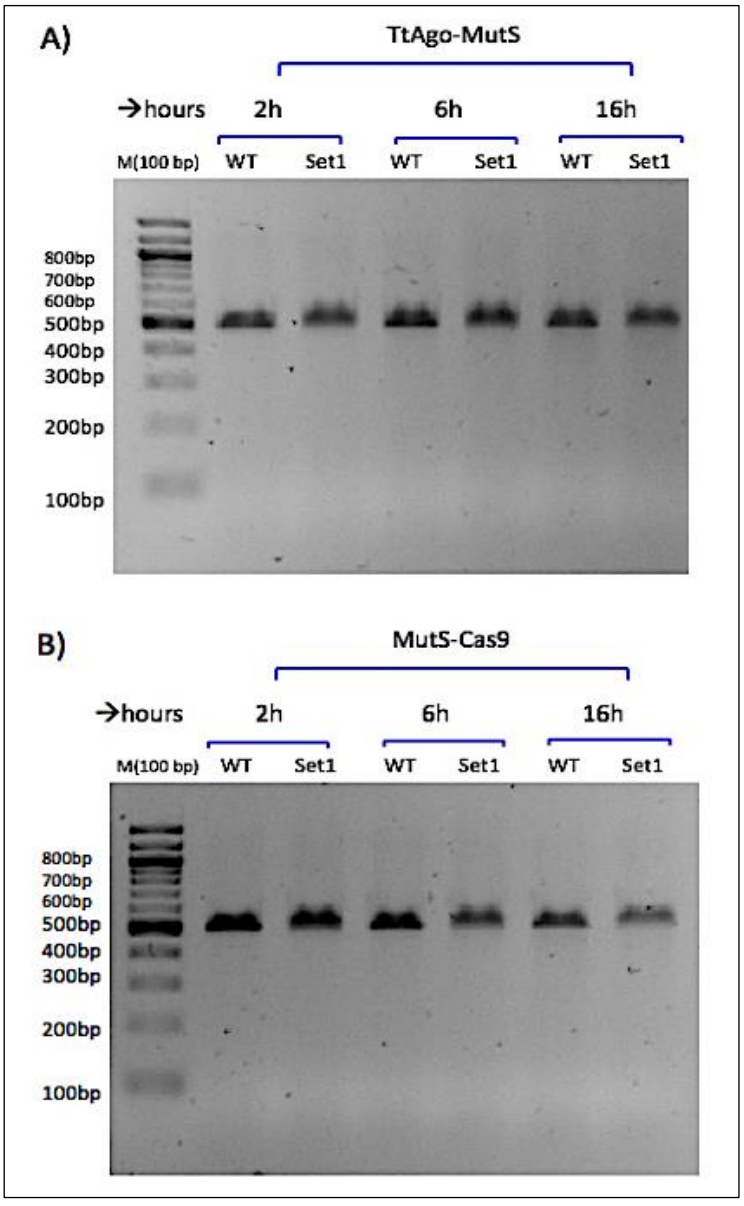

Figure 5.27: 2% agarose gel electrophoresis of the cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 2, 6, and 16 hours. The heteroduplex and homoduplex products were incubated with TtAgo-MutS (shown in A) and MutS-Cas9 (shown in B) at 37 °C for 2, 6, and 16 hours. WT is homoduplex MCF10A WT DNA, Set1 is heteroduplex that contains MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA. The heteroduplex products which are shown as Set1 were degraded by TtAgo-MutS and MutS-Cas9.

The quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 2, 6, and 16 hours was shown (Figure 5.28).

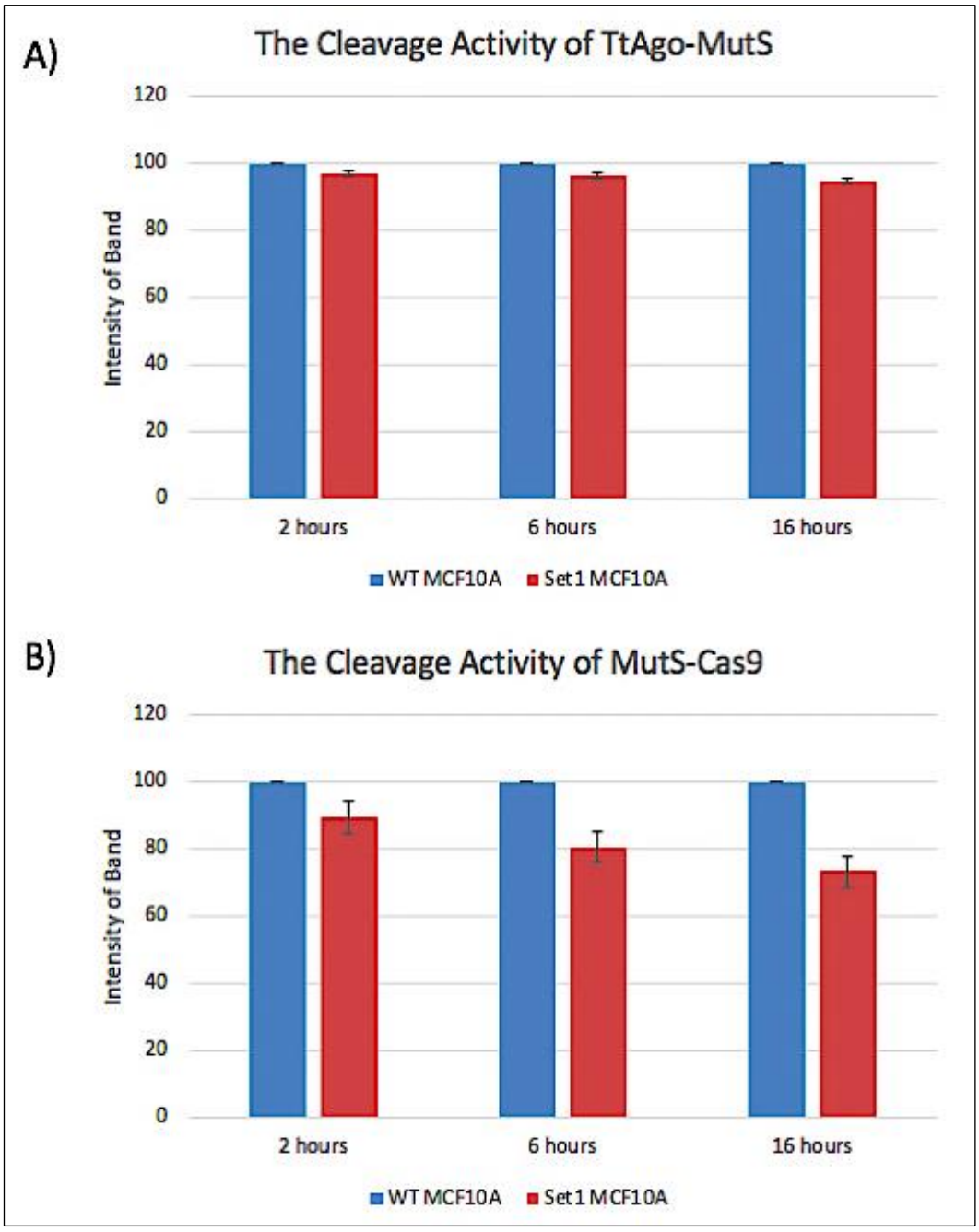

Figure 5.28: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 2, 6, and 16 hours. Graphics show the decreased band intensity due to cleavage activity of TtAgo-MutS (shown in A) and MutS-Cas9 (shown in B) on Set1 MCF10A heteroduplex product (containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA).

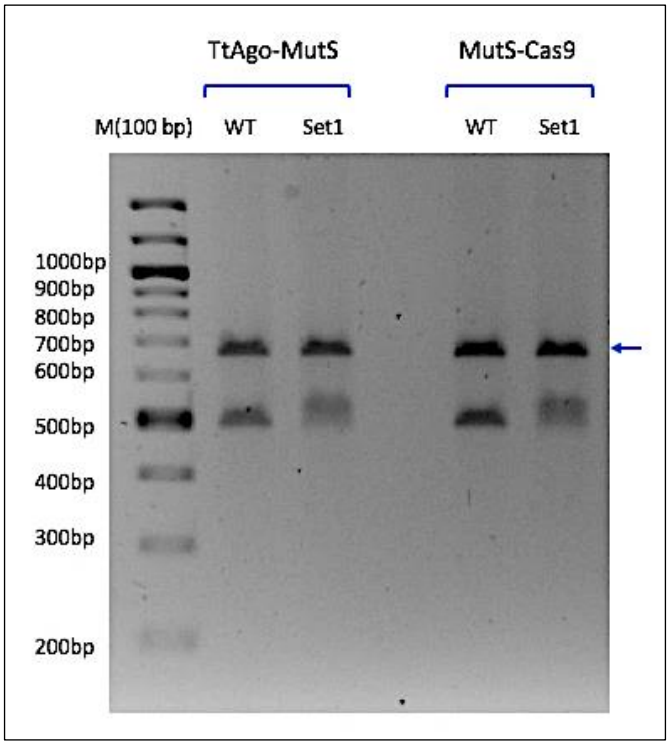

Figure 5.29: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS and MutS-Cas9 at 37 °C for 6 hours (with internal control). The heteroduplex and homoduplex products were incubated with TtAgo-MutS and MutS-Cas9 at 37 °C for 6 hours. WT is homoduplex MCF10A WT DNA, Set1 is heteroduplex that contains MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA. pAc5.1 HisV5 dCRY WT DNA was added into all products as internal control and it was shown with arrow. The heteroduplex products which are shown as Set1 were degraded by TtAgo-MutS and MutS-Cas9.

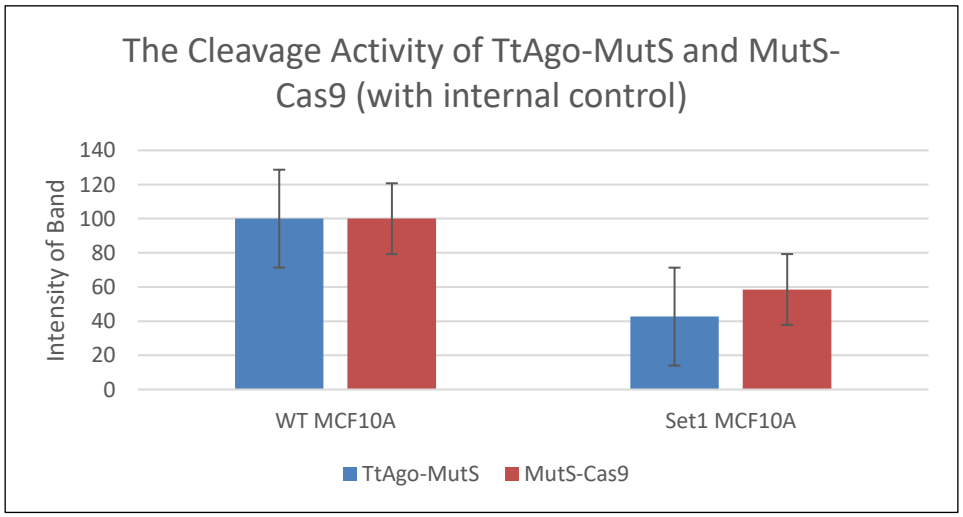

Figure 5.30: Quantification of TtAgo-MutS and MutS-Cas9 cleavage activity at 37 °C for 6 hours (with internal control). Graphic shows the decreased band intensity due to cleavage activity of TtAgo-MutS and MutS-Cas9 on Set1 MCF10A heteroduplex product (containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA).

These results show 6 hour is appropriate incubation time for cleavage activity of TtAgo-MutS and MutS-Cas9 (Figure 5.29 and Figure 5.30). To optimize incubation temperature, products were incubated with TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C.

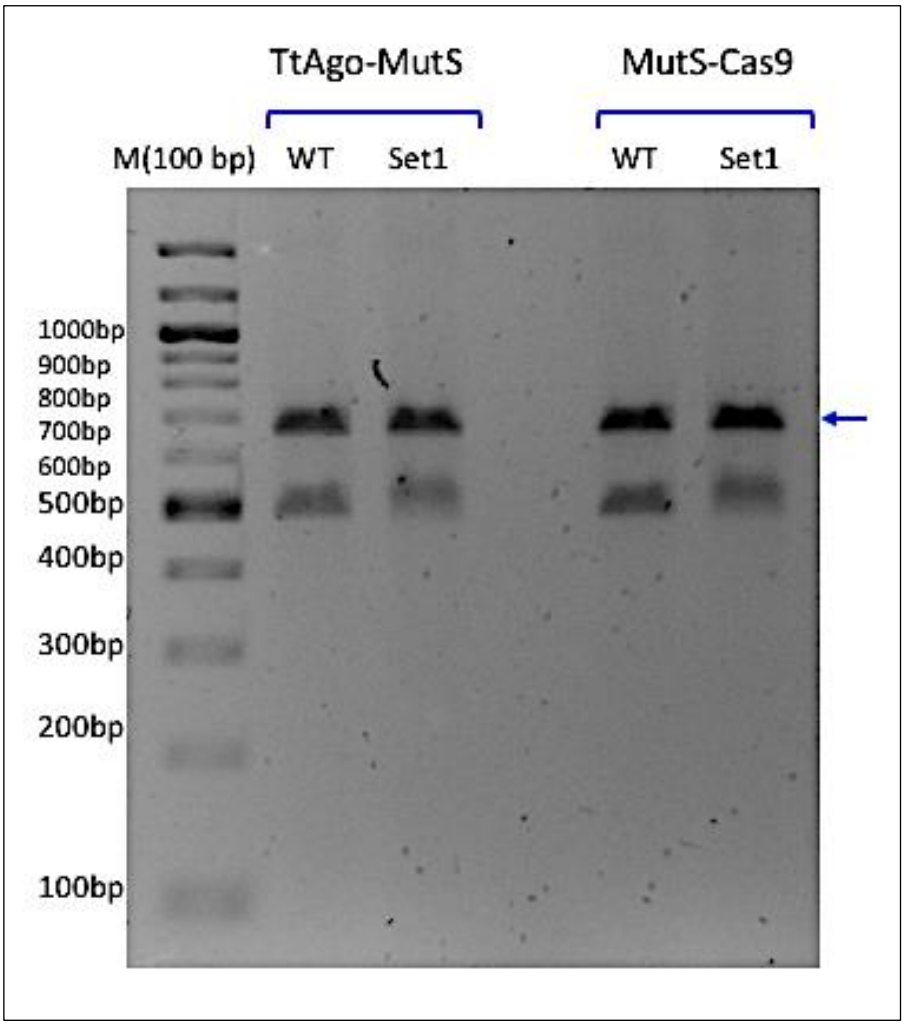

Figure 5.31: 2% agarose gel electrophoresis of cleavage activity of TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours (with internal control). The heteroduplex and homoduplex products were incubated with TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours. WT is homoduplex MCF10A WT DNA, Set1 is heteroduplex that contains MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA. pAc5.1 HisV5 dCRY WT DNA was added into all products as internal control and it was shown with arrow. The heteroduplex products which are shown as Set1 were degraded by TtAgo-MutS and MutS-Cas9.

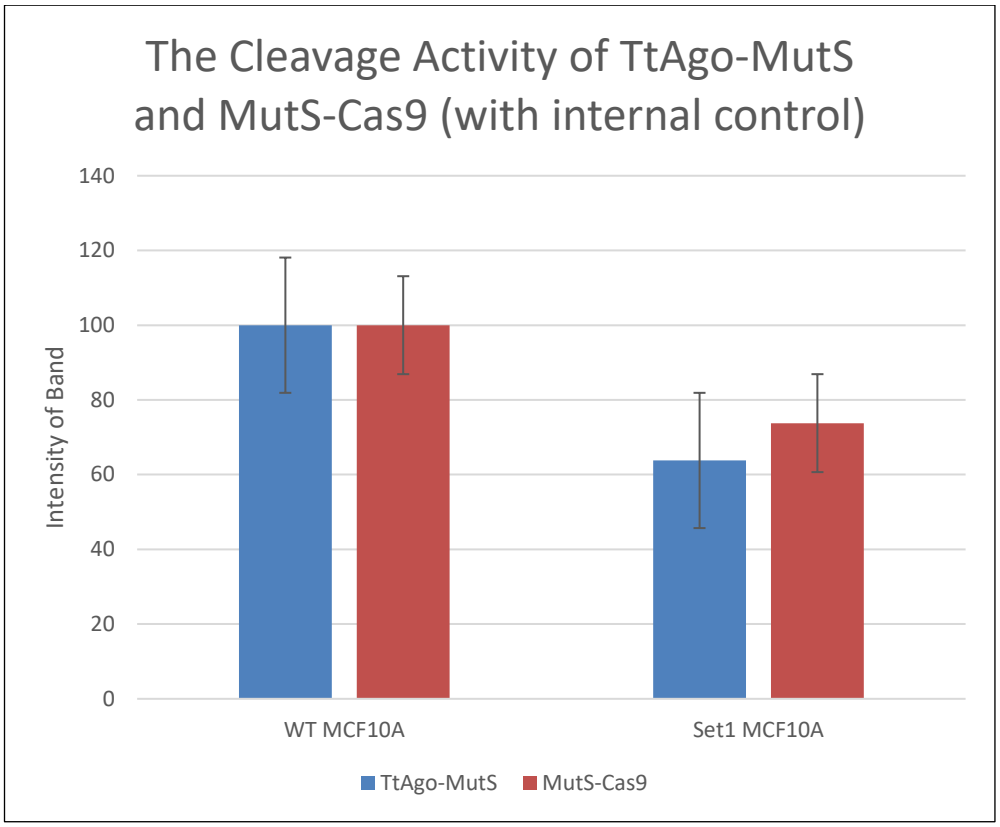

Figure 5.32: Quantification of cleavage activity of TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours (with internal control). Graphic shows the decreased band intensity due to cleavage activity of TtAgo-MutS and MutS-Cas9 on Set1 MCF10A heteroduplex product (containing MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA).

The agarose gel image shows TtAgo-MutS and MutS-Cas9 have cleavage activity both at 37 °C and 60 °C (Figure 5.31). The quantification of cleavage activity of TtAgo-MutS at 60 °C and MutS-Cas9 at 37 °C for 24 hours shows TtAgo-MutS and MutS-Cas9 have cleavage activity both at 37 °C and 60 °C (Figure 5.32).

Additionally, to observe the effect of high temperature on cleavage activity of TtAgo-MutS and MutS-Cas9, products were incubated with these enzymes at 72 °C.

According to this result, the cleavage activity of TtAgo-MutS and MutS-Cas9 at high temperature exhibits low performance (Figure 5.33).

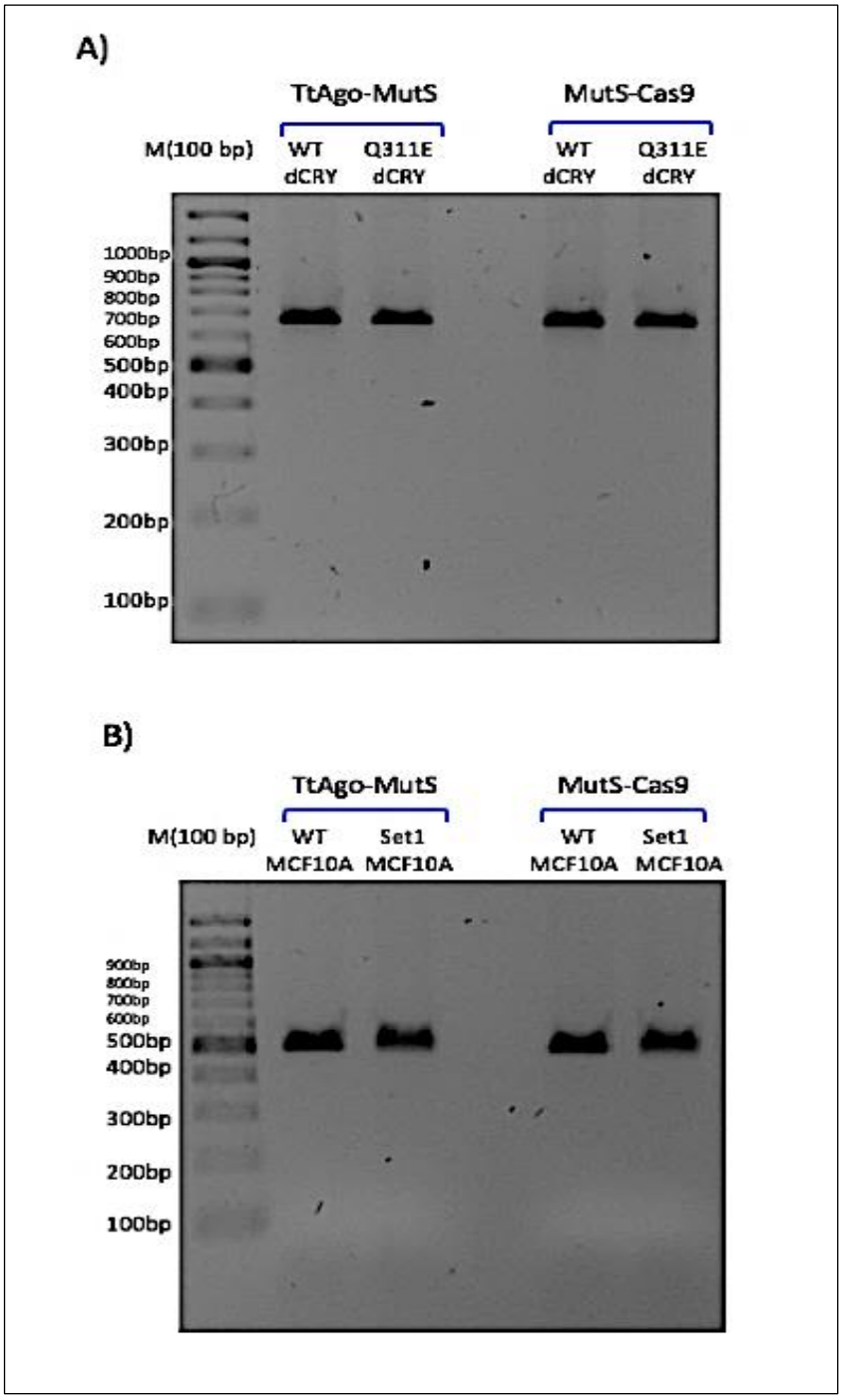

Figure 5.33: 2% agarose gel electrophoresis of the cleavage activity of TtAgo-MutS and MutS-Cas9 at 72 °C for 3 hours. The heteroduplex and homoduplex products were incubated with TtAgo-MutS and MutS-Cas9 at 72 °C for 3 hours. WT dCRY is homoduplex pAc5.1 HisV5 dCRY WT DNA, Q311E dCRY is heteroduplex that contains pAc5.1 HisV5 dCRY WT DNA and pAc5.1 HisV5 dCRY Q311E DNA (shown in A). WT MCF10A is homoduplex MCF10A WT DNA, Set1 MCF10A is heteroduplex that contains MCF10A WT DNA and MCF10A *BMAL1* k.o Set1 DNA (shown in B).

6. DISCUSSION

The development of CRISPR/Cas9 system which uses customizable sequence-guided nuclease has revolutionized genome editing research. CRISPR/Cas9 system provides to form double-strand breaks (DSBs) at the desired region in the genome for genome editing. After the formation of DSBs, non-homologous end-joining (NHEJ) repair mechanism is usually used for resolving the DSBs. NHEJ mechanism is error-prone and resulting in small insertions and/or deletions (indels) at the site of the break. The formation of indels results in the generation of frameshift mutation and it can knock out the function of the gene [14]. In CRISPR/Cas9 genome editing researches, using of sequence-specific nucleases causes an important challenge: detection and validation of the desired on-target mutation. All methods for the analysis of CRISPR/Cas9-induced mutations have advantages and disadvantages [23]. The choice of the best method in any condition depends on the expected frequency and length of the mutations, the sensitivity and the cost of the method.

In the first step of the study, we evaluated the most widely-used three methods for the CRISPR/Cas9-induced mutation detection with regards to sensitiveness, workflow, advantages, and disadvantages.

One of the most widely used methods for detection of CRISPR/Cas9-induced mutations is the T7E1 cleavage assay. T7E1 detects mismatches and accurately cleaves heteroduplexes. Digested DNA fragments are run on an agarose gel. In this way, T7E1 cleavage assay presents a predictable fragmentation pattern which consisted of cleaved and uncleaved DNA fragments [24]. In this study, we used T7E1 cleavage assay to detect both point mutation and CRISPR/Cas9-induced mutation. We optimized the incubation time of T7E1 enzyme as 2 hours to obtain a clear fragmentation pattern. To estimate the efficiency of gene editing, band brightness can be used and the more efficient gene editing causes the brighter cleaved product [32]. We evaluated the fragmentation pattern observing band brightness and calculated the indel ratio in percentage. According to the fragmentation pattern and the indel ratio in percentage of MCF10A set and pAc5.1 HisV5 dCRY set (Figure 5.10 and Figure 5.11), T7E1 has low sensitiveness for point mutation detection. The

fragmentation pattern of pAc5.1 HisV5 dCRY set was observed unclear and indel ratio in percentage was calculated as low. However, the fragmentation pattern of MCF10A set was observed clear and indel ratio in percentage was calculated as higher. Therefore, T7E1 cleavage assay is an efficient method for detection of CRISPR/Cas9-induced mutations. T7E1 cleavage assay is advantageous due to it is simple, cost effective and fast. In spite of these advantages, T7E1 can overlook the single nucleotide changes.

Secondly, the HRM method analyzes the melting curve of PCR products in the presence of an intercalating fluorescent dye. A melt curve is generated showing the temperature-dependent denaturation profile of the PCR products thanks to fluorescent properties of dye [36]. Wild type and mutant products generate a different melt curve. CRISPR/Cas9-induced mutations are detected by analysis of different HRM curve profiles. In this study, HRM analysis was used to detect both point mutation and CRISPR/Cas9-induced mutation. In HRM curve profiles of MCF10A set and pAc5.1 HisV5 dCRY set, heteroduplex products have lower Tm than wild type products. This Tm shifting provides to detect both point mutation and CRISPR/Cas9-induced mutation efficiently. HRM analysis is simple, cost-effective and faster than T7E1 cleavage assay. HRM analysis can be used to detect single nucleotide changes. In spite of these advantages, set up cost for machinery can be high.

Thirdly, MutS is efficiently used as a tool for mutation detection [43]. The electrophoretic mobility shift assay (EMSA) is used to determine MutS-DNA interaction [46]. We quantitatively evaluated the binding ability of MutS to heteroduplex mutant samples. The observation of retarded and shifted heteroduplex samples in agarose gel demonstrates the formation of MutS-mutant sample complex. This binding activity of MutS was used for detection of both point mutation and CRISPR/Cas9-induced mutation. Binding of MutS requires 15 minutes and due to time-efficient, this method is more advantageous than using of T7. However, MutS-based mutation detection method cannot show the estimated mutation site, hence indel ratio in percentage cannot be calculated. Therefore, using of T7E1 cleavage assay is more sensitive than MutS based mutation detection method.

Moreover, we tested the rapid agarose electrophoretic mobility shift assay using agarose gel instead of the polyacrylamide gel. This method requires simply altering the running buffer using of TB instead of TBE. Reducing the conductivity of the running buffer increases the resolving potential of agarose gel by providing for the use of higher voltages and shorter gel run times [41]. The successful application of rapid agarose electrophoretic mobility shift assay instead of classical EMSA presents an alternative method which is efficient, fast, inexpensive and safe.

Additionally, to control MutS binding activity, HRM analysis was performed with PCR products which contain MutS. Control group (contains dialysis buffer) of both two sets has similar Tm shift profile with normal HRM curve. According to results, the formation of MutS-heteroduplex DNA was detected by HRM analysis for MCF10A set. However, HRM analysis could not detect the MutS-heteroduplex DNA for pAc5.1 HisV5 dCRY set.

In summary, each mutation detection method has different advantages depending on different conditions, and these advantages change according to mutation type (point mutation or indel). The fact that mutation sequence cannot be observed and is a disadvantage for all of these methods. The appropriate method may be preferred considering these properties.

In the second step of the study, we tested reprogramming of Cas9 for PAM independence for detecting mismatches by replacing the DNA binding region with MutS. We benefited from the research of split Cas9 design [20]. The reengineered MutS-Cas9 enzyme was generated by replacing the PI domain with MutS. We used MutS-Cas9 enzyme as an alternative to available mismatch cleavage enzyme to detect point mutation and indel. According to results, MutS-Cas9 can detect indel but can overlook the single nucleotide changes. Therefore, we optimized the cleavage activity of MutS-Cas9 using MCF10A set. The using of MutS-Cas9 for mutation detection requires lots amount of enzyme and long incubation time. The using of MutS-Cas9 can not present fragmentation pattern like T7E1. We analyzed the cleavage activity of MutS-Cas9 by quantifying band intensity. Therefore, to detect the mutation, the efficiency of MutS-Cas9 is lower than T7E1. Moreover, we used TtAgo-MutS enzyme to compare with MutS-Cas9 with regards to mismatch cleavage activity. TtAgo-MutS enzyme detects mutations thanks to mismatch

recognizing the activity of MutS and DNA chopping activity of TtAgo. In addition, the cleavage activity of reengineered MutS-Cas9 enzyme shows that Cas9 can be reprogrammable for PAM independence.

To sum up, briefly, we generated a new reengineered MutS-Cas9 enzyme which has mismatch cleavage activity on mutant samples to detect the CRISPR/Cas9-induced mutations. This enzyme presents an alternative to available mutation detection methods and it gives a point of view in the current Cas9 reengineering studies. This study may be developed for research of split Cas9 enzyme design.

REFERENCES

[1] Scherer S., Davis R. W., (1979), "Replacement of chromosome segments with altered DNA sequences constructed in vitro", Proceedings of the National Academy of Sciences, 76(10), 4951-4955.

[2] Rudin N., Sugarman E., Haber, J. E., (1989), "Genetic and physical analysis of double-strand break repair and recombination in Saccharomyces cerevisiae", Genetics, 122(3), 519-534.

[3] Yang J., Zimmerly S., Perlman P. S., Lambowitz A. M., (1996), "Efficient integration of an intron RNA into double-stranded DNA by reverse splicing", Nature, 381(6580), 332.

[4] Chevalier B. S., Kortemme T., Chadsey M. S., Baker D., Monnat Jr R. J., Stoddard, B. L., (2002), "Design, activity, and structure of a highly specific artificial endonuclease", Molecular Cell, 10(4), 895-905.

[5] Kim Y. G., Cha J., Chandrasegaran S., (1996), "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain", Proceedings of the National Academy of Sciences, 93(3), 1156-1160.

[6] Christian M., Cermak T., Doyle E. L., Schmidt C., Zhang F., Hummel, A., Bogdanova A. J., Voytas, D. F., (2010), "Targeting DNA double-strand breaks with TAL effector nucleases", Genetics, 186(2), 757-761.

[7] Ishino Y., Shinagawa H., Makino K., Amemura M., Nakata A., (1987), "Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product", Journal of Bacteriology, 169(12), 5429-5433.

[8] 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.

[9] 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.

[10] Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J. A., & Charpentier E., (2012), "A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity", Science, 337(6096), 816-821.

[11] Makarova K. S., Haft D. H., Barrangou R., Brouns S. 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.

[12] Doudna J. A., Charpentier E., (2014), "The new frontier of genome engineering with CRISPR-Cas9", Science, 346(6213), 1258096.

[13] Cong L., Ran F. A., Cox D., Lin S., Barretto R., Habib N., Hsu P. D., Wu X., Jiang W., Marrafini L. A., Zhang, F., (2013), "Multiplex genome engineering using CRISPR/Cas systems", Science, 339(6121), 819-823.

[14] Gratz S. J., Cummings A. M., Nguyen J. N., Hamm D. C., Donohue L. K., Harrison M. M., Wildonger J., O'Connor-Giles K. M., (2013), "Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease", Genetics, 194(4), 1029-1035.

[15] Zhang F., Wen Y., Guo X., (2014), "CRISPR/Cas9 for genome editing: progress, implications and challenges", Human Molecular Genetics, 23(R1), R40-R46.

[16] Qi L. S., Larson M. H., Gilbert L. A., Doudna J. A., Weissman J. S., Arkin A. P., Lim W. A., (2013), "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression", Cell, 152(5), 1173-1183.

[17] Ibraheem D., Elaissari A., Fessi H., (2014), "Gene therapy and DNA delivery systems", International Journal of Pharmaceutics, 459(1-2), 70-83.

[18] Jinek M., Jiang F., Taylor D. W., Sternberg S. H., Kaya E., Ma E., Anders C., Hauer M., Zhou K., Lin S., Kaplan M., Iavarone A. T., Charpentier E., Nogales E., Doudna J. A., (2014), "Structures of Cas9 endonucleases reveal RNA-mediated conformational activation", Science, 343(6176), 1247997.

[19] Nishimasu H., Ran F. A., Hsu P. D., Konermann S., Shehata S. I., Dohmae N., Ishitani R., Zhang F., Nureki, O., (2014), "Crystal structure of Cas9 in complex with guide RNA and target DNA", Cell, 156(5), 935-949.

[20] Wright A. V., Sternberg S. H., Taylor D. W., Staahl B. T., Bardales J. A., Kornfeld J. E., Doudna J. A., (2015), "Rational design of a split-Cas9 enzyme complex", Proceedings of the National Academy of Sciences, 112(10), 2984-2989.

[21] Richter F., Fonfara I., Bouazza B., Schumacher C. H., Bratovič M., Charpentier E., Möglich A., (2013), "Engineering of temperature-and light-switchable Cas9 variants", Nucleic Acids Research, 44(20), 10003.

[22] Ma D., Peng S., Xie, Z., (2016)., "Integration and exchange of split dCas9 domains for transcriptional controls in mammalian cells", Nature Communications, 7, 13056.

[23] Zischewski J., Fischer R., Bortesi L., (2017), "Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases", Biotechnology Advances, 35(1), 95-104.

[24] Qiu P., Shandilya H., D'Alessio J. M., O'Connor K., Durocher J., Gerard G. F., (2004), "Mutation detection using Surveyor™ nuclease", Biotechniques, 36(4), 702-707.

[25] Ran F. A., Hsu P. D., Lin C. Y., Gootenberg J. S., Konermann S., Trevino A. E., Scott D. A., Inoue A., Matoba S., Zhang Y., Zhang F., (2013), "Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity", Cell, 154(6), 1380-1389.

[26] Kim J. M., Kim D., Kim S., Kim J. S., (2014), "Genotyping with CRISPR-Cas-derived RNA-guided endonucleases", Nature Communications, 5, 3157.

[27] Yu C., Zhang Y., Yao S., Wei Y., (2014), "A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish", PLoS One, 9(6).

[28] Liu W., Xie X., Ma X., Li J., Chen J., Liu Y. G., (2015), "DSDecode: a web-based tool for decoding of sequencing chromatograms for genotyping of targeted mutations", Molecular Plant, 8(9), 1431-1433.

[29] Hendel A., Fine E. J., Bao G., Porteus M. H., (2015), "Quantifying on-and off-target genome editing", Trends in Biotechnology, 33(2), 132-140.

[30] Bauer D. E., Canver M. C., Orkin S. H., (2015), "Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9", Journal of Visualized Experiments: JoVE, (95).

[31] Ramlee M. K., Yan T., Cheung A. M., Chuah C. T., Li S., (2015), "High-throughput genotyping of CRISPR/Cas9-mediated mutants using fluorescent PCR-capillary gel electrophoresis", Scientific Reports, 5, 15587.

[32] Ran F. A., Hsu P. D., Wright J., Agarwala V., Scott D. A., Zhang F., (2013), "Genome engineering using the CRISPR-Cas9 system", Nature Protocols, 8(11), 2281.

[33] Mashal R. D., Koontz J., Sklar J., (1995), "Detection of mutations by cleavage of DNA heteroduplexes with bacteriophage resolvases", Nature Genetics, 9(2), 177.

[34] Freeman A. D., Déclais A. C., Lilley D. M., (2013), "The importance of the N-terminus of T7 endonuclease I in the interaction with DNA junctions", Journal of Molecular Biology, 425(2), 395-410.

[35] Vouillot L., Thélie A., Pollet N., (2015), "Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases", G3: Genes, Genomes, Genetics, 5(3), 407-415.

[36] Li Y. D., Chu Z. Z., Liu X. G., Jing H. C., Liu Y. G., Hao D. Y., (2010), "A cost-effective high-resolution melting approach using the EvaGreen dye for DNA polymorphism detection and genotyping in plants", Journal of Integrative Plant Biology, 52(12), 1036-1042.

[37] Thomas H. R., Percival S. M., Yoder B. K., Parant J. M., (2014), "High-throughput genome editing and phenotyping facilitated by high resolution melting curve analysis", PloS One, 9(12), e114632.

[38] Dahlem T. J., Hoshijima K., Jurynec M. J., Gunther D., Starker C. G., Locke A. S., Wis A. M., Voytas D. F., Grunwald D. J., (2012), "Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome", PLoS Genetics, 8(8), e1002861.

[39] Zhu X., Xu Y., Yu S., Lu L., Ding M., Cheng J., Song G., Gao X., Yao L., Fan D., Meng S., Zhang X., Hu S., Tian Y., (2014), "An efficient genotyping method for genome-modified animals and human cells generated with CRISPR/Cas9 system", Scientific Reports, 4, 6420.

[40] Shui B., Matias H.L., Guo Y., Peng Y., (2016), "The rise of CRISPR/Cas for genome editing in stem cells", Stem Cells International, 2016.

[41] Ream J. A., Lewis L. K., Lewis K. A., (2016), "Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions", Analytical Biochemistry, 511, 36-41.

[42] Modrich P., Lahue R., (1996), "Mismatch repair in replication fidelity, genetic recombination, and cancer biology", Annual Review of Biochemistry, 65(1), 101-133.

[43] Stanisławska-Sachadyn A., Sachadyn P., (2005), "MutS as a tool for mutation detection", Acta Biochimica Polonica, 52(3), 575-83.

[44] Su S. S., Lahue R. S., Au K. G., Modrich P., (1988), "Mispair specificity of methyl-directed DNA mismatch correction in vitro", Journal of Biological Chemistry, 263(14), 6829-6835.

[45] Nakahara T., Zhang Q. M., Hashiguchi K., Yonei S., (2000), "Identification of proteins of Escherichia coli and Saccharomyces cerevisiae that specifically bind to C/C mismatches in DNA", Nucleic Acids Research, 28(13), 2551-2556.

[46] Lishanski A., Ostrander E. A., Rine J., (1994), "Mutation detection by mismatch binding protein, MutS, in amplified DNA: application to the cystic fibrosis gene", Proceedings of the National Academy of Sciences, 91(7), 2674-2678.

[47] Sachadyn P., Stanisawska A., Kur J., (2000), "One tube mutation detection using sensitive fluorescent dyeing of MutS protected DNA", Nucleic Acids Research, 28(8), e36.

[48] Hunt E. A., Evans Jr T. C., Tanner N. A., (2018), "Single-stranded binding proteins and helicase enhance the activity of prokaryotic argonautes in vitro", PloS One, 13(8), e0203073.

[49] O'Geen H., Ren C., Coggins N. B., Bates S. L., Segal D. J., (2018), "Unexpected binding behaviors of bacterial Argonautes in human cells cast doubts on their use as targetable gene regulators", PloS One, 13(3), e0193818.

[50] Swarts D. C., Szczepaniak M., Sheng G., Chandradoss S. D., Zhu Y., Timmers E. M., Zhang Y., Zhao H., Lou J., Wang Y., Joo C., Van Der Oost J., (2017), "Autonomous generation and loading of DNA guides by bacterial Argonaute", Molecular Cell, 65(6), 985-998.

[51] Willkomm S., Makarova K. S., Grohmann D., (2018), "DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids", FEMS Microbiology Reviews, 42(3), 376-387.

[52] Ozturk N., Selby C. P., Annayev Y., Zhong D., Sancar A., (2011), "Reaction mechanism of Drosophila cryptochrome", Proceedings of the National Academy of Sciences, 108(2), 516-521.

BIOGRAPHY

Asena CANTÜRK was born in Adana, September 14, 1992. She received her B. Sc. degree in Molecular Biology and Genetics from the Gebze Technical University in 2016. Asena joined the MSc. Program in Gebze Technical University, Graduate School of Natural and Applied Science, Molecular Biology and Genetics Department. She worked for 2.5 years under the supervision of Assoc. Prof. Dr. Nuri ÖZTÜRK.

APPENDICES

Appendix A: DNA and Protein Markers

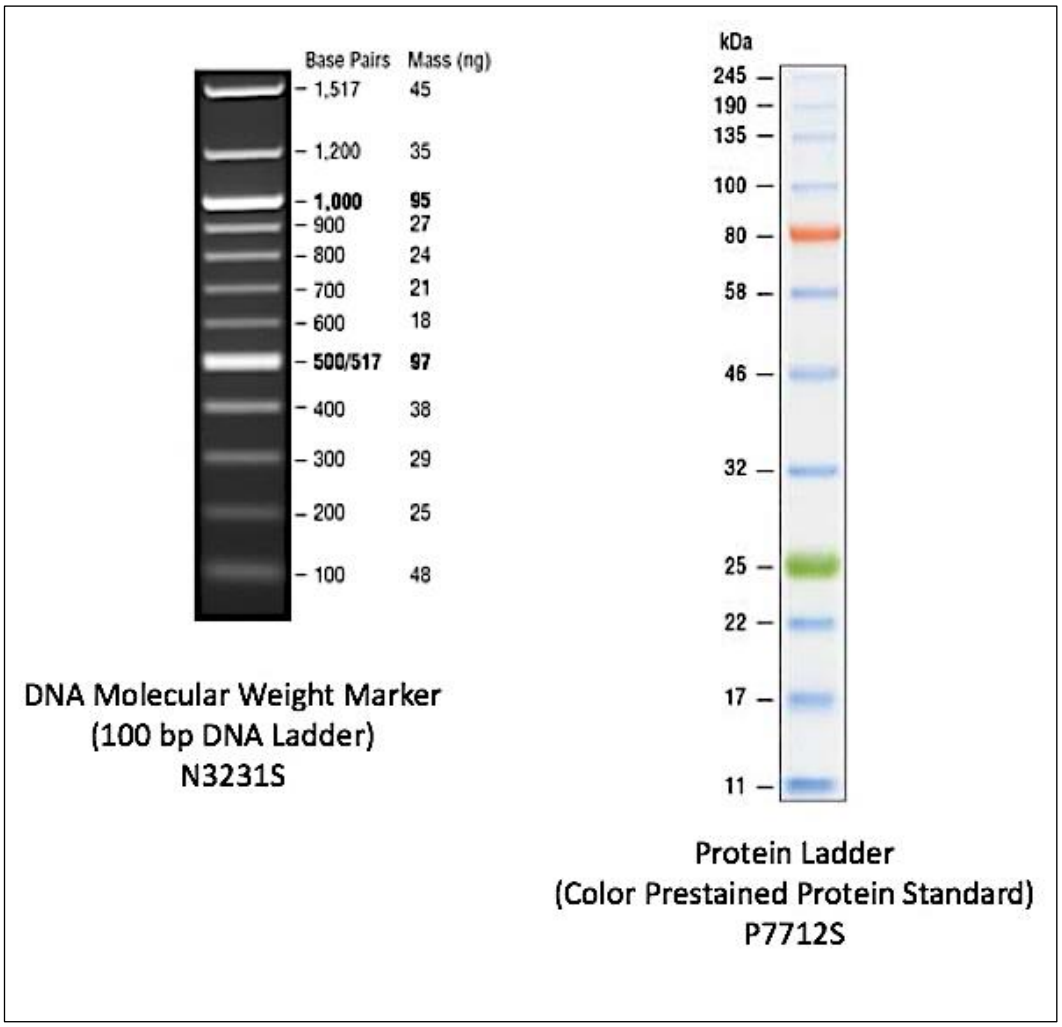

Figure A1.1: DNA Molecular Weight Marker and Protein Marker.