DEVELOPMENT OF GENE EDITING STRATEGIES FOR HUMAN β-GLOBIN (HBB)
GENE MUTATIONS

by
Batuhan Mert Kalkan

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
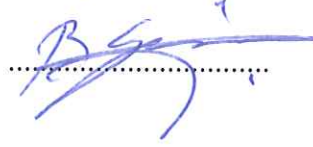Biotechnology

Yeditepe University
2018

DEVELOPMENT OF GENE EDITING STRATEGIES FOR HUMAN β-GLOBIN (HBB)
GENE MUTATIONS

APPROVED BY:

Assoc. Prof. Dr. Fatih Kocabaş
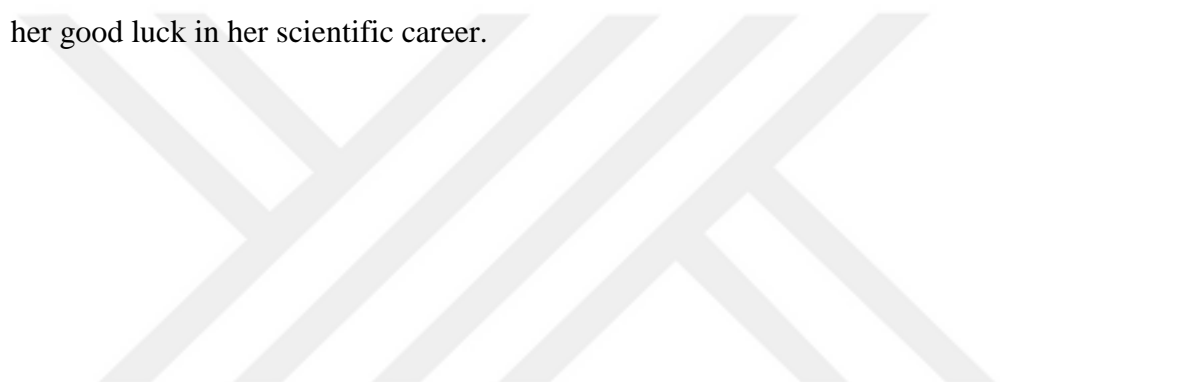(Thesis Supervisor)

Assist. Prof. Dr. Bilge Güvenç Tuna

Assist. Prof. Dr. Ceyda Açılan Ayhan

DATE OF APPROVAL: ..../..../2018

ACKNOWLEDGEMENTS

ABSTRACT

DEVELOPMENT OF GENE EDITING STRATEGIES FOR HUMAN β-GLOBIN (HBB) GENE MUTATIONS

Recent developments in gene editing technology have enabled scientists to modify DNA sequence by using engineered endonucleases. These gene editing tools are promising candidates for clinical applications, especially for treatment of inherited disorders like sickle cell disease (SCD). SCD is caused by a point mutation in human β-globin gene (HBB). Clinical strategies have demonstrated substantial success, however there is not any permanent cure for SCD available in clinics. CRISPR/Cas9 platform uses a single endonuclease and a single guide RNA (sgRNA) to induce sequence-specific DNA double strand break (DSB). In this study, it was aimed to target HBB gene via CRISPR/Cas9 genome editing tool to introduce nucleotide alterations for efficient genome editing and correction of point mutations causing SCD in HEK293T cell line, by Homology Directed Repair (HDR). We have achieved to induce target specific nucleotide changes on HBB gene in the locus of mutation causing SCD. The effect on on-target activity of longer sgRNA was examined and observed that longer sgRNA has the same performance for targeting and Cas9 induced DSBs. HDR mechanism was triggered by co-delivery of donor DNA repair templates in circular plasmid form. In conclusion, we have suggested methodological pipeline for efficient targeting and creating desired modifications on HBB gene.

ÖZET

İNSAN β-GLOBIN (HBB) GEN MUTASYONLARI İÇİN GEN DÜZENLEME STRATEJİLERİNİN GELİŞTİRİLMESİ

Gen düzenleme teknolojisindeki son ilerlemeler, programlanabilir endonükleazlar kullanılarak DNA dizisinde istenilen değişikliklerin yapılabilmesini olası kılmıştır. Bu yeni gen düzenleme araçları, β-globin genindeki (HBB) nokta mutasyonların neden olduğu kalıtsal bir monogenik bozukluk olan orak hücre hastalığı (SCD) gibi genetik bozuklukların tedavisinde, klinik uygulamalar için umut verici adaylardır. SCD için destekleyici tedavi yöntemleri geliştirilmiş olsa da hala uzun süreli evrensel tedavi eksikliği söz konusudur. Tip II CRISPR-Cas9 sistemleri, tek bir endonükleaz, Cas9, ile diziye özgü DNA çift iplikli kopmaya (DSB) katılan tek kılavuz RNA (sgRNA) olarak adlandırılan kısa bir RNA molekülü ile hareket eder. Bu çalışmada, HEK293T hücre dizisinde, Homoloji ile Yönlendirilmiş Onarım (HDR) ile etkili genom düzenleme ve SCD'ye neden olan nokta mutasyonlarının düzeltilmesi için nükleotid değişikliklerinin tanıtılması amacıyla HBB geninin CRISPR / Cas9 genom düzenleme aracı ile hedeflenmesi amaçlanmıştır. Çalışma sonucunda SCD'ye neden olan mutasyonun lokusunda HBB geninde hedef spesifik nükleotid değişikliklerini indüklenebildi. Uzun sgRNA'nın hedef aktivitesi üzerindeki etkisi incelendi ve uzun sgRNA'nın hedefleme ve Cas9 ile indüklenen DSB'ler için aynı performansa sahip olduğu gözlendi. HDR mekanizması, donör DNA onarım şablonlarının plazmid formunda birlikte verilmesiyle tetiklendi. Sonuç olarak, etkili hedefleme ve HBB geninde istenen modifikasyonlar için metodolojik bir yol haritası önerildi.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

AAVs	Adeno-associated viral vectors
AdVs	Adenovirus vectors
BSA	Bovine serum albumin
Cas	CRISPR-associated protein
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DSBs	Double strand breaks
ESC	Embryonic stem cell
FBS	Fetal bovine serum
GVHD	Graft versus host disease
HBB	human β-globin gene
HDR	Homology directed repair
HR	Homologous recombination
HSPCs	Hematopoietic progenitor stem cells
IDLVs	Integrase deficient lentivirus vectors
iPSCs	Induced pluripotent stem cells
M	Molar
ml	Millilitre
mM	Millimolar
NHEJ	Non-homologous end joining
PAM	Protospacer adjacent motif
PKc	Protein kinase
SCD	Sickle cell disease
sgRNA	Synthetic guide RNA
TALENs	Transcription activator-like effector nucleases
ZFNs	Zinc finger nucleases
µl	Microliters

1.   INTRODUCTION

1.1. THE HISTORY OF GENOME EDITING

Genome editing has recently developed in the most rapid way and became a routine protocol in hundreds of research laboratories globally. Despite having a short time course of development, genome editing technology has such an intense history that can globally gather researchers around for robust and efficient alterations of mammalian and plant genomes. Studies of classical genetics have been relied on analysis and discovery of spontaneous mutations. The very first steps of DNA sequence changes spans to the findings showing that chemicals and radiation could enhance the rate of mutagenesis [1, 2]. However, both chemical and radiation mutagenesis cause base changes at random genomic sites. Targeted genomic alterations were first created taking the advantage of homologous recombination (HR) as a precise but inefficient way compare to the previous methods [3, 4]. Nevertheless; recovery, characterization and selection of targeted cells were challenging and inefficient [5]. Moreover, targeted genome modifications were not applicable for the species other than mouse due to lack of culturable ESCs of mammalians.

Rapid improvement of genome editing technology has solved the issues mentioned above and provided site directed genomic modifications in all types of cells, tissues and organisms [6, 7]. In the past decade, genome engineering has steered towards the use of engineered nucleases composed of DNA binding and cleavage domains, providing sequence specific targeting and precise modifications on genome [8]. Working principle of these nucleases relies on the induction of double strand breaks (DSBs) on DNA, stimulating DNA repair mechanisms in cells [9]. Essentially, DNA cleavage domain of nucleases create a DSB and facilitate the generation of insertions, deletions and substitutions desired at the genomic site of interest. Various platforms of engineered nucleases have been in use for genome editing studies.

1.2. GENE EDITING TOOLS AND STRATEGIES FOR BLOOD DISORDERS

The novel approach called genome editing has been widely used in the research field of gene therapy, functional genomics and development of transgenic organisms for the past years. Gene editing is simply based on the usage of engineered, programmable and target specific nucleases inducing point specific modifications in the genome. These programmable nucleases are composed of a motive or sequence specific DNA binding domain and a DNA cleavage domain. DNA cleavage domain creates a double strand break (DSB) and facilitates the generation of insertions, deletions and substitutions desired at the genomic site of interest. Various platforms of engineered nucleases have been in use for genome editing studies. One of the most widely used and pioneer gene editing system is Zinc Finger Nucleases (ZFNs) [10-13]. ZFNs consist of a global Cys_2-His_2 DNA binding domain and a DNA cleavage domain named FokI endonuclease [14]. Another popular tool of genome editing technology is transcription activator-like effector nucleases (TALENs), a protein originated from pathogenic bacteria named Xanthomonas [15-17]. TALENs provide specific nucleotide recognition by their DNA binding domain composed of amino acid motives. Each of these conserved motives, which are robustly programmable in a target specific manner, recognize a particular nucleotide [18]. The most recently, clustered regularly interspaced short palindromic repeats (CRISPR) / CRISPR-associated protein (Cas) 9 was introduced to dethrone TALENS and ZFNs for gene editing [19, 20]. Unlike a peptide – DNA interaction to provide targeting specificity as in ZFN and TALEN approach, CRISPR/Cas9 system has a basis of guide RNA – DNA complementation to ensure a higher performance of sequence specific targeting of any genomic location. A 20 base long guide RNA sequence co-delivered with Cas9 protein is the only requirement for specific targeting and DNA cleavage [21]. CRISPR technology promises a faster, easier and cheaper design compare to ZFN, TALEN and related genome editing techniques. Comparing the efficiencies of CRISPR/Cas9 and TALEN systems applied on the same cell line, it was observed that CRISPR/Cas9 is more robust, and promising method for effective genome editing [22].

CRISPR/Cas9 system was first discovered as an acquired immunity machinery in bacteria. Invader DNA is recognized by CRISPR RNA (crRNA) and cleaved by Cas9 nuclease [23]. In bacterial and archaeal genomes CRISPR locus is made of strictly conserved repetitive DNA sequences interspaced with specific sequences called spacers. Spacer sequences are

generated through cleavage of invader's DNA into small fragments and integration into CRISPR locus of the host genome. These spacer sequences are then used as DNA templates to produce crRNA targeting viral or phage DNA, acting as bacterial immunity library members. There are different CRISPR/Cas9 systems which have been still identified and also engineered, based on amino acid sequences and tertiary structures of Cas9 protein. Major classes of CRISPR/Cas9 system are I, II and III. It was described that class II CRISPR/Cas9 system requires only a Cas9 protein with two nuclease domains named RuvC and HNH, incorporated with a guide RNA. Thus, class II CRISPR/Cas9 has been pointed as a relatively simpler, efficient and easily designable system for gene editing studies [20].

CRISPR/Cas9 gene editing system is based on generation of a DSB followed by the process of cellular DNA repair. The original CRISPR/Cas9 system is guided to target site by the combination of mature crRNA and trans activating crRNA (tracrRNA) which is partially complementary to crRNA and provides to maturation of crRNA [24]. In research applications, a chimeric RNA, containing both crRNA and tracrRNA sequences, called as guide RNA (gRNA) is used [21]. Guide RNAs varying between 20-24 nucleotides are able to be designed using variety of tools, providing an easy application ability. Target specific cleavage of DNA also requires another component called protospacer adjacent motif (PAM) which is a 2-6 base length DNA sequence located in the downstream of target site [25]. PAM sequence is essential for successful binding and cleavage of targeted genomic loci[21, 26, 27]. The most commonly used PAM sequence is 5'-NGG-3' associated with Cas9 nuclease of Streptococcus Pyogenes and the researchers are still studying to identify different PAM sequences to achieve improved targeting in a wider range of sites on genome [28, 29]. Together with gRNA and PAM sequences, CRISPR/Cas9 system can target up to 30 base lengths on target site, which is theoretically a unique sequence on whole genome of different organisms. However, it has been reported that CRISPR/Cas9 system has a tolerance to mismatches observed between the guide and target sequences, which would lead to off-target mutagenesis [19, 30, 31].

The fundamental logic behind genome modification using engineered nucleases is the generation DSB near target site, triggering a subsequent DNA repair process [14, 17, 23]. There are two main endogenous repair mechanisms for DSB, which are non-homologous end joining (NHEJ) and homology directed repair (HDR). In NHEJ, broken ends of DNA

are directly ligated back together in brief. In most of the cases, NHEJ repair mechanism results in small insertions or deletions (in-dels) at the site of DNA break. These in-dels would result in small sized mutations causing gene silencing. The second repair mechanism to get rid of DSB is HDR. In this mechanism, a homology containing sequence of DNA serving as a template is required to synthesize new DNA repair the break by homologous recombination. Naturally, a sister chromatid is the template for HDR in case of DSB repair. However, it is reported that HDR mechanism can also work with the presence of an externally introduced DNA bearing homology regions, called as donor template [24, 32].

Genome editing has shown to have a remarkable potential to cure genetic diseases through permanent correction of mutations [33-35] or insertion of recuperative DNA sequences as done in gene therapy [36, 37] . Gene editing technology enables targeted genome modifications with higher precision and adaptability.

Gene editing of blood cells as therapeutic approach to cure blood disorders is a simple and conceptual idea, which has been intensively focused for the last several years in the field of research. Initial scientific approach for genome editing had been based on use of viral vectors such as retrovirus and lentivirus derived vectors. Integration of these vectors into host genome is operated in an uncontrolled manner, resulting in unexpected side effects. Several studies have underlined that lentiviral and retroviral vectors do not demonstrate random genomic integration but biased integrative fashion[38]. Various concerns have emerged due to lentiviral vector usage in treatments of blood disorders such as leukaemia and lymphoma caused by proto-oncogene activation upon insertion of viral genomic content. However, it has not been reported that there is a high risk of leukaemia in humans and lymphoma formation in mice, remaining a risk of long-term latency.

Discovery of genome editing strategies has provided more precise mutation repair in comparison to gene therapy. Engineered nuclease systems are strikingly promising tools for inherited disease therapeutics. Instead of gene addition, precise repair of mutant genes has appeared to be a safer approach. Tremendous effort is currently made to adopt and optimize the gene editing systems into iPSCs and HSPCs, carrying onward to globally accessible clinical applications. Inducing controlled double strand breaks on target locus and activating homology directed repair mechanisms in cells, engineered nucleases provide much more precise and effective way of disease therapy. Depending on the therapeutic strategy, gene

correction, knock-in and safe harbour integrations are applicable in nuclease mediated gene editing technology. Although NHEJ mechanism is more frequent than HDR, it is currently known that cells tend to favour HDR mechanism during S and G2 phases of cell cycle [39]. Therefore, hematopoietic cells may be simultaneously induced to shift proliferative state and gene correction by delivery of engineered nucleases and repair templates to achieve high efficiency gene editing.

As with the gene therapy, delivery of gene editing components to HSPCs is a major bottleneck. To avoid genotoxicity and off-target mutations, transient expression or controlled inhibition of nucleases is currently in demand. Delivery methods of genome editing tools can be categorized into two, as viral and non-viral delivery. Considering non-viral delivery approach, cells can be transfected with plasmid DNA containing and expressing gene editing components, in vitro transcribed mRNA to induce translation of nuclease in host cell, or direct delivery of purified nucleases with donor repair templates [40]. Donor template delivery is also versatile that it could be introduced to target cells in the form of plasmid DNA, dsDNA or ssDNA linear oligo [41, 42]. In non-viral delivery of gene editing tools, transfections are supported by driving forces which are cationic polymers, lipids, calcium phosphate and electroporation. However, the non-viral transfection efficiency varies between cell type. For instance, hematogenic cells are harder to transfect in comparison to many cell lines. Combination of chemical supplements and transfection methods would result in an increase in efficiency but also toxicity and stress related cell death. Delivery of in vitro transcribed mRNAs encoding engineered nucleases (ZFN, TALEN or Cas9+gRNA) appears to be more advantageous due to lower genotoxicity and transient nuclease activity [40]. Direct delivery of engineered nucleases as purified proteins using electroporation has been applied as another alternative approach to achieve temporary and safe genome editing [43]. However, the size of nuclease would affect the passage through cell membrane. To overcome this challenge, recombinant proteins with smaller size would be more convenient. In a recent study, genome editing of hematopoietic stem cells have been efficiently performed using CRISPR/Cas9 system in the form of ribonucleoprotein which is purified Cas9 protein and target specific gRNA complex [44].

Apart from non-viral delivery strategies of genome editing tools, viral based methods are alternatively used. Advances in genome engineering have brought in improvements in viral

transfer tools and methods. Generation of non-integrating viral vector such as adenovirus vectors (AdVs), adeno-associated viral vectors (AAVs) and integrase deficient lentivirus vectors (IDLVs) enabled successful transfer of genes required for the expression of gene editing tool in host cells in both in vitro and in vivo [40]. Each viral vector mentioned above possesses different characteristics in terms of transduction efficiency, packaging size and target cell type. AdVs are dsDNA viruses enabling a packaging capacity up to 37 kb which provides enough room for genes encoding nucleases and donor repair templates. As a proof of concept, CD34+ T cells had previously been genome modified using ZFNs packaged in AdVs to acquire HIV-1 resistance [45]. AAVs are currently the most widely used ssDNA viruses coinfected with a partner such as adenoviruses or herpes simplex viruses. AAVs are advantageous due to their low immunogenicity and low frequency of random integration into host genome. However, the packaging capacity is quite low, around 4.7 kb, which might be incompatible for large sized nucleases and donor DNA [46]. For instance, spCas9 can barely fit into a typical AAV, leaving no space for other elements required for gene editing. In comparison to Cas9 and TALENs, ZFNs are encoded by smaller sequence that can be packaged and delivered using AAVs [47]. IDLVs has been used to deliver ZFNs and donor templates [48]. Regarding to their competence in transducing non-dividing cells and integration deficiency, IDLVs are considered as advantageous tools for packaging [49]. Nevertheless, HSPCs are hard to transduce since they require higher titers of viral particles which is challenging to obtain for IDLVs. Therefore, IDVL production protocols ought to be developed to eliminate this drawback present in case of hematopoietic cell editing.

1.3. ENGINEERED NUCLEASES FOR GENOME EDITING

The urgent need for safe, efficient and site-specific gene editing has emerged nuclease-based approaches. As a broad definition, nucleases are the enzymes catalysing the hydrolysis of nucleic acids by the cleavage of phosphodiester bonds within the chain structure composed of nucleotide subunits [50]. There exist two major classes of nucleases based on their activity, exonucleases and endonucleases. Exonucleases cleave nucleic acid chains from the ends of recognition site while endonucleases show activity on the site of recognition and binding [51]. Figure 1.1 illustrates the catalytic activity difference between the two major types of nucleases.
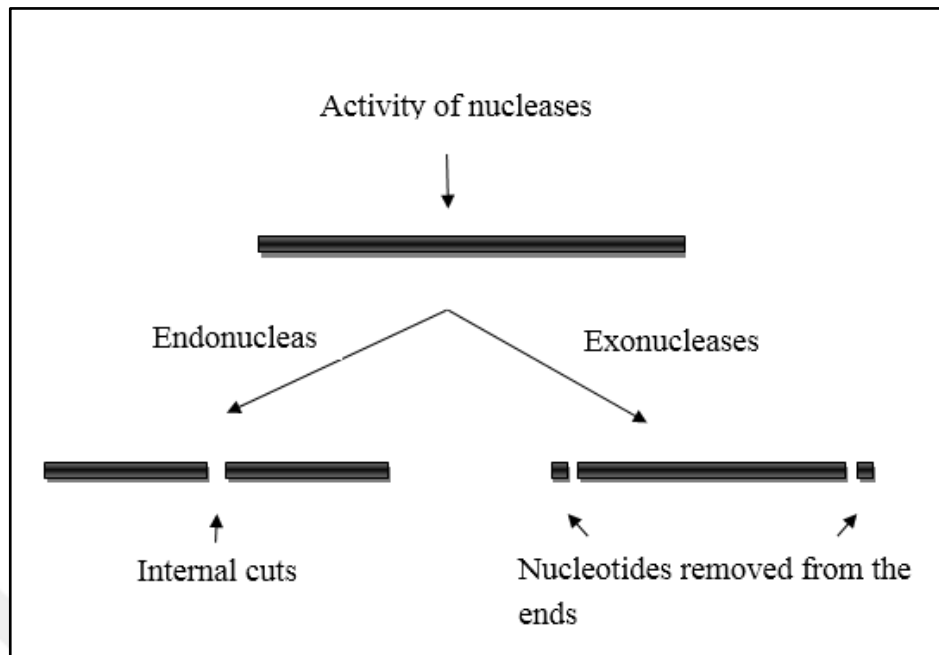
Figure 1.1. Activity of two major types of nucleases.

In genome engineering applications, endonucleases are used as site specific genome editing tools. In comparison to gene therapy and similar classical methods, nuclease mediated genome editing provides a significant increase in targeted cell percentage [52]. Since the first reports of development process in genome editing, five different classes of nucleases have been investigated so far. As the beginning steps of endonuclease mediated genome editing technology, three types of nucleases were identified; chemical endonucleases, meganucleases and Zinc Finger Nucleases (ZFNs) [53]. Later, this classification has been updated with the discovery of TALE nucleases and finally CRISPR/Cas9.

The fundamental logic behind genome modification using engineered nucleases is the generation DSB near target site, triggering a subsequent DNA repair process [14, 17, 23]. There are two main endogenous repair mechanisms for DSB, which are NHEJ and HDR. In NHEJ, broken ends of DNA are directly ligated back together in brief. In most of the cases, NHEJ repair mechanism results in small insertions or deletions (in-dels) at the site of DNA break. These in-dels would result in small sized mutations causing gene silencing. The second repair mechanism to get rid of DSB is HDR. In this mechanism, a homology containing sequence of DNA serving as a template is required to repair the DSB by homologous recombination. Naturally, a sister chromatid is the template for HDR in case of DSB repair. However, in genome engineering applications an exogenous DNA repair

template is introduced to generate desired genomic modifications [24, 32]. Figure 1.2 demonstrates the basic flow of nuclease mediated genome editing.
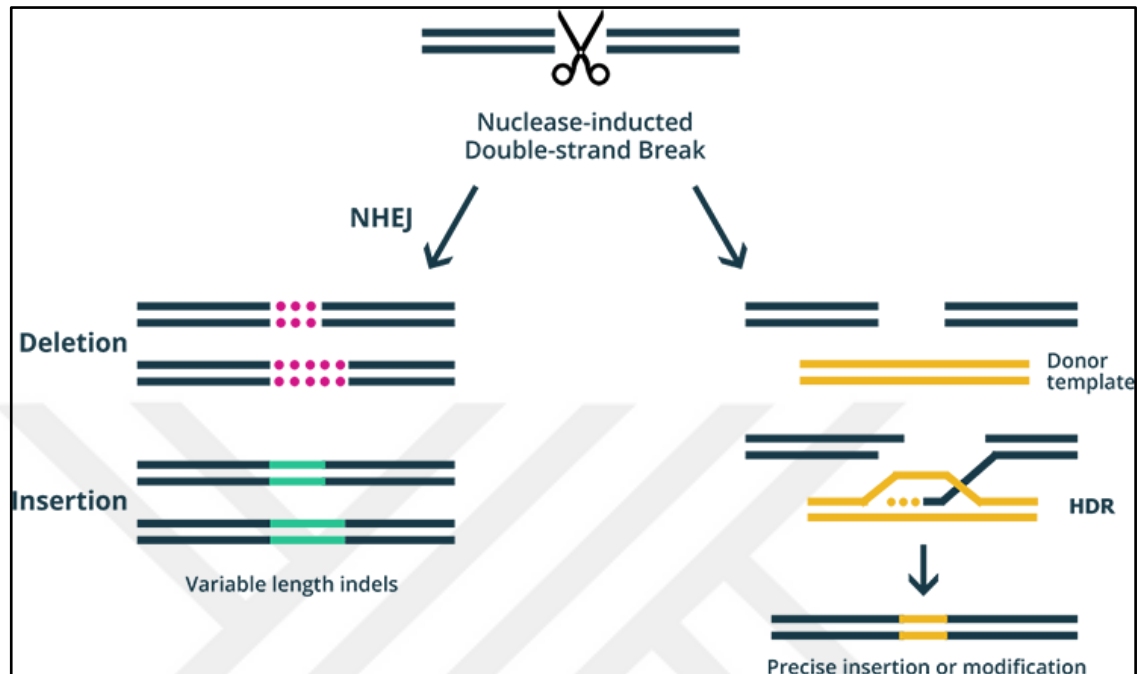
Figure 1.2. Genome editing using engineered endonucleases

Naturally, DSBs are frequently occurring events in eukaryotic cells and they are repaired with two major molecular pathways. These repair mechanisms are critical for both pathological and physiological DSBs generated during recombination events. In case of a DSB, cells urgently attempt to restore the integrity of DNA. If the cells are diploid, HDR mechanism can step in. HDR mechanism basically relies on the basis of the homology bridge established between donor and acceptor DNA [54]. In non-dividing haploid cells or in diploid organisms which are not in the S phase of cell cycle, a homology donor may not be nearby to conduct HDR. During the S phase, sister chromatids may act as homology donor templates since they are physically close. Outside of S/G2 phase, NHEJ repair mechanism become the only option to provide the opportunity to survive with a mechanistic flexibility [55]. NHEJ is mediated by a variety of enzymes peptides [56].

Once a DSB is induced, Ku is known to be the first protein to bind the cleavage site on DNA [57]. Ku forms a protein-DNA complex and serves as a harbour for nucleases, polymerases and ligases to dock and manage the NHEJ mechanism [58]. Ku can be considered as a toolbelt protein where various proteins can dock as in PCNA during DNA replication. DSB

creates two DNA ends and Ku-DNA complex is formed at each ends of DNA to be re-joined. Each Ku-DNA complex can recruit polymerase, nuclease and ligase in any combination and order [59]. Ku undergoes conformational change to form stable complexes with protein kinase (PKc), lambda and mu polymerases, and DNA ligase IV when bound to DNA ends [60, 61]. The Artemis DNA-PKc complex possesses 5' endonuclease and 3' and endonuclease activities to cleave damaged DNA overhangs [62]. The polymerases (pol mu and lambda) are able to bind Ku-DNA complex from their N terminal domain [59]. DNA ligase is described as the most modular ligase, forming the complex with XLF and XRCC4. This complex has the ability to ligate gaps and incompatible ends of DNA [63].

HDR mechanism relies on recombinational DNA repair to maintain genomic integrity. The process is homology directed in which a second chromosomal copy involves. The key concept of HDR is highly conserved in eukaryotes [64]. As previously describes, DSBs can be repaired by NHEJ instead of HDR. Nevertheless, NHEJ is highly error prone and will possibly result in indel mutations and translocations. The template dependent mechanism of HDR ensures high probability of correct restoration of DSB site. When the DSB is induced and a repair template is available. The first step of HDR is the resection by nuclease and helicase to generate single stranded DNA (ssDNA) defining the homology arms flanked. The ssDNA acts as the scaffold for the assembly of RecA/RAD51 filament and this complex promotes the search for homology [65]. Once the homology template is found, RecA/RAD51 filament catalyses the exchange of DNA strands. The DNA strand exchange is stimulated by the relationship between RecA/RAD51 and SSB/RPA which are the members of ssDNA binding proteins [66]. Finally, the Holliday structures formed during the exchange of strands are disjoined by nucleolytic resolution [67].

Genome editing has shown to have a remarkable potential to cure genetic diseases through permanent correction of mutations [33-35] or insertion of recuperative DNA sequences as done in gene therapy [36, 37]. Gene editing technology enables targeted genome modifications with higher precision and adaptability.

1.3.1. Chemical Endonucleases

Chemical endonucleases are the class of nucleases which are synthetically derived by the fusion of DNA binding polymers and DNA reactive agents. Triplex forming oligonucleotides (TFOs) are a specific example of DNA binding polymers used for gene targeting and modification [68]. Since the design and production process of DNA binding polymers is based on theoretical deduction, it has been believed to eliminate the broad screening for the identification of potential novel DNA binding polymers through expensive wet lab protocols. The fusion of TFOs and active topoisomerase domains such as psoralen, camptothecin, bypridine or variety of restriction enzyme domains had been described as potential tools for sequence specific gene editing [69-72]. Despite being promising in principle, chemical endonucleases have remained inefficient for cellular delivery, efficiency of genome modifications and failed to alleviate the safety concerns for the consideration in clinical trials.

1.3.2. Meganucleases

The meganuclease family of engineered endonucleases has been widely used for more than a decade to study and develop sequence specific DNA targeting and cleavage, inducing gene modifications. Meganucleases can be classified into 5 subfamilies based on their recognition and activity motifs on DNA. Essentially, each short sequence of DNA contributes a three-dimensional structure called DNA motif. Meganucleases are grouped according the structure motifs for recognition. The five well-known meganucleases are LAGLIDADG, GIY-YIG, HNH, His-Cys box and PD-(D/E)XK [73, 74]. LAGLIDADG meganucleases are the most commonly studied one among the other types and have been found in classes of the organisms. Although their functional roles have not been clearly understood, LAGLIDADG meganucleases have been described as highly specific nucleases which are able to recognize and cleave exon-exon junction, facilitating the splicing of their own intron [52]. LAGLIDADG meganucleases specifically recognize genomic sequences ranging between 14 and 40 base length. The protein-DNA interactions were revealed using DNA co-crystal structure [75, 76].
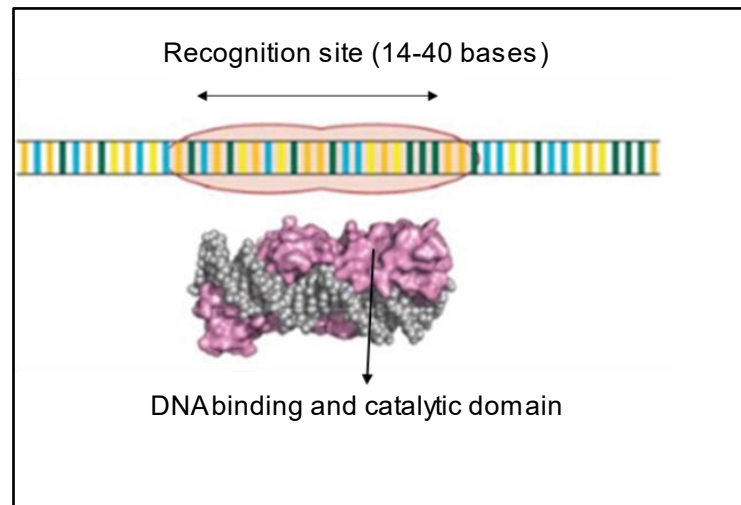
Figure 1.3. Meganucleases recognise specific DNA motifs and induce DSBs.

The establishment of gene editing strategies using meganucleases has launched with I-SceI yeast meganuclease, a member of LAGLIDADG meganucleases (Figure 1.3). It was first discovered by Jacquer and Dujon as an endonuclease with 18 base pairs recognition length and the function of intron homing in mitochondria [77]. Short after its discovery, I-SceI meganuclease was announced to be a potential tool for promoting homologous recombination mechanisms in mammalian cells [78]. This was the first study to demonstrate that meganuclease induced double strand breaks could be repaired with the presence of donor DNA templates carrying homology arms to the target locus in mouse cells. However, subsequent studies have shown that the efficiency of DSBs and homologous recombination may vary depending on the locus of interest [79]. Studies with mammalian embryos have revealed the technical and practical limitations and focused on overcoming these drawbacks using modified meganucleases conjugated with nuclear localization signal peptides [80]. Another limitation was the scarcity of I-SceI suitable genes for efficient targeting. To overcome, researchers have invested labour in producing meganucleases with custom recognition sites for different genes in human, rat and mouse genomes [81, 82]. Nowadays, the term called engineered nuclease mediated genome editing brings ZFNs, TALENs and CRISPR/Cas9 into minds. However, meganucleases were the first actors of genome editing technology in mammalian cells. In comparison to previous studies without engineered nucleases, the induction of homologous recombination and endogenous DNA repair mechanisms have elevated with nearly thousand-fold estimation [83].

1.3.3. Zinc Finger Nucleases (ZFNs)

After the first attempts of genome editing via meganucleases and unsatisfactory yield of success, researchers have started a search for novel tools. A pioneering methodological approach was released based on the engineering and fusion of a restriction enzyme FokI for induction of DSB and zinc finger proteins (ZFP) for DNA motif recognition and binding, together forming zing finger nuclease (ZFN) [84]. In ZFNs, FokI is the nuclease domain derived from bacterial restriction enzyme and this domain requires to be dimerized to cleave target DNA [85]. Therefore, a pair of ZFN monomers must be available in the system to form a complete and functional endonuclease. Each monomer of ZFPs bind adjacently but opposite sides of DNA while FokI nuclease domains sit and dimerize on 5-7 base length sequence called spacer (Figure 1.4). ZFNs cover longer DNA fragments compare to meganucleases and chemical nucleases previously used. Thus, ZFNs provide higher sequence specificity due to longer coverage, 23-25 base pair (up to 43) which is considerably unique sequence in genome.
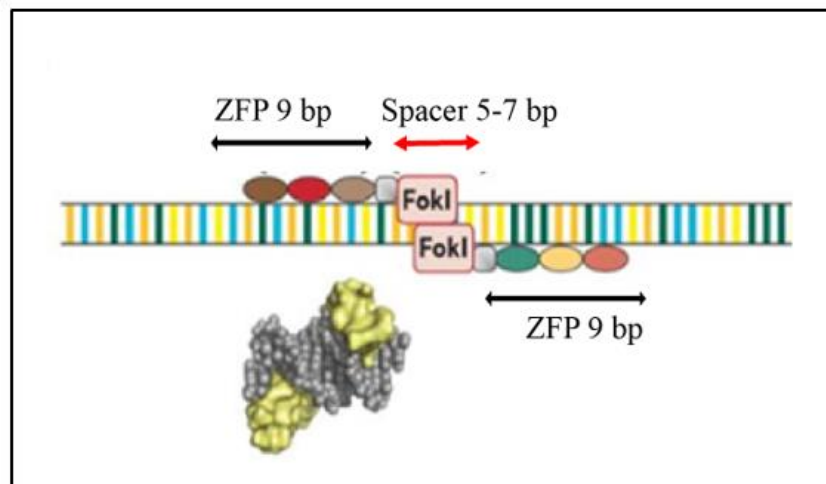
Figure 1.4. Zinc finger nuclease dimers covering 23-25 base length on target locus

The sequence specificity of ZFNs is provided by ZFPs which are essentially composed of Cys_2-His_2 zinc finger tandem arrays [86]. Cys_2-His_2 zinc fingers are defined as the most common DNA binding motif encountered in eukaryotes [87]. Each motif is compatible with 3 bases of DNA and generally 3-6 zinc fingers are used to produce a monomer of ZFN [86]. Even though ZFN-DNA interactions are flexible in nature, each zinc finger motif recognizes

and interacts with a specific 3 base pair of DNA sequence [88]. Zinc finger nucleases possess a modular nature that could be further customized by alterations using mutagenesis. Nevertheless, studies have reported that modified ZFNs may have cytotoxic effects or induce off target mutations out of the target genomic site [89]. Still, ZFNs are noted to have poor targeting efficiency due to the lack in available collection of 64 ZFPs to cover all possible combinations of 3 base pair DNA sequence [90].

The development of ZFNs as a promising genome engineering tool provided the generation of the first targeted knock out mammalian in 2009 [91]. In this study, rat embryos were microinjected with ZFNs to achieve knock out of Immunoglobulin M (IgM) and Rab38 genes. Like meganucleases, ZFNs were shown to efficiently induce DSBs and promote HR mechanisms at specific sequences [83]. In 2005, Urnov et al. has achieved to correct SCID mutation in human cells with a success rate of 18% using the advantage of ZFN induced DSB and HR provided by the co-transfection of exogenous DNA donor template [92]. ZFN technology has rapidly been adapted for the generation of transgenic animals in numerous laboratories. Successful outcomes of genome editing in mammalian embryos including mouse, rat, pig and cattle have been released one after the other [93-96].

In parallel to promising results, unwanted outcomes of ZFN mediated genome editing have been observed. One major and frequently reported issue has been the off-target effects of ZFNs. It was reported that ZFNs have failed to maintain a strict specificity when similar genomic sequences are available in the genome [97]. Despite the raising safety concerns and technical challenges, some of the pre-clinical studies have achieved to proceed through clinical trials. Clinical trials aiming the inactivation of CCR5, HIV co-receptor, expression has been granted with the approval of US regulatory authorities [98, 99]. After all, ZFN technology is currently accessible but not easily affordable for standard laboratories. ZFN technology has not been adapted for several cell types yet. Limited success and high off-target effects have been made this genome editing platform inconvenient even though it is patented and commercially available [100].

1.3.4. Transcription Activator-like Effector Nucleases (TALENs)

TALENs were reported in the field of genome engineering as a potential alternative to ZFNs in 2011. Like the ZFN technology, TALENs were first used to achieve efficient knockouts in rats [101]. Soon after, TALEN was extended to genome engineering in mouse and numerous new knockout models have been generated [102-104]. TALEN technology has been also applied in pig, cattle and primate embryos recently [105-107]. Beside sharing similarities with ZFNs, TALENs have some charming distinctions. Both ZFNs and TALENs are chimeric peptides engineered by the fusion of DNA binding domain and endonuclease domain. While they share the same endonuclease domain, FokI, TALENs have a structurally different binding domain known as transcription activator-like effectors (TALEs) derived from the plant pathogenic Xanthomonas spp. Bacterium [108]. TALEs are composed of 33-35 amino acid tandem repeats and each of these repeats recognises a single base in DNA major grove [109]. In TALENs, the nucleotide specificity is determined by the amino acids located in the 12th and 13th positions [110]. Another common feature of ZFNs and TALENs is that they both function as dimers. Therefore, a pair of TALENs must be designed to target the genomic location of interest. Figure 1.5 illustrates the schematic representation of a pair of TALENs. Target DNA sequences of a pair of TALENs are commonly 30-40 base in length excluding spacers which are 12-21 base pair length [111].
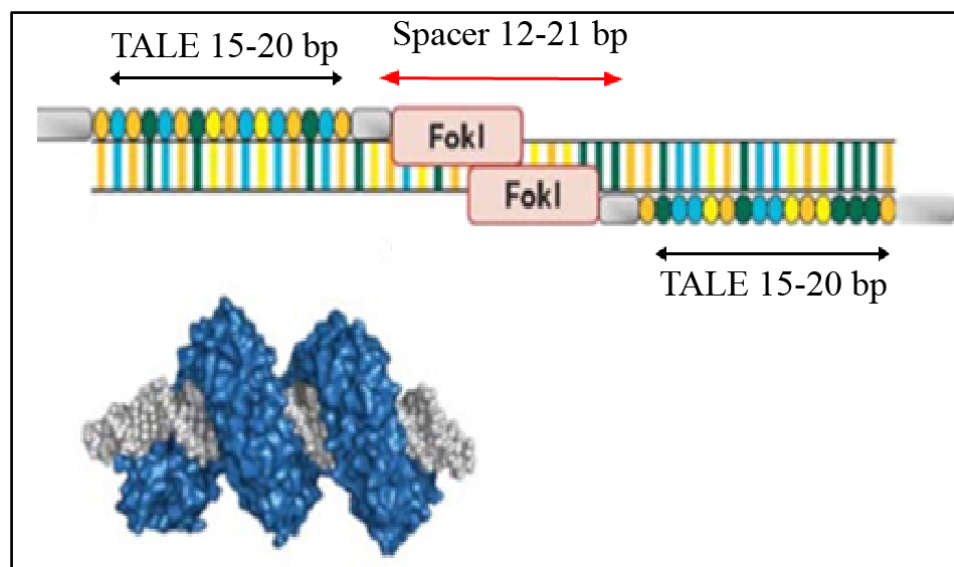
Figure 1.5. TALEN dimers covering 42-61 base length on target locus

Increased power of computational tools has enabled TALEN technology to present a wider repertoire for DNA targeting, design and engineering new TALENs [112]. In contrast to ZFNs, custom engineered TALENs are easier to generate in a basic molecular biology laboratory. Large collection of intermediate reagents are commercially available to construct a customized TALEN locally [113, 114].

In comparison to previous technology, TALENs can be engineered to target a wider range of DNA sequences, providing a remarkable advantage. The only limitation of designing TALENs is the requirement for a thymine at the 5' end of the target sequence [115]. Another drawback of TALEN technology is that TALENs are not able to induce DSB on a DNA sequence containing methylated cytosines [116].

TALEN technology has exhibited higher genome editing efficiency as a tool for the correction of in mouse in comparison to previous attempts [117, 118]. TALENs have also been a promising tool for gene therapy applications including preclinical and clinical trials. Recently, human iPSCs have been investigated to create genomic modifications and corrections for inherited disorders using TALEN technology [119-121].

1.3.5. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR Associated Proteins (Cas)

A novel genome editing tool with high efficiency was discovered in 2013, enabling the scientific community to drive genomic modifications more precisely, easier and cheaper. In nature CRISPR and Cas are the components of an adaptive immune system of prokaryotes and this system was discovered more than 2 decades ago [122]. CRISPR/Cas was first tested and used for gene editing in cell culture and subsequently mouse studies have been done [123, 124].

Briefly, CRISPR/Cas system is composed of a short RNA to be paired with the DNA sequence at the target site and a DNA endonuclease to induce DSB. The short RNA strand corresponds two parts in prokaryotes, crRNA and tracrRNA. The crRNA is responsible to pair with the target sequence while tracrRNA is the part which interacts with Cas endonuclease. Currently, the CRISPR/Cas system of *Streptococcus pyogenes* (Sp) is the

most commonly used tool for genome editing and the endonuclease of this system is Cas9 [125].

CRISPR/Cas is an RNA guided DNA cleavage system in prokaryotes which provides adaptive immunity against invading phages or plasmids [126]. Within this immune system, organisms capture small DNA fragments (around 20 bp) from the DNA of invaders and insert these sequences in CRISPR locus in their own genome. These captured archived sequences within CRISPR locus are named as protospacers. In type II CRISPR systems, CRISPR locus is transcribed as pre-CRISPR RNA (pre-crRNA) and matured into target specific crRNA. Trans-activating crRNA (tracrRNA) is also transcribed from the same locus. tracrRNa is invariable, target independent and contributes to the processing of pre-crRNA [127]. Following the transcription and processing, crRNA and tracrRNA forms a complex with CRISPR-associated protein 9 (Cas9) to produce an active RNA guided endonuclease. The active endonuclease is able to target and cleave 23 base pair length of DNA which is composed of 20 base pair guide sequence in crRNA derived from protospacer and 3 base pair of 5'-NGG-3' sequence called as protospacer adjacent motif (PAM) [128]. The 5'-NGG-3' PAM sequence is recognized by spCas9 while different PAM sequences are available for the Cas9 proteins derived from species other than *Streptococcus pyogenes* [129, 130]. In genome engineering applications, crRNA and tracrRNA is linked form a single synthetic guide RNA (sgRNA) to simplify the CRISPR/Cas tool [131]. Figure 1.6 shows the prokaryotic (left) and engineered CRISPR/Cas9 system (right) used in genome editing applications.
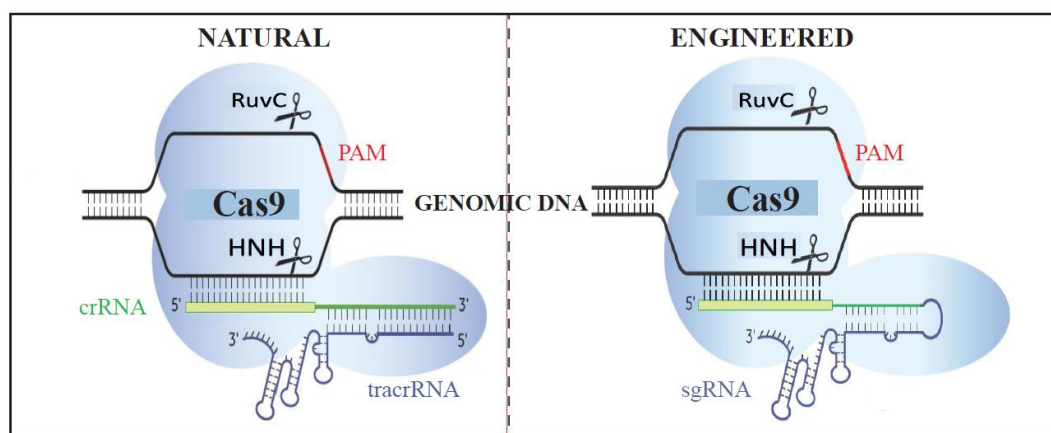
Figure 1.6. Components of CRISPR/Cas9 system.

The CRISPR locus is discovered to be present in around 45% of bacteria and 85% of archaea [132]. Essentially, the CRISPR locus is an array of tandem repeats separated by spacer sequences. Studies have shown that CRISPR can be found in both genomic and plasmid DNA [133]. The spacer sequences of CRISPR are derived from genomic material of phage or foreign plasmid, proving that CRISPR is a prokaryotic immune system [134, 135]. The spacers are the key elements for the recognition of genetic material which belongs to invaders. CRISPR locus has plasticity to provide the addition new spacers for the recognition of the new viruses attempting to infect the host cell. Since genome is modified after each addition of a new spacer, the record of the immune system is inherited to the offspring. Activity of CRISPR immune system is dependent on the presence of CRISPR associated (Cas) genes coding the proteins required for the immune response [136]. The CRISPR/Cas immune system can be mainly divided into three stages (Figure 1.7) [133]. As the first step, which is called adaptation, new spacer sequences are added to CRISPR locus. The second stage, expression, Cas genes are expressed, and pre-crRNA is transcribed from the CRISPR locus. The third and the last stage, interference, involves in the recognition and destruction of target nucleic acid.
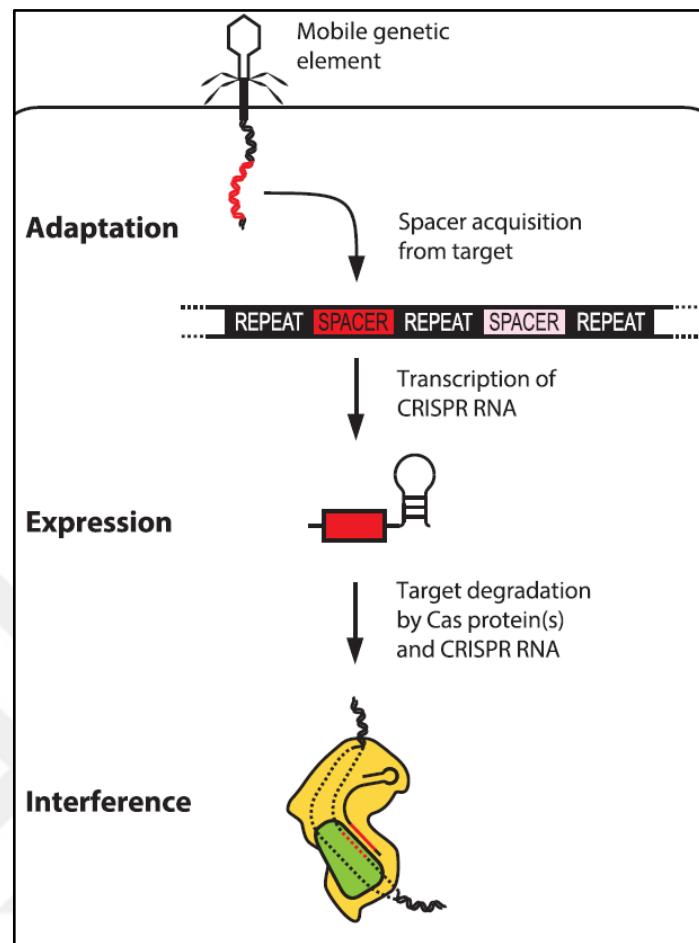
Figure 1.7. The key steps of CRISPR/Cas immune system in prokaryotes [133].

Cas proteins encoded by CRISPR locus have a high diversity in structure and function such as nuclease, helicase and RNA binding protein [137]. The proteins Cas1 and Cas2 are universal for CRISPR/Cas systems and are involved in the adaptation stage of immunity. The remaining Cas proteins are mostly associated with certain types of CRISPR/Cas systems in nature. The CRISPR/Cas systems have been classified according to variety of Cas proteins and the composition of CRISPR locus. Currently, there are 3 types of CRISPR/Cas systems identified; Type I, Type II and Type III [138]. The processing of crRNA and the interference stage show dissimilarities between CRISPR/Cas types (Figure 1.8) [133]. In Type I system, Cas5 and Cas6 mediate the maturation of pre-crRNA into crRNA. The interference stage requires Cas3 protein along with Cascade and crRNA. Cas3 has both helicase and DNase domain to induce DSB on target DNA [139]. Type II system uses RNase III and tracrRNA for maturation and 5' end trimming of crRNA. Cas9 is the signature protein for Type II system for DNA targeting and cleavage. Cas9 involves in adaptation stage of immunity and

participates in crRNA processing and induces DSB on target DNA guided by crRNA [131]. Type III system also has Cas6 for crRNA processing with the partnership of an unknown factor performing 3' end trimming. In Type III system, Csm/Cmr protein complex is responsible for nucleic acid targeting. Type III system also contains a signature protein called Cas10 however its function is still unclear. While Type I and Type II systems can only target DNA, Type III system can target both DNA and RNA [140]. An organism can host multiple CRISPR/Cas systems.
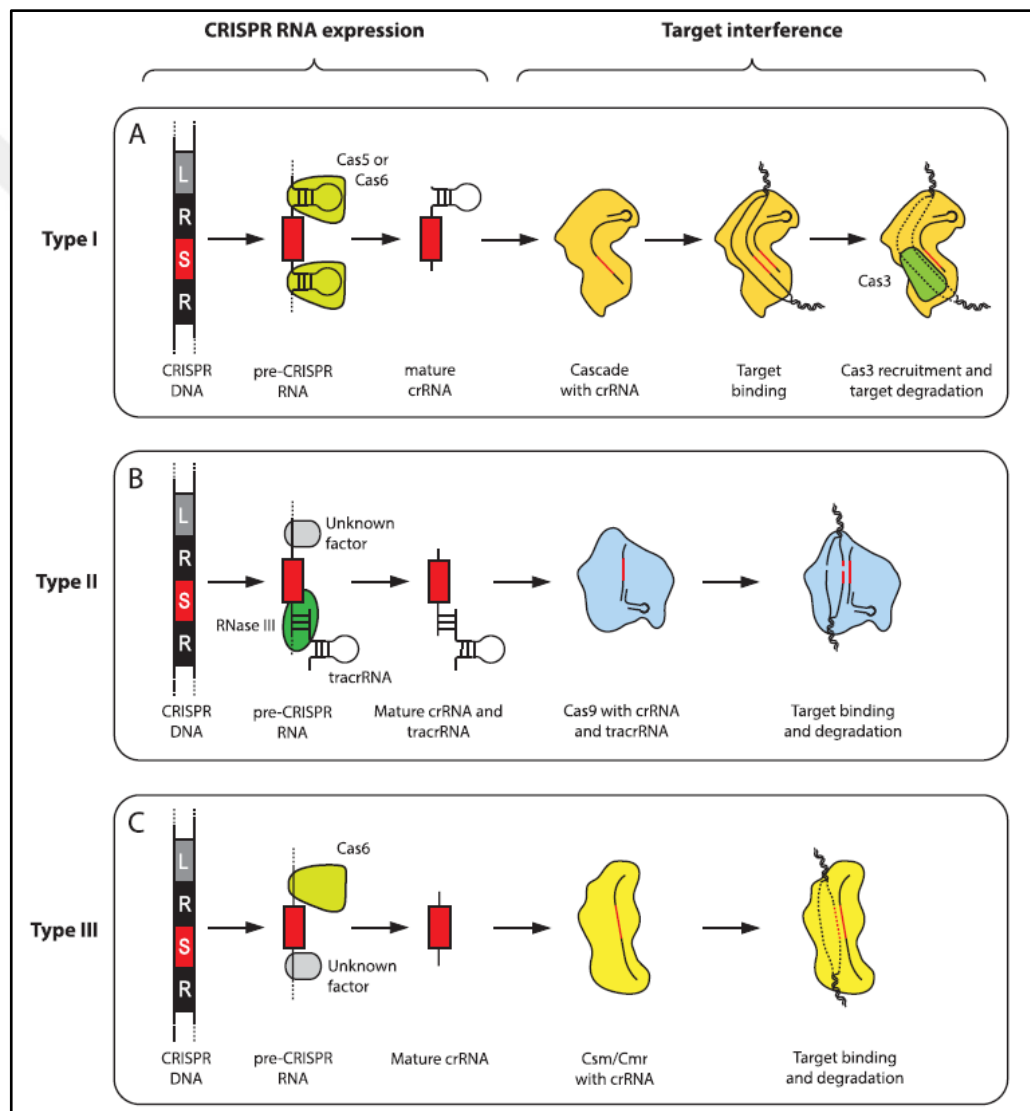
Figure 1.8. The comparison of CRISPR/Cas systems [133]

The mechanism of CRISPR for targeting and cleaving the DNA is analogous to that observed in ZFNs and TALENs. Essentially, these tools are binary systems generated by the partnership of an endonuclease domain and an element to direct it to the given DNA sequence. By contrast to ZFNs and TALENs which shares a common endonuclease domain named FokI, CRISPR system uses its own and natural endonuclease Cas, usually Cas9. Whereas ZFN and TALEN technology relies on DNA-protein interaction for the detection of target DNA sequences, CRISPR/Cas system is based on the nucleotide pairing code of the sgRNA molecule conjugated with the Cas9 endonuclease [141].

CRISPR/Cas technology promises simplicity and plasticity for design along with the proven efficiency to promote gene disruptions and editing applications [142]. These advantages have positioned CRISPR/Cas technology as the most popular genome editing tool and the most preferred method. Using CRISPR tools it is very easy to inactivate a locus through indels induced by NHEJ, or generate large deletions, insertions, knock-ins, inversions, duplications, substitutions or chromosomal rearrangements by HDR in numerous species [143-150].

CRISPR/Cas tool has rapidly been adapted to gene therapy applications and still there are a lot to explore. CRISPR has recently been used to correct or generate intentional mutations in human induced pluripotent stem cells for disease modelling to study molecular mechanisms and disease pathology [151]. Furthermore, human genetic disorders such as retinitis pigmentosa and Duchene muscular dystrophy have recently been studied to achieve restoration of gene functions using CRISPR/Cas9 components delivered by AAVs to mouse models [152, 153]. Preliminary studies to use CRISPR/Cas technology for genome editing in mammalian embryos have reported that mosaicisms and off targets effects are still major drawbacks as it was reported for the previous nuclease based technologies [154, 155].

CRISPR/Cas platform offers numerous advantages over meganucleases, ZFNs and TALENs in both somatic and pluripotent stem cells. CRISPR/Cas system is more user friendly than the previous nuclease platforms. Since the specificity of CRISPR/Cas9 is related to gRNAs, design and synthesis of wide range of possibilities is much easier and cheaper to conduct [156]. It is possible to design and try multiple gRNAs as many as the PAM sequence is available on the target sequence. The low cost of plasmid mediated CRISPR/Cas9 applications makes it more economical compare to ZFNs and TALENs. CRISPR/Cas9 tool

is currently the fastest option for genome editing. A typical application can successfully be completed within 2 weeks using CRISPR/Cas9 [157]. So far CRISPR/Cas9 has exhibited the highest editing efficiency in human stem cells. In human pluripotent stem cells, researchers have shown that CRISPR/Cas9 system achieved up to 80% editing [22].

In biological and medical research, CRISPR/Cas9 based genomic manipulations such as knock-in, knock-out, activation and interference have been widely utilized [158]. Stems cells are known as the main players in tissue repair and homeostasis. Therefore, stem cell-based cell and gene therapies are crucial for clinical applications. Remarkable effort and financial sources have been invested into the studies combining stem cell and CRISPR/Cas9 technologies [159].

1.4. CHOOSING THE CORRECT GENE EDITING TOOL

Each of the endonucleases previously discussed have its own advantages and drawbacks, therefore researchers should wisely choose which gene editing platform they will use. The very first feature of each endonuclease to consider is the success rate and activity. It is easy to presume that not all of the synthesized and engineered endonucleases would be equally efficient and functional. For instance, many ZFNs fail to induce DSBs at target locus both *in vivo* and *in vitro*. Despite the improvements in ZFN technology, it has been experienced that commercially available ZFNs usually function better than the ones generated locally [160]. In contrast to ZFNs and meganucleases, TALENs have shown nearly 90% success rate in mammalian cells unless there is a heavy methylation on target DNA sequence [116]. However, the mutation frequencies in mammalian cells targeted with TALEN technology has been in a range between 1% and 60% [161]. Similar to TALENs, CRISPR/Cas9 technology has a broad range of gene editing activity which is between 2% and 80% in mammalian cells [22, 162]. Briefly, the success rate of endonucleases highly depends on the cell type and method of delivery. Currently, there is no strict and reliable rules for the prediction of endonuclease activities without the validation by wet lab experiments.

Target specificity is another key factor to consider when choosing the type of nuclease for genome editing applications. The specificity of both ZFNs and TALENs can be improved by altering the number of zinc fingers and TALEs. Theoretically, ZFNs and TALENs

containing more modules can recognise longer sequences of DNA. Therefore, more specificity for the target site can be achieved. Nevertheless, there is always a possibility that these endonucleases may interact with other genomic sites having similar sequences and therefore partial interactions. Since both ZFNs and TALENs require dimerization for functionality of the FokI domains, high specificity is promised. In theory, endonucleases recognizing sequences more than 20 base pairs would have no off-target results in eukaryotic genomes. However, in reality, all three genome editing platforms exhibit off-target effects due to similarities in DNA sequences and the tolerance of mismatches [30, 163, 164]. Locally produced ZFNs have been reported as cytotoxic, which might be due to their excessive off-target effects [89]. TALENs and CRISPR/Cas9 are generally much less cytotoxic [165, 166]. The excitement for CRISPR/Cas9 technology was tempered due to several studies reporting off-target effects in mammalian cells [30, 128, 167, 168]. The off-target mutations caused by CRISPR/Cas9 system have reported to induce chromosomal translocations [169]. In CRISPR/Cas9 system, which is modulated by RNA guidance, the mismatches between the gRNA and target DNA can be tolerated in the 5' region. Thus off-target effects are inevitable within the current technology.

Multiplex genome editing application can be done using engineered endonucleases. CRISPR/Cas9 is the most convenient tool for the modification of multiple targets simultaneously. The reason is that both ZFNs and TALENs may form mismatched dimers since the users have to introduce pairs for each target [170]. CRISPR/Cas9 tool does not have such a limitation, therefore it overtakes other endonuclease-based tools in case of multiplex genome editing. This advantage promised by CRISPR/Cas9 system enables the rapid generation of multigene knockout laboratory animals for unlimited number of scientific studies [124, 171, 172].

The mutation signature is another significant subject while choosing the genome editing tool. Both ZFNs and TALENs have the same nuclease domain while their mutation patters are remarkably diverse. In the absence of homology repair template, TALENs produce deletion mutations much more than insertion mutations while ZFNs produce deletions and insertions with similar frequencies [173]. In CRISPR/Cas9 system, small insertions are more prevalent than deletions and big indels [131, 174].

1.5. DELIVERY OF ENDONUCLEASES AND DONOR TEMPLATES

For a successful gene editing process, endonucleases and the donor templates must efficiently be delivered to target cells. Engineered endonucleases can be delivered to cells in culture, embryos and whole organisms in various forms. These forms include plasmid DNA, mRNA, viral vectors and proteins. For all three types of endonucleases, method of delivery can be the same. Liposome mediated transfection or electroporation of plasmid DNA containing the endonuclease sequence can be performed to achieve transient expression of ZFNs, TALENs and Cas9 in various cell types [111, 175]. *In vitro* transcribed mRNAs encoding endonucleases can be injected into cells and embryos of several organism to induce genomic modifications [172, 176, 177]. Viral delivery is another option for the delivery of engineered endonucleases both *in vitro* and *in vivo*. Integrase deficient lentiviral vectors, adenoviruses and adeno associated viruses can be used for the delivery of ZFNs, TALENs and Cas9 partnered with gRNA. DNA donor templates can also be cloned and packaged into these vectors to perform HDR. Since ZFNs are relatively smaller in size, they can be packaged into smaller vectors such as adeno associated viral vectors. However, the size limitation of this vector does not allow to carry the templates for HDR. Instead, lentiviral vectors can be used in such applications [178]. Integrase deficient lentiviral vectors are not compatible for the delivery of TALENs since the TALE repeats may lead to undesired recombination in target cells due to the nature of vector [179]. Adenoviruses are attractive carriers due to their large cargo capacity for the delivery of endonuclease and repair template [179, 180]. Lentiviral vectors with genomic integration ability can be used for stable expression of Cas9 and gRNA in mammalian cell lines. Nevertheless, this method facilitates high rates of off-target mutations [181]. ZFNs, TALENs and Cas9 can be delivered to target cell in the form of purified proteins. In previous studies, hard to transfect cells were successfully delivered ZFN proteins [182, 183]. Purified Cas9 protein accompanied with gRNA, called as ribonucleoprotein, can directly be injected into embryos or delivered to cultured cells via electroporation/nucleofection [32, 184, 185].

1.6. SICKLE CELL DISEASE

Sickle cell disease (SCD) is known as a monogenic blood disease caused by a point mutation in human β-globin gene (HBB) encoding two subunits of tetrameric protein haemoglobin. Haemoglobin is a bidirectional carrier in respiratory system. It mediates transportation of O_2 from lungs to tissues. Moreover, haemoglobin facilitates the return transport of CO_2. Haemoglobin has four polypeptide chains composed of identical αβ dimers (α1β1 and α2β2) associate to form the haemoglobin tetramer with each having a heme group, a substituted porphyrin with a central iron. The subunits of haemoglobin binds O_2 via the heme group [186]. Figure 1.9 illustrates the tetrameric structure of haemoglobin.
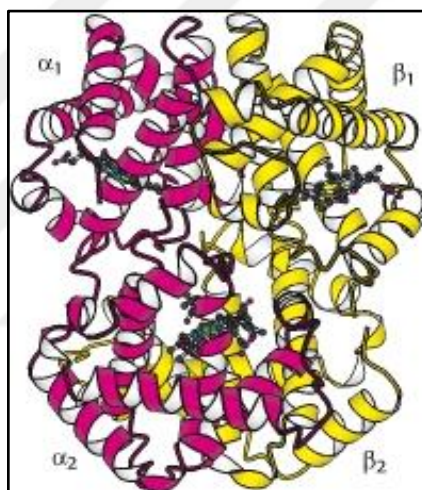
Figure 1.9. Molecular structure of haemoglobin tetramer [187]

Malformation of the haemoglobin protein structure leads formation of abnormally shaped red blood cells. SCD is caused by mutant copies of HBB called haemoglobin S (HbS). The point mutation, substation of A to T, in the sixth codon of HBB gene results in conversion of glutamic acid to valine and consequently an abnormal folding of haemoglobin emerges [188]. Abnormal HbS haemoglobin results in aggregation and polymerization of the protein, forming sickle shaped red blood cells. Unlike doughnut shaped, elastic normal red blood cells, sickle cells have a stiff and sharp sticky structure that easily aggregate and stick on narrow blood vessel interior surface (Figure 1.10). As the outcome of occlusion, insufficient oxygen is delivered to tissues and therefore organ damages are observed in long term of disease progress [189]. Moreover, sickle shaped red blood cells have remarkably shorter lifespan compare to normal red blood cells, causing chronic anaemia as a further

pathological impact on patients health [190]. Clinically, the presence of homozygous variant (HbSS) is the most severe case for SCD patients, in comparison to heterozygous mutants [188]. It has been reported that around 100,000 patients have been diagnosed with SCD in USA [191]. Approximately 300,000 children is born with SCD worldwide each year [192]. Despite of high frequency of SCD, still there is not any definitive treatment for this disease. Current treatments are predominantly available as supportive agents to reduce disease severity and background complications. Mostly used clinical applications for SCD patients are blood transfusion, hydroxyurea therapy and vaccinations to prevent the risk of severe infections which SCD patients are prone [193].
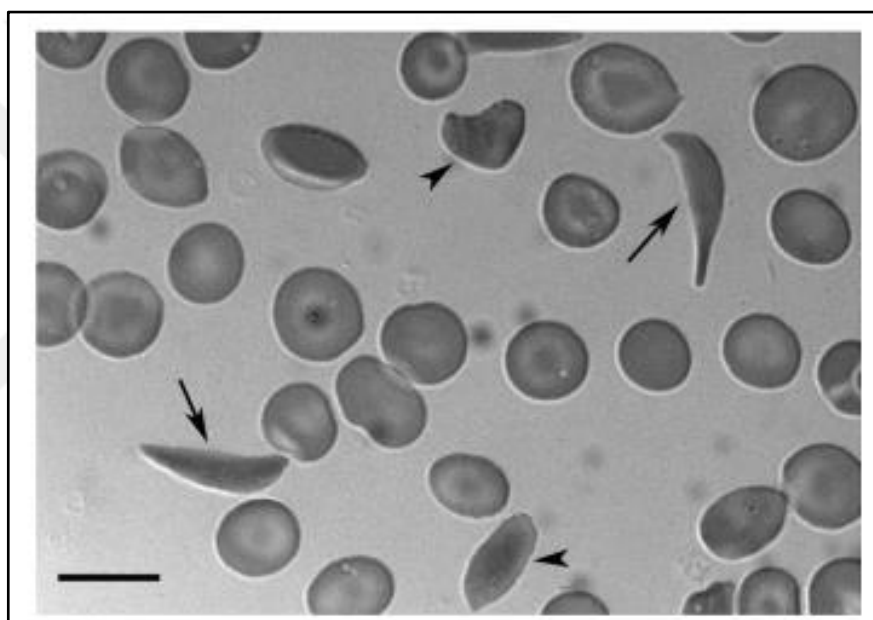
Figure 1.10. Normal and sickle shaped red blood cells [188]

Regarding that mature blood cells are derived from hematopoietic stem and progenitor cells (HSPCs), a clinical approach called allogenic stem cell transplantation has been used as a promising treatment for SCD and similar blood disorders. This technique relies on finding suitable donors, harvesting healthy HSPCs and transplanting to the patients [194]. Despite the successful results of this method, unfortunately it is not a universal treatment due to lack of available donors, immune response effects such as graft versus host disease (GVHD) and other side effects with several toxicities [195]. To eliminate the drawbacks of allogenic transplantation, autologous transplantation of ex vivo corrected HSPCs has been proposed as a promising method. Monogenic blood diseases such as SCD could be cured by direct

correction of mutations using genome editing tools which are also called engineered nucleases. These nucleases have demonstrated high potential for therapeutic applications in previous studies [196].

Previously, ZFNs and TALENs had been reported to successful for targeting and correction of SCD mutation on HBB gene up to a certain extend. Patient derived iPSCs with SCD mutation have been targeted using ZFNs with 9.8% efficiency [33]. A similar study have reported that iPSCs of SCD patients have been corrected without disturbing the subsequent differentiation efficiency into erythroid cell line however β-globin expression levels remained lower than healthy subjects [35]. Hoban et al. recently performed delivery of ZFNs in CD34+ HSPCs with up to 65% DSB induction rate. Despite the level of gene correction was 10-20%, repopulation of engrafted hematopoietic cells in bone marrow and spleen of immunocompromised mice remained insufficient for long term consideration [197]. The therapeutic potential of TALENs has also been investigated for SCD. Engineered TALENs was introduced the cells to induce DSB around the SCD mutation in HBB gene to demonstrate efficient targeting [198]. A follow up study was published using iPSCs with promising results of targeting efficiency [199]. Another group have showed that TALENs was a promising genome editing tool to correct SCD mutation in patient derived iPSCs. Corrected iPSCs were further differentiated into erythroid cells and the results demonstrated that 30-40% of the cell population with heterozygous wild type phenotype, which is clinically sufficient [200]. There has not been any study published concerning the SCD mutation correction in HSPCs using TALEN platform.

Recently, type II CRISPR/Cas9 system is the most fashionable tool for genome editing and promising approach for the direct correction of mutations causing monogenic diseases as SCD [201]. In comparison to ZFNs and TALENs, CRISPR/Cas9 system exhibited higher efficiency and lower cost while controlled targeting of HBB gene in iPSCs and K562 cell line derived from patient suffering chronic myeloid leukaemia [202, 203]. Huang et al. successfully corrected the mutation causing SCD, applying CRISPR/Cas9 system in iPSCs of SCD patients. In this study, researchers accomplished to preserve differentiation ability of edited iPSCs into erythrocyte cells with improved levels of β-globin expression [202]. Liang et al. gave a new impulse to HBB gene editing technology applying CRISPR/Cas9 system in human zygotes. However, various off-target mutations induced by Cas9 activity

and low efficiency of HDR were recorded, requiring more investigation for improvement [154]. In a recently published study, researchers attained a significant success at HDR efficiency through application of CRISPR/Cas9 system in CD34+ HSPCs [204]. DeWitt et al. introduced Cas9 and gRNAs as ribonucleoprotein complex along with ssDNA donor templates for desired correction. Researchers concluded that their study resulting in high gene editing rates for HBB locus and clinically significant recovery of WT β-globin production. Contrast to the previous attempts, lower but detectable off-targeting activity was observed in both HSPCs and K562 cells. Although the previous and recent scientific work has held promise for SCD patients, more research is required to alleviate off-targeting activity of CRISPR/Cas9 system and upgrade HDR yield in order to carry genome editing one step forward to clinical trials and become a universal therapeutic application.

2. MATERIALS

2.1. INSTRUMENTS

- ChemiDOC XRS+Gel Imaging System (Biorad, USA)
- 80 °C freezer (Thermo, USA)
- Centrifuge (Sigma, Germany)
- Electrophoresis Chamber (Biorad, USA)
- CO_2 Incubator (Thermo Scientific, YSA)
- NanoDrop Spectrophotometer (Thermo, USA)
- ZEISS Axio Vert.A1 – Inverted Microscope (ZEISS, Germany)
- CytoFLEX Flow Cytometer Platform (Beckman Coulter, USA)
- Vortex-Genie (USA Scientific)
- Water Bath (Memmert, Germany)
- Laminar flow cabinet (ESCO, Singapore)

2.2. EQUIPMENT

- Eppendorf tubes 2 mL (Isolab)
- Centrifuge tubes 15 mL, 50 mL (Isolab)
- Serological pipettes 5 mL, 10 mL, 25 mL (Corning)
- Micropipettes (Eppendorf)
- T150 Tissue Culture Flasks (TPP, 90150)
- 6 Well Plates (TPP)
- Cryopreservation vials (Thermo-Fischer, 374081)
- Haemocytometer (Sigma, Z359629)

2.3. CHEMICALS AND REAGENTS

- Dulbecco's phosphate-buffered saline (DPBS) (Gibco, 14190094)
- Trypsin-EDTA (0.25%) (Gibco, 25200056)
- Dimethyl Sulfoxide (DMSO) (Sigma, D2650)
- Dulbecco's Modified Eagle's Medium (DMEM), high glucose (Sigma, D5671)
- Fetal Bovine Serum (FBS) (Gibco, 10270106)
- Penicillin-Streptomycin (Gibco, 15070063)
- T4 Polynucleotide Kinase (NEB, M0201S)
- FastDigest BpiI (Thermo Scientific, FD1014)
- FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific, EF0654)
- TBE Buffer (Tris-borate-EDTA) (10X) (Thermo Scientific, B52)
- T4 DNA Ligase (Thermo, EL0014)
- LB Broth (Sigma, L3522)
- LB Agar (Sigma, 61746)
- Kanamycin Sulfate (Gibco 15160054)
- Ampicillin (Sigma-Aldrich, A1593)
- 50bp DNA Step Ladder (Promega, G4521)
- Lipofectamine® 2000 Transfection Reagent (Thermo, 11668027)
- Polyethylenimine (PEI) (Aldrich, 408727)
- Q5® Hot Start High-Fidelity 2X Master Mix (NEB, M0494S)
- Gel Loading Dye, Purple (6X) (NEB, B7024S)
- NEBuffer™ 2 (NEB, B7002S)
- Agarose I (Thermo, 17852)
- Taq 2X Master Mix (NEB, M0270L)
- EcoRI (NEB, R0101S)
- CloneAmp HiFi PCR Premix (Takara, 639298)
- FastDigest HindIII (Thermo, FD0504)
- Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550S)
- Quick-Load® 1 kb Extend DNA Ladder (NEB, N3239S)
- SCR7 pyrazine (Sigma, SML1546)

2.4. KITS

- NucleoBond® Xtra Midi (MN, 740410.50)
- PureLink Genomic DNA Mini Kit (Invitrogen, K182001)
- EnGen™ Mutation Detection Kit (NEB, E3321S)
- In-Fusion® HD Cloning Plus (Takara, 638910)
- NucleoSpin® Gel and PCR Clean-up Kit (MN, 740609.50)
- TA Cloning™ Kit (Invitrogen, K202020)

2.5. CELL LINES

- Human Chronic Myelogenous Leukemia K-562 (ATCC, CCL-243)
- Human embryonic kidney cell line HEK293T (ATCC, CRL-3216)

3.   METHODS

3.1. CELL CULTURE

3.1.1.   Cell Lines and Culture Conditions

Human embryonic kidney cell line (HEK293T) and Human Chronic Myelogenous Leukaemia cell line (K562) were purchased from ATCC (American Type Cell Collection). HEK293T cells were cultured in high glucose (4.5 g/L) Dulbecco's Modified Eagle Medium supplemented with 10% Fetal Bovine Serum (FBS) and 1% Penicillin-Streptomycin. K652 cells were cultured in RPMI-1640 medium supplemented with 10% Fetal Bovine Serum (FBS) and 1% Penicillin-Streptomycin. Cells were maintained at 37°C, 5% CO_2 incubator.

3.1.2.   Cell Passaging

HEK293T cells were passaged into a new flask when they reach 70-80% confluency. Inside a laminar hood, medium was removed from the flask and washed once with 1X Phosphate Buffered Saline (PBS). Trypsin was used to detach the cells from the surface of the flask. After addition of trypsin, flask was maintained in the incubator for 5-10 minutes. Detachment of the cells were monitored using light microscope. Fresh medium was added, and cells were collected into a falcon tube. Tube was centrifugated for 5 minutes at 300 x g. Supernatant was removed, and pellet was resuspended in fresh medium. Cells were counted using a haemocytometer under light microscope at 10X magnification. Adequate number of cells were passaged into a new flask and placed into the incubator.

K562 cells were passaged by collecting the medium into a falcon tube and centrifuge at 300 x g for 5 minutes. No trypsinization was required since it is suspension culture. Cells were resuspended in fresh medium and counted by haemocytometer. Adequate number of cells were transferred into a new flask and returned to the incubator.

3.1.3. Cell Freezing and Thawing

Freezing cells, or cryopreservation, were achieved by using a freezing mixture composed of 90% FBS and 10% DMSO. Cell suspension was prepared as explained in section 3.1.2 and centrifuged at 300 x g for 5 minutes. Pellet was resuspended in freezing mixture and aliquoted as 1 ml into cryopreservation vials. Vials were stored in -80°C, then transferred into liquid nitrogen tank. For cell thawing, cryopreservation vial was removed from storage, either -80°C or liquid nitrogen tank, and warmed up to 37°C, immediately. Fresh medium was added into the vial drop by drop to prevent cell damage, and cell suspension was transferred into falcon tube for centrifugation at 300 x g for 5 minutes. Cell pellet was resuspended in fresh medium and transferred into a flask. Medium was changed the next day to remove remaining DMSO.

3.2. DESIGN OF GRNAS FOR TARGET LOCUS

Exact genomic location of the most common mutation (rs334) causing SCD was determined using online SNP database of National Centre for Biotechnology Information (NCBI). Reference genomic sequence of HBB gene and neighbouring sequences were obtained from NCBI gene database. Using SnapGene desktop tool, exact site of rs334 was located and 100 bp length target sequence centring the possible mutation site was picked. Guide RNA sequences were determined using online gRNA design tool (available http://crispr.mit.edu/) by Zhang Lab, MIT. 13 candidate gRNA sequences were suggested by the tool, listing the on-target scores and possible off-target numbers for each gRNA. The gRNA sequence which has the highest score and closest distance to the rs334 was picked. Moreover, 7 bp longer version of this gRNA sequence was designed. The gRNA sequence obtained as 20 bp length was elongated from its 5' direction, resulting in 27 bp long gRNA (Table 3.1).

Table 3.1. gRNA sequences designed for targeting

Name	Sequence 5'-3'	PAM	Score
guide-SCD	GTCTGCCGTTACTGCCCTGT	GGG	71
guide-SCD-long	TGGAGAAGTCTGCCGTTACTGCCCTGT	GGG	71

Complementary sequences of the gRNAs were determined using SnapGene and adapter sequences were added to facilitate cloning of designed oligos before ordering, as shown in the (Figure 3.1).
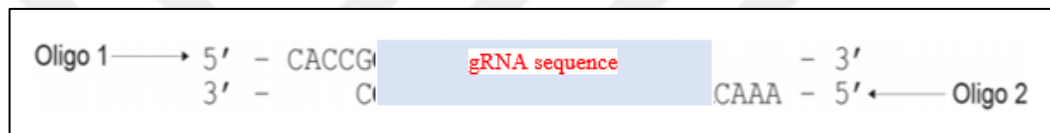
Figure 3.1. The final forms of gRNA sequences to be ordered as forward and reverse ssDNA oligos, with adapter sequences (grey) for cloning

3.3. CLONING sgRNA INTO PSPCAS9(BB)-2A-GFP

3.3.1. Annealing of Oligonucleotides

Annealing reaction mixture of ssDNA oligos were prepared in PCR tube as given below using T4 Polynucleotide Kinase:

Table 3.2. Reaction for ssDNA oligo annealing.

REAGENT	AMOUNT
Forward Oligo (100 µM)	1 µl
Reverse Oligo (100 µM)	1 µl
5X Rapid Ligation Buffer	2 µl
T4 PNK	1 µl
Nucleotide Free Water	5 µl

Prepared reactions were placed into the thermocycler and temperatures were set as below:

Table 3.3. Annealing conditions for thermocycler.

TEMPERATURE (°C)	TIME
37	30 minutes
95	5 minutes
25	Ramp down 95-25°C with 0.1°C /sec

Afterwards, annealed oligos were diluted 1:200 using nuclease free water and stored in -20°C immediately.

3.3.2. Digestion of the Plasmid with Restriction Enzymes

Digestion of pSpCas9(BB)-2A-GFP plasmid by BpiI enzyme was conducted by setting up the reaction mix given below:

Table 3.4. Digestion reaction of pSpCas9(BB)-2A-GFP.

REAGENT	AMOUNT
10X FastDigest Green Buffer	2 µl
Plasmid DNA	Up to 5 µg
BpiI FastDigest Enzyme	1 µl
Nuclease Free Water	Complete to 20 µl

Reaction mix was placed in thermocycler for 30 minutes at 37°C for incubation. Following incubation, FastAP Thermosensitive Alkaline Phosphatase was added into the digestion mix as given below:

Table 3.5. Alkaline phosphatase treatment.

REAGENT	AMOUNT (µl)
Alkaline Phosphatase	4
10X FastAP Buffer	3
Nuclease Free Water	3

Tube was placed back into the thermocycler and set to 37°C for 20 minutes followed by 80°C for 15 minutes.

3.3.3. Gel Purification

1%(w/v) agarose gel was prepared in TBE buffer. Digested and dephosphorylated plasmid mix were loaded into the well. Electroporation was conducted for 40 minutes at 150 V. After 40 minutes, agarose gel was visualized under UV transilluminator. Band around 9.2 kb was dissected using a scalpel and placed into Eppendorf tube, and the weight was measured.

NucleoSpin® Gel and PCR Clean-up Kit was used for plasmid purification from the gel. To solubilize the gel slice, 200 µl of NTI buffer was added for each 100 mg of agarose gel. Sample was incubated for 5-10 minutes at 50°C until the gel was dissolved, and vortexed briefly. Column was placed into a collection tube and up to 700 µl of sample was loaded. Sample was centrifugated at 11,000 x g for 30 seconds. Flow-through was discarded. To wash the silica membrane, 700 µl of NT3 Buffer was added into the column and centrifugated for 30 seconds at 11,000 x g. Flow-through was discarded. Column was centrifugated at 11,000 x g for 1 minute to dry the membrane. Column was placed into a clean Eppendorf tube and 30 µl of prewarmed NE buffer was added directly over the membrane. Sample was incubated at room temperature for 1 minute, then centrifugated for additional 1 minute at 11,000 x g. Concentration was measured using NanoDrop.

3.3.4. Ligation of the Plasmid Backbone and sgRNA Insert

Ligation reaction was set according to the Table given below:

Table 3.6. Ligation reaction.

REAGENT	AMOUNT
Purified Plasmid Backbone	50 ng
Annealed Oligos (1:200)	1 µl
T4 DNA Ligase	1 µl
10X T4 DNA Ligase Buffer	2 µl
Nuclease Free Water	Complete to 20 µl

Reaction mix was incubated in thermocycler at 22°C for 90 minutes. Tubes were left overnight at RT. Mix was used for bacterial transformation.

3.3.5. Bacterial Transformation by Heat Shock

Chemically competent DH5α cells (50 µl) were thaw on ice. Ligated plasmids were placed on ice for 10 minutes. 7 µl of ligation mix was added into bacteria by pipetting gently. Bacteria mixture was incubated on ice for 25 minutes. Tubes were placed on heat block adjusted to 42°C for 45 seconds and immediately placed on ice again for 3 minutes to provide heat shock. 250 µl of SOC medium was added to each tube. Cells were incubated in shaking incubator at 37°C, 150 rpm for 120 minutes. Inoculum was spread over Ampicillin (+) agar plates. Plates were incubated overnight in 37°C incubator. Colonies were observed the next day.

3.3.6. Plasmid Isolation via MidiPrep

To prepare the overnight culture for midiprep, 200 ml of LB Broth and 200 µl of Ampicillin was added into a flask. A colony was picked from the plate prepared in section 3.4 and mixed with LB Broth. Flask was placed into shacking incubator at 37°C overnight. Before starting on the plasmid isolation, glycerol stock was prepared by mixing 500 µl of bacteria sample and 500 µl of 50% glycerol into a cryopreservation vial. Vials were kept in -80°C. NucleoBond® Xtra Midi kit was used to perform plasmid isolation from bacteria culture. The inoculum was taken from the incubator and the volume was divided into 4 falcon tubes. Tubes were centrifuged at 6,000 x g, 10 minutes at 4°C. Supernatant was discarded. 8 ml of RES buffer was added into the falcon tube and pellet was resuspended. 8 ml of LYS buffer was added to lyse the cells. Tube was inverted 5 times and incubated at RT for 5 minutes. During this incubation, the column filter and column was prepared and equilibrated by addition of 12 ml of EQU buffer. 8 ml of NEU buffer was added into the tube and tube was inverted until sample turns colourless. Sample was loaded into the column. 5 ml of EQU buffer was added from the edges to wash the column. Filter was removed carefully, and 8 ml of Wash buffer was added into the column. The falcon tube at the bottom was changed and 5 ml of ELU buffer was added into the column. 0.7 volumes of RT isopropanol were added into the eluted sample. Tube was vortexed then incubated at RT for 2 minutes. Finalizer was attached to the syringe, pump was removed then sample was loaded in the syringe, pump was placed back and pushed slowly. Flow through was discarded. Finalizer

was removed then plug was removed and finalizer was placed again, 2 ml of %70 ethanol was added, pump was placed again and pushed slowly, flow through was discarded. Finalizer was dried until no ethanol was present. 300 µl of TRIS buffer was added into the syringe and DNA was eluted into Eppendorf tube. Concentration was measured using NanoDrop.

3.3.7. Polymerase Chain Reaction (PCR) Verification of sgRNA insert

The primers listed in the Table **3.7** were used for PCR amplification to test the presence of sgRNA inserts on plasmids. The primer Cloning Forward was common for both gRNA inserted plasmids, annealing U6 promoter upstream of the gRNA cloning site. For each sgRNA, original reverse oligos used in section 3.3.1 were used as reverse primers for PCR reaction.

Table 3.7. Primer sequences for PCR reaction

Primer Name	Sequence 5'-3'
Cloning Forward	GACTATCATATGCTTACCGT
Guide-SCD Reverse	AAACACAGGGCAGTAACGGCAGACC
Guide-SCD-long Reverse	AAACACAGGGCAGTAACGGCAGACTTCTCCAC

Stock primers (100 µM) were diluted to 1:20 using nuclease free water. PCR reaction was set according to the instructions given below:

Table 3.8. PCR reaction for sgRNA insert detection.

REAGENT	AMOUNT (µl)
2X Master Mix	10
Plasmid DNA	1
Forward Primer	1
Reverse Primer	1
Nuclease Free Water	7

Thermocycler was adjusted as given below:

Table 3.9. Thermocycler settings for PCR.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Initial denaturation	95	60	1
Denaturation	95	15	
Annealing	55	15	25
Extension	68	10	
Hold	4	∞	-

2% agarose gel was prepared using TBE buffer. Into each PCR product, 4 µl of 6X loading dye was added. 4 µl of 50 bp DNA Step Ladder was loaded into the first well. PCR products were loaded as 10 µl in each well. Gel was run at 150 V for 35 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.3.8 Verification of the sgRNA insert by Sanger Sequencing

Plasmids isolated by midiprep protocol was sent to Medsantek Company, Turkey for sequencing service. Singer directional Sanger sequencing was performed using 10 µM primer Cloning Forward given in Table 3.7 was used. The plasmid DNA concentrations was adjusted to 100 ng/µl.

3.4. OPTIMIZATION OF MAMMALIAN CELL TRANSFECTION

3.4.1. Lipofectamine

3x105 HEK293T cells were seeded to 6 well plates and cultured in 2 ml DMEM high glucose supplemented by 10% FBS, no antibiotics. Cells were incubated overnight at 37oC, 5% CO2 and 95% RH. In a 1.5 ml tube, 1 µg pSpCas9(BB)-2A-GFP-sgRNA was diluted in 100 µl serum-free DMEM. Control mix was prepared with pSpCas9(BB)-2A-GFP without sgRNA. 15 µl of Lipofectamine 2000 was diluted in 100 µl serum-free DMEM. Diluted plasmids were added to diluted Lipofectamine 2000 in 1:1 ratio. The mixture was incubated at room temperature for 15 minutes. The mixture was added to the culture media dropwise and mixed without disturbing the cells. Cells were placed in the incubator and incubated for 16-24 hours. GFP expression was checked under fluorescent microscope. If 70% or more cells expressing GFP was observed, the media was replaced with fresh DMEM (high glucose, 10%FBS, no antibiotics) and the cells were returned to the incubator and incubated for additional 72-96 hours. GFP expression and confluency was analysed. Cells were harvested by trypsinization for downstream applications.

3.4.2. Polyethylenimine (PEI)

HEK293T and K562 cells were seeded in 6 well plate with $3x10^5$ density and cultured in 2 ml DMEM high glucose supplemented by 10% FBS, no antibiotics. Cells were incubated overnight at 37ºC, 5% CO_2 and 95% RH. In a 1.5 ml tube, 1 µg pSpCas9(BB)-2A-GFP-sgRNA was diluted in 200 µl serum-free DMEM. Control mix was prepared with

pSpCas9(BB)-2A-GFP without sgRNA. 2 µg PEI (1 mg/ml stock) was added and mixed well (DNA: PEI ratio is 1:2). The mixture was incubated at room temperature for 15 minutes. The mixture was added to the culture media dropwise and mixed without disturbing the cells. Cells were placed in the incubator and incubated for 16-24 hours. GFP expression was checked under fluorescent microscope. If 70% or more cells expressing GFP was observed, the media was replaced with fresh DMEM (high glucose, 10% FBS, no antibiotics) and the cells were returned to the incubator and incubated for additional 72-96 hours. GFP expression and confluency was analysed. Cells were harvested by trypsinization for downstream applications.

3.5. SCREENING OF TRANSFECTED CELLS

3.5.1. Imaging by Fluorescent Microscope

24 hours after transfection, cells were examined under inverted fluorescence microscope (Zeiss Axio Vert.A1). Filter was set to FITC mode and imaging was performed with 5X and 10X magnifications. Images were recorded in TIFF format.

3.5.2. Flow Cytometer

Measurement of GFP expressing cell population was performed using Beckman Coulter's CytoFLEX flow cytometry device. HEK293T cells were harvested by trypsin and washed twice with PBS. Since K562 cells were suspension cells, treatment with trypsin was not done. Harvested cells were counted using haemocytometer and diluted to 5×10^5 cells in 200 µl PBS. Cell suspensions were transferred into 96 well plates and the plates were placed to the plate reader part of the device. Acquisition was set to 1×10^5 events read in live population. Live population was determined according to distinct aggregation on forward to side scatter plot. Gate was taken covering live population and GFP expression was analysed at FITC-A / FSC-A dot plot. Non-transfected cells were used as GFP negative control. Results were recorded as histogram and dot plot.

3.6. INDEL DETECTION OF SHORT AND LONG gRNAs

3.6.1. Genomic DNA Isolation

Cells harvested in section 3.4.2. are resuspended in 200 µl PBS. 20 µl of Proteinase K were added into the sample followed by addition of 20 µl of RNase A and mixed well via vortex and incubated at RT for 2 minutes. 200 µl of Lysis/Binding Buffer was added and mixed by vortexing. Sample was kept at 55°C for 10 minutes. 200 µl of absolute ethanol was added to the lysate and mixed well by vortexing for 5 seconds. Lysate was transferred into spin column. The column was centrifugated at 10,000 x g for 1 minute. Collection tube was discarded, and column was placed into a new collection tube. 500 µl Wash Buffer 1 was added to the column and centrifugated at 10,000 x g for 1 minute. Collection tube was replaced and 500 µl of Wash Buffer 2 was added to the column. Column was centrifugated at maximum speed for 3 minutes. Collection tube was discarded, and column was placed in a 1.5 Eppendorf tube. 50 µl of Genomic Elution Buffer was added to the column and incubated at RT for 1 minute. Tube with column was centrifugated at max speed for 1 minute. Elution step was repeated for maximum recovery of DNA. Column was discarded and DNA within the tube was measured with NanoDrop.

3.6.2. PCR Amplification of Target Locus

The primers listed in the Table 3.10 were used for PCR amplification of endonuclease targeted locus of HBB gene. Primers were designed to produce 1 kb products centring the expected target site.

Table 3.10. Primer sequences for PCR reaction

Primer Name	Sequence 5'-3'
SCD-Target-F	TCTGGAGACGCAGGAAGAGA
SCD-Target-R	ACGATCCTGAGACTTCCACAC

EnGen™ Mutation Detection Kit was used for PCR amplification. PCR reaction was established using genomic DNA isolated in section 3.6.1 as the following:

Table 3.11. PCR reaction of target locus.

REAGENT	AMOUNT
Q5 Hot Start High-Fidelity 2X Master Mix	10 µl
10 µM Forward Primer	1 µl
10 µM Reverse Primer	1 µl
Genomic DNA	300 ng
Nuclease-free water	To 20 µl

Annealing temperatures for the primers were found at https://tmcalculator.neb.com/ website and reaction were set on thermocycler accordingly.

Table 3.12. PCR settings for target locus.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Initial denaturation	98	30	1
Denaturation	98	7	
Annealing	Variable	12	35
Extension	72	20	
Final extension	72	120	1
Hold	4	∞	-

By using the control template and primer mix supplied by the kit, a separate PCR reaction was set as following:

Table 3.13. Control PCR reaction.

REAGENT	AMOUNT (µl)
Q5 Hot Start High-Fidelity 2X Master Mix	10
Control Template and Primer Mix	2
Nuclease-free water	8

Thermocycler was set for the control reaction as given below:

Table 3.14. PCR settings for control reaction.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Initial denaturation	98	30	1
Denaturation	98	7	
Annealing	65	12	35
Extension	72	20	
Final extension	72	120	1
Hold	4	∞	-

After the reactions, 5 µl of PCR product was spared for further applications and remaining was mixed with 1 µl Purple Gel Loading dye (6X). 1% agarose gel was prepared with TBE buffer and samples were loaded. Gel was run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.6.3. T7E Assay

From the 5 µl PCR product spared from section 3.6.2, heteroduplex was formed with the following reaction:

Table 3.15. Heteroduplex Formation Reaction.

REAGENT	AMOUNT (µl)
PCR product	5
10X NEBuffer 2	2
Nuclease-free water	12

Thermocycler was used to denature and anneal PCR product.

Table 3.16. Thermocycler settings for heteroduplex formation.

STEP	TEMPERATURE (°C)	RAMP RATE (ºC/min)	TIME (sec)
Denaturation	95	-	300
Annealing	95-85	1.5	5
Annealing	85-25	0.1	5
Hold	4	-	∞

To continue the heteroduplex digestion, 1 µl T7 Endonuclease I was added into the mixture and incubated for 20 minutes at 37°C in the thermocycler. Afterwards, 1 µl of Proteinase K was added to inhibit T7E activity and incubated 37°C for 5 minutes in the thermocycler. Sample was mixed with 4 µl of Purple Gel Loading dye(6X) and loaded on 1% agarose gel prepared with TBE Buffer and run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.7. DESIGN AND PRODUCTION OF HDR TEMPLATES

3.7.1.   PCR Amplification of WT gDNA with Varying Lengths

The primers listed in the Table 3.16 were used for PCR amplification of wild type genomic DNA sequence of HBB gene. Primers were designed to produce 1 kb, 2 kb and 4kb long products centring the expected target site.

Table 3.17. Primer sequences for PCR reaction

Primer Name	Sequence 5'-3'
HBB-1kb-F	TCTGGAGACGCAGGAAGAGA
HBB-1kb-R	ACGATCCTGAGACTTCCACAC
HBB-2kb-F	CCTACCCCTACTTTCTAAGTCA
HBB-2kb-R	GATACATTGTATCATTATTGCCCT
HBB-4kb-F	CAGGGATGTGAAACAGGGTCT
HBB-4kb-R	GGAGAAACCATCTCGCCGTA

PCR reaction was prepared using *Taq* 2X Master Mix to create 3 different PCR product lengths as 1 kb, 2 kb and 4 kb. Reaction was prepared as given below:

Table 3.18. PCR amplification from WT gDNA.

REAGENT	AMOUNT
Taq 2X Master Mix	12.5 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
Genomic DNA	50 ng
Nuclease-free water	To 25 µl

PCR reaction was set on thermocycler as given below:

Table 3.19. Thermocycler set-up for PCR amplification.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Initial denaturation	95	30	1
Denaturation	95	15	
Annealing	47	15	30
Extension	68	240/kb	
Final extension	68	120	1
Hold	4	∞	-

PCR products were mixed with 4 µl of Purple Gel Loading dye(6X) and loaded on 1% agarose gel prepared with TBE Buffer and run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.7.2. TA Cloning of PCR Fragments into pCRv2.1 Vector

After section 3.7.1, 1 kb and 2 kb PCR products were used for TA cloning into pCRv2.1 Vector. For this purpose, the following reaction was prepared:

Table 3.20. Ligation reaction for TA cloning of PCR products into pCRv2.1 vector.

REAGENT	AMOUNT (µl)
PCR Product	2
5X T4 DNA Ligase Reaction Buffer	2
pCR2.1 vector (25 ng/µL)	2
ExpressLink T4 DNA Ligase (5 units)	1
Nuclease-free water	3

Ligation reaction was incubated at RT for 15 minutes.

3.7.3. Bacterial Transformation and Blue-White Screening

Bacterial transformation of TA cloning products from section 3.7.2 was performed as explained in 3.3.5 using Stellar Competent Cells. After incubation at the shaking incubator for 120 minutes, bacteria were spread over agar plate containing 100 µg/ml Kanamycin, 40 µl of 40 mg/ml X-Gal and 40 µl of 100 mM IPTG. Plates were incubated at 37°C overnight and observed the next day.

3.7.4. Plasmid Isolation via MidiPrep

After transformation in section 3.7.3, 2 white colonies from 1 kb and 2 white colonies from 2 kb were picked and plasmid isolation via MidiPrep was performed as explained in section 3.3.6.

3.7.5. Size Determination by Agarose Gel Electrophoresis

Samples obtained from section 3.7.4, were loaded on 1% agarose gel to verify the insertion of the PCR products by product sizes. Gel was run at 150 V for 45 minutes and images were obtained from Biorad Chemidoc XRS device.

3.7.6. Diagnostic Digest of the Vector

For diagnostic digest of the pCRv2.1 Vector with 1 kb and 2 kb inserts, EcoRI restriction enzyme was used. Reaction was set as given below:

Table 3.21. Diagnostic digest reaction for WT inserts into pCRv2.1 vector.

REAGENT	AMOUNT
Plasmid	500 ng
10X Buffer	2 µl
EcoRI (10 U/µl)	1 µl
Nuclease-free water	To 20 µl

Reaction was incubated at 37°C for 1 hour. Sample was mixed with 4 µl of Purple Gel Loading dye(6X) and loaded on 1% agarose gel prepared with TBE Buffer and run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.7.7. Site Directed Mutagenesis of WT Fragments

Primers for site directed mutagenesis were designed using online In-Fusion Cloning Primer Design Tool (available https://www.takarabio.com/learning-centers/cloning/in-fusion-cloning-tools). The primers listed in the Table 3.21 were used for plasmid linearization and induction of sequence changes.

Table 3.22. Primer sequences for inverse PCR reaction to be used for in fusion cloning

Primer Name	Sequence 5'-3'
INF-SDM-F	CCTGTAAGCTTAGGTGAACGTGGATGAAGTTGG
INF-SDM-R	ACCTAAGCTTACAGGGCAGTAACGGCAGAC

PCR mix was prepared using the plasmids obtained from TA cloning in 3.7.4 and CloneAmp HiFi PCR Master Mix as given below:

Table 3.23. PCR Mix for linearization of pCRv2.1 with 1 kb and 2 kb inserts.

REAGENT	AMOUNT
CloneAmp HiFi PCR Premix	12.5 µl
Forward Primer	200 nM
Reverse Primer	200 nM
Plasmid	5 ng
Nuclease-free water	To 25 µl

Vector was linearized using inverse PCR with the mix prepared with the adjusted thermocycler as given below:

Table 3.24. Thermocycler settings for vector linearization.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Denaturation	95	15	
Annealing	47	15	30
Extension	78	240/kb	
Hold	4	∞	-

Following PCR, product was mixed with 4 µl of Purple Gel Loading dye(6X) and loaded on 1% agarose gel prepared and run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

In fusion reaction for site directed mutagenesis was prepared according to the Table below:

Table 3.25. In-Fusion reaction.

REAGENT	AMOUNT
Linear Product	100 ng
In-Fusion Enzyme	2 µl
Nuclease-free water	To 10 µl

Reaction was incubated at 50 °C for 15 minutes.

3.7.8. Bacterial Transformation

Bacterial transformation of Site Directed Mutagenesis products from section 3.7.7 was performed as explained in 3.3.5 using Stellar Competent Cells. Plates were observed the next day.

3.7.9. Plasmid Isolation via MidiPrep

After transformation in section 3.7.8, one colony was picked and plasmid isolation via MidiPrep was performed as explained in section 3.3.6.

3.7.10. Verification of SDM via Restriction Digest

Plasmid obtained from section 3.7.9 was digested using HindIII restriction enzyme with the reaction given below:

Table 3.26. Restriction digest reaction of SDM plasmid with HindIII.

REAGENT	AMOUNT (µl)
10X FastDigest Green buffer	2
FastDigest Enzyme	1
SDM Plasmid	10
Nuclease-free water	17

Reaction mixture was placed in thermocycler for 15 minutes at 37°C. Afterwards, products were loaded on 1% agarose gel and run at 150 V for 1 hour. Results are obtained from Biorad Chemidoc XRS device

3.8. TRANSFECTION OF HEK293T CELL LINE WITH SPCAS9(BB)-2A-GFP AND HDR TEMPLATES

3.8.1. PEI Transfection

Transfection HEK293T cells with SDM products (HDR template) obtained from section 3.7.9 and spCas9(BB)-2A-GFP was performed as explained in section 3.4.2. Mixtures were prepared as given in Table below and introduced into the cells.

Table 3.27. Mixtures for transformation via PEI.

TUBE #	spCas9(BB)-2A-GFP (μg)	1 kb donor (μg)	PEI (μg)	TUBE #	spCas9(BB)-2A-GFP (μg)	2 kb donor (μg)	PEI (μg)
1	0	1	2	7	0	1	2
2	1	0	2	8	1	0	2
3	0,5	0,5	2	9	0,5	0,5	2
4	0,5	1	3	10	0,5	1	3
5	1	1	4	11	1	1	4
6	1	2	6	12	1	2	6

12 hours after transfection, culture media was changed with fresh complete DMEM supplemented by SCR7 with final concentration of 1μM.

3.8.2. Flow Cytometry

Evaluation of GFP expressing cell population was performed according to the protocol given in 3.5.2. Procedure was conducted without any change.

3.8.3. Genomic DNA Isolation

Genomic DNA was isolated from transfected cells as explained in section 3.6.1.

3.8.4. PCR Amplification of Donor Template

The primers listed in the Table **3.26** were used for PCR amplification endonuclease targeted genomic DNA sequence of HBB gene. Primers were designed to produce 1 kb long products centring the expected target site.

Table 3.28. Primer sequences for PCR reaction

Primer Name	Sequence 5'-3'
HBB-1kb-F	TCTGGAGACGCAGGAAGAGA
HBB-1kb-R	ACGATCCTGAGACTTCCACAC

PCR was used to amplify the donor template on isolated gDNA in section 3.8.3. Reaction was set as given below:

Table 3.29. PCR amplification from isolated gDNA after transfection.

REAGENT	AMOUNT
Taq 2X Master Mix	12.5 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
Genomic DNA	50 ng
Nuclease-free water	To 25 µl

PCR reaction was set on thermocycler as given below:

Table 3.30. Thermocycler set-up for PCR amplification.

STEP	TEMPERATURE (°C)	TIME (sec)	CYCLES
Initial denaturation	95	30	1
Denaturation	95	15	
Annealing	50	15	30
Extension	68	60	
Final extension	68	300	1
Hold	4	∞	-

PCR products were mixed with 4 µl of Purple Gel Loading dye(6X) and loaded on 1% agarose gel prepared with TBE Buffer and run at 150 V for 45 minutes. Gel image was obtained from Biorad Chemidoc XRS device.

3.8.5. Restriction Digest with HindIII and Agarose Gel Electrophoresis

PCR product obtained from section 3.8.4 was digested using HindIII restriction enzyme with the reaction given below:

Table 3.31. Restriction digest reaction of PCR product with HindIII.

REAGENT	AMOUNT (µl)
10X FastDigest Green buffer	2
FastDigest Enzyme	1
PCR product	10
Nuclease-free water	17

Reaction mixture was placed in thermocycler for 15 minutes at 37°C. Afterwards, products were loaded on 1% agarose gel and run at 150 V for 1 hour. Results are obtained from Biorad Chemidoc XRS device.

4.   RESULTS

4.1. VERIFICATION OF SGRNA INSERTS INTO PSPCAS9(BB)-2A-GFP

To confirm the successful cloning of sgRNA sequences into pSpCas9(BB)-2A-GFP plasmids, PCR amplification for the cloning site was performed. In PCR reaction, the primer named Cloning-F was used as the standard forward primer annealing the U6 promoter located upstream of the sgRNA cloning site. The reverse strands of oligo DNAs previously designed and used for sgRNA insert production were used as the reverse primer for each sgRNA inserted plasmid. It was expected to obtain a 100 bp PCR product if the sgRNA sequence is successfully cloned in to plasmid. In the absence of sgRNA insertion, reverse primer would not anneal to template plasmid DNA therefore there would not be any PCR product visualized in agarose gel. Cloning results were also verified with Sanger sequencing method.

4.1.1.   PCR Validation of sgRNA inserts

PCR reactions to assess whether the sgRNA cloning was successful or not were run on 1% (w/v) agarose gel and visualized in Biorad Chemidoc system (Figure 4.1).
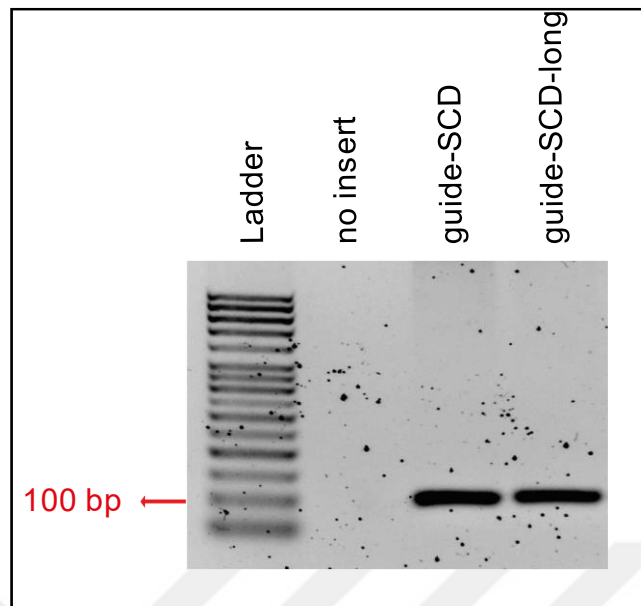
Figure 4.1. Agarose gel image of PCR products. HyperLadder™ 50bp – Bioline was used DNA ladder for size determination.

PCR result indicates the successful cloning process of two sgRNA sequences, guide-SCD and guide-SCD-long, into pSpCas9(BB)-2A-GFP. The plasmid without any sgRNA insert did not produce any 100 bp PCR products while gRNA cloned plasmids did, indicating the presence of gRNA sequences downstream of the U6 promoter.

4.1.2. Sanger Sequencing of gRNA Cloned Plasmids

Service for sequencing the plasmids was purchased from Medsantek Company, Turkey. Sanger sequencing results verified the PCR results and confirmed correct orientation of the inserts (Figure 4.2).
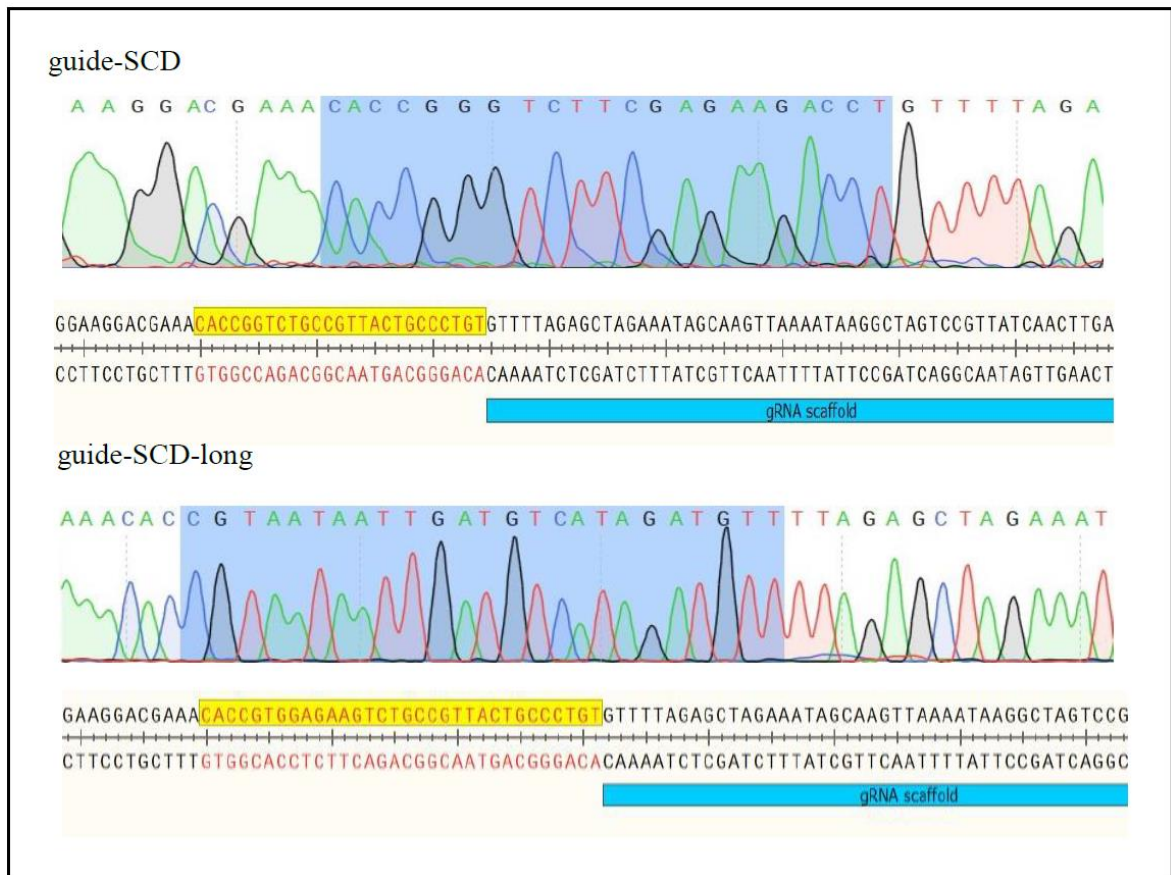
Figure 4.2. Sanger sequencing results for the sgRNA cloned plasmids

The results have shown that both gRNA sequences were successfully cloned into plasmids and ready for downstream applications. Locations of each insert are the correct sites to facilitate sgRNA and gRNA scaffold expressions in transfected cells.

4.2. OPTIMIZATION AND SCREENING OF MAMMALIAN CELL TRANSFECTION

Equal number of cultured HEK293T cells were transfected with the same amount of pSpCas9(BB)-2A-GFP plasmids using Lipofectamine 2000 and PEI transfection reagents to figure out the most efficient way to deliver Cas9 endonuclease, sgRNA and HDR templates. 24 hours post-transfection, cells were examined under fluorescent microscope with 5X magnification (Figure 4.3).
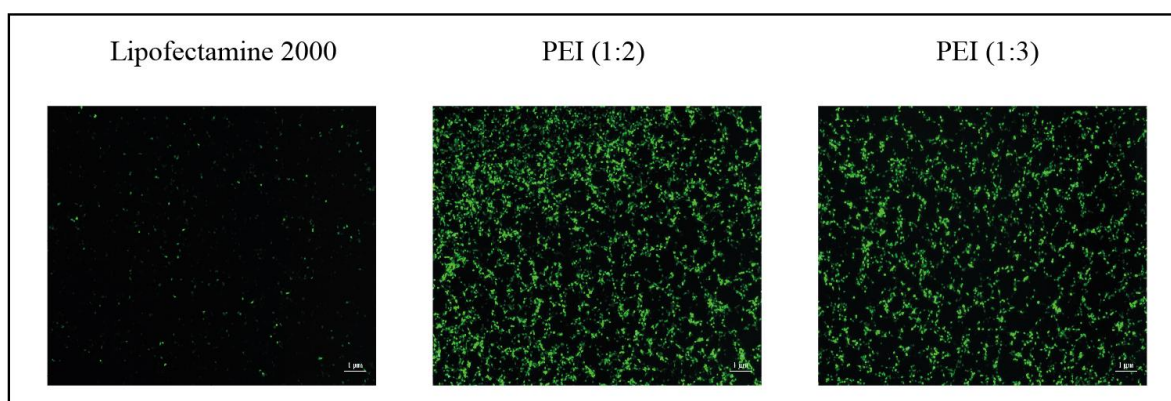
Figure 4.3. GFP expressing HEK293T cells transfected with pSpCas9(BB)-2A-GFP plasmid using Lipofectamine and PEI with two different DNA:PEI ratios.

Each well containing 300,000 cells were transfected with 1 µg of plasmid. For the transfection with Lipofectamine reagent, 5 µl of the reagent was used. In PEI transfections, to obtain 1:2 and 1:3 DNA:PEI ratios, 2 µg and 3 µg of PEI were used respectively. Based on the observations made using fluorescent microscopy imaging, Lipofectamine 2000 demonstrated the weakest efficiency for transfection. PEI has shown much better performance in term of transfection efficiency and GFP expression. In comparison to GFP expressing cells transfected using Lipofectamine, cells transfected using PEI has showed higher expression of GFP. Although the highest transfection efficiency was demonstrated by PEI, Lipofectamine was observed less toxic to the cells. The increased amount of PEI resulted in an increase in cell death. To conclude, it was decided to use PEI as a transfection reagent with 1:2 DNA to PEI ratio for HEK293T cells in this study.

The same protocol of PEI transfection (1:2) was applied to K562 cell line in suspension culture. 48 hours after transfection, cells were collected and the GFP expression was examined using flow cytometry (Figure 4.4).
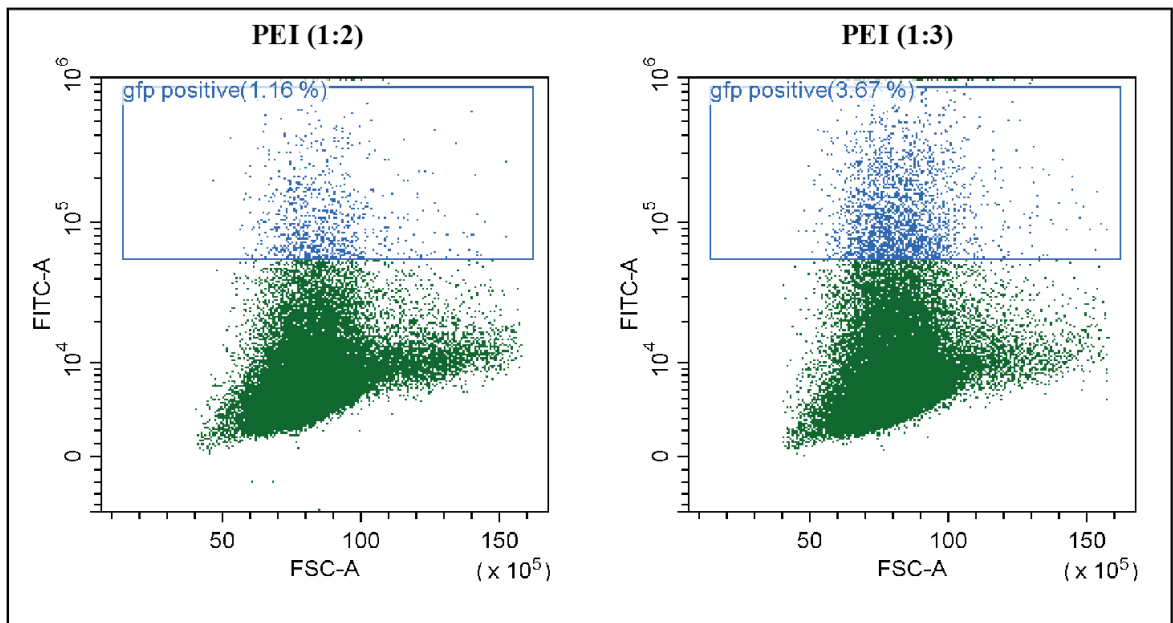
Figure 4.4. GFP expressing K562 cells transfected with pSpCas9(BB)-2A-GFP using PEI with two different DNA:PEI ratios

In K562 cell line, PEI transfection method was not as successful as it was in HEK293T cells. We observed over 2-fold increase in GFP expressing population when PEI was increased from 2 µg to 3 µg. Since PEI was expected to be highly toxic, higher doses of the reagent was not taken into consideration. Nonetheless, 1-4% transfection yield was not enough to proceed through downstream or high throughput application using K562 cells and PEI. In conclusion, the data states that PEI transfection is efficient for HEK293T cells but not suitable work with K562 cell line.

PEI transfection (1:2) was repeated for HEK293T cells and sgRNA cloned plasmids. 24 hours after transfection cells were harvested by trypsinization and GFP expression was analysed using flow cytometry (Figure 4.5).
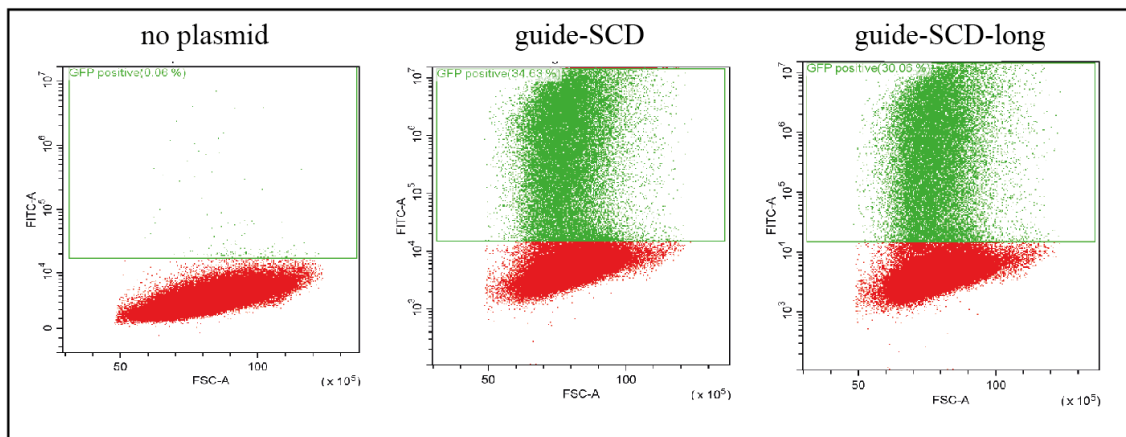
Figure 4.5. GFP expressing HEK293T cells transfected with sgRNA cloned pSpCas9(BB)-2A-GFP plasmids using PEI

wo plasmids containing guide-SCD and guide-SCD-long showed similar transfection efficiencies, 34.6 % and 30% respectively. The data indicates that both gRNA expressing plasmids are functional and exhibit close performance in terms of transfection yield, GFP therefore Cas9 and sgRNA expressions.

4.3. DETECTION OF ON-TARGET GENOME EDITING

T7 Endonuclease (T7E) based mutation detection kit was used to evaluate the activity of both sgRNAs on target locus. Briefly, genomic DNAs were isolated from HEK293T cells transfected with guide-SCD and guide-SCD-long and Cas9 expressing cells and the target locus was amplified by high fidelity PCR. Primers for the PCR was designed to produce 1 kb product, centring the DNA sequence targeted by gRNAs. It was expected to generate 500 bp band on agarose gel after T7E treatment to PCR products if the gRNAs successfully induce DSBs resulting in indel mutations. PCR products derived from genome targeted cells were proceeded to mutation detection assay, followed by agarose gel electrophoresis (Figure 4.6).
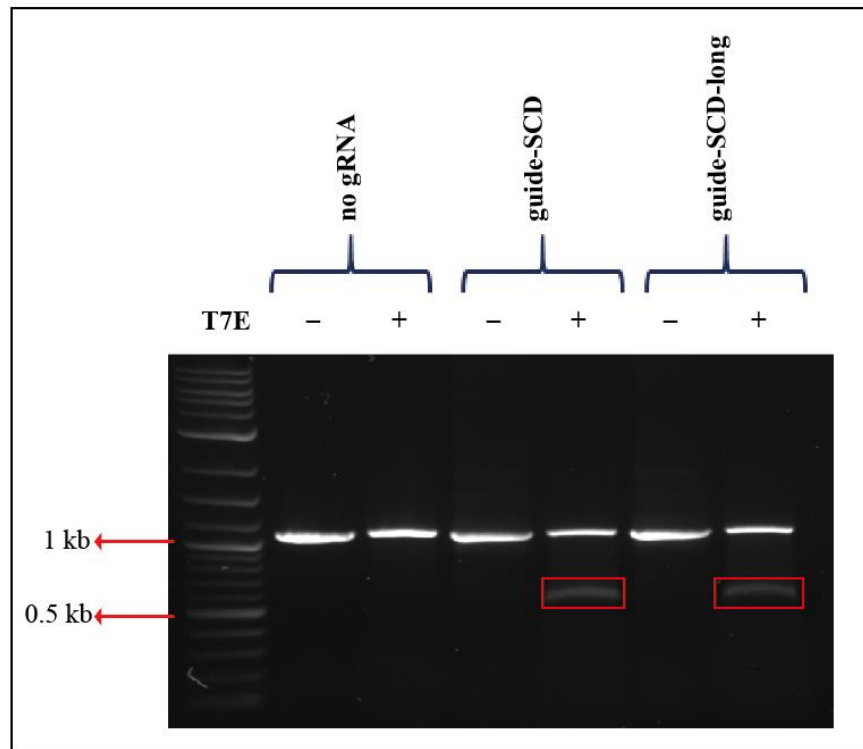
Figure 4.6. T7E assay for HEK293T cells transfected with pSpCas9(BB)-2A-GFP only, pSpCas9(BB)-2A-GFP + guide-SCD and pSpCas9(BB)-2A-GFP +guide-SCD-long respectively

The data indicates that both guide-SCD and guide-SCD-long successfully functioned and showed equal activity of cleavage on target DNA sequence. Cells transfected with Cas9 only did not show any DSB on target locus since there was not a gRNA for gene targeting. The experiment was performed with initially equal amount of genomic DNA and PCR products to be able to evaluate relative activities of each guide RNA. According to the gel image, longer gRNA has showed indistinguishably performance of targeting and cleavage on target locus, in comparison to standard length gRNA. In theory, longer gRNAs would have less off-target effects on genome. Since long gRNA has the same targeting efficiency and functionality on target site, it could be preferred to use for future application rather than using standard length gRNA. Thus, it would be possible to reduce off-target mutations without losing on-target activity.

4.4. GENERATION OF WT DNA FRAGMENTS AS STARTING MATERIALS FOR HDR TEMPLATES

Wild type genomic DNA from HEK293T cells were isolated and target locus was amplified by PCR. Primer pairs were designed to produce 1kb, 2kb and 4kb DNA fragments centring the gRNA binding site. PCR reactions were prepared, and product sizes were examined by agarose gel electrophoresis (Figure 4.7).
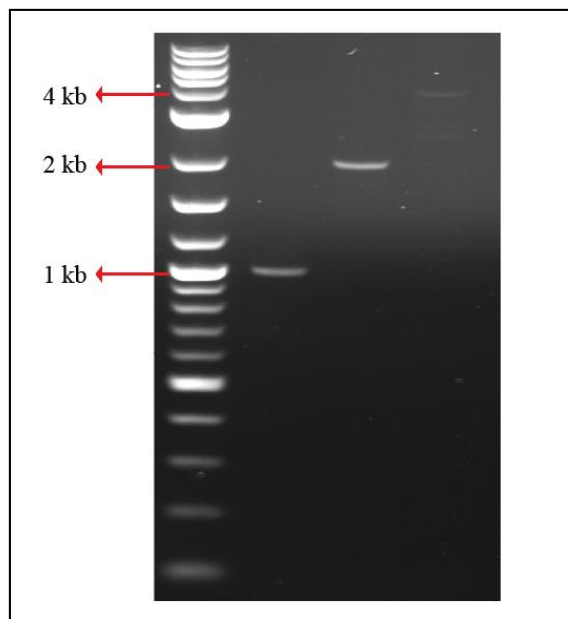
Figure 4.7. Agarose gel image for PCR products

The result indicates that 1 kb and 2 kb fragments were successfully produced in correct size and without any unspecific products. However, 4 kb fragment could not be properly produced after repetitive trials. The main cause of this problem might be due to the limitations of Taq polymerase used in this experiment. In this experiment, usage of standard Taq polymerase was required to obtain products with TA overhangs. Taking the advantage of this property of Taq polymerase, PCR products could serve as inserts for TA cloning. Overall, downstream procedures were done using 1 kb and 2kb fragments due to inefficient amplification of the 4 kb fragment.

4.5. TA CLONING OF WT DNA FRAGMENTS

1 and 2 kb PCR products obtained from previous section were directly used for TA cloning protocol. The main reason to clone these fragments into a plasmid was to produce higher amounts through bacterial transformation and prepare for site directed mutagenesis protocol to create desired modifications on donor templates. Briefly, wild type DNA fragments were ligated to linearized pCR 2.1 plasmid and bacterial transformation was performed. This plasmid has both ampicillin and kanamycin resistance genes for selection. The cloning site of this plasmid is in LacZalpha sequence, enabling the blue-white screening of transformed colonies (Figure 4.8).
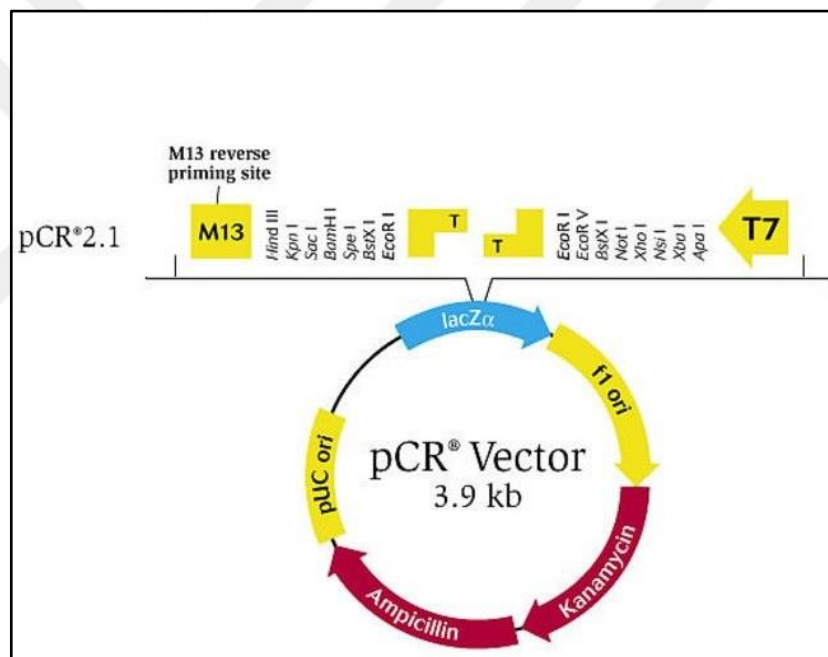
Figure 4.8. Basic plasmid map of pCR 2.1

Bacteria transformed with ligation reaction mixture were spread to agar plates containing kanamycin for selection and IPTG/X-gal for the blue-white screening (Figure 4.9). In theory, blue colonies are indicators of successful insertion in LacZalpha sequence, interrupting the production of functional beta-galactosidase enzyme. Non-functional enzyme lacks the ability to catalyse the breakdown of X-gal into blue product.
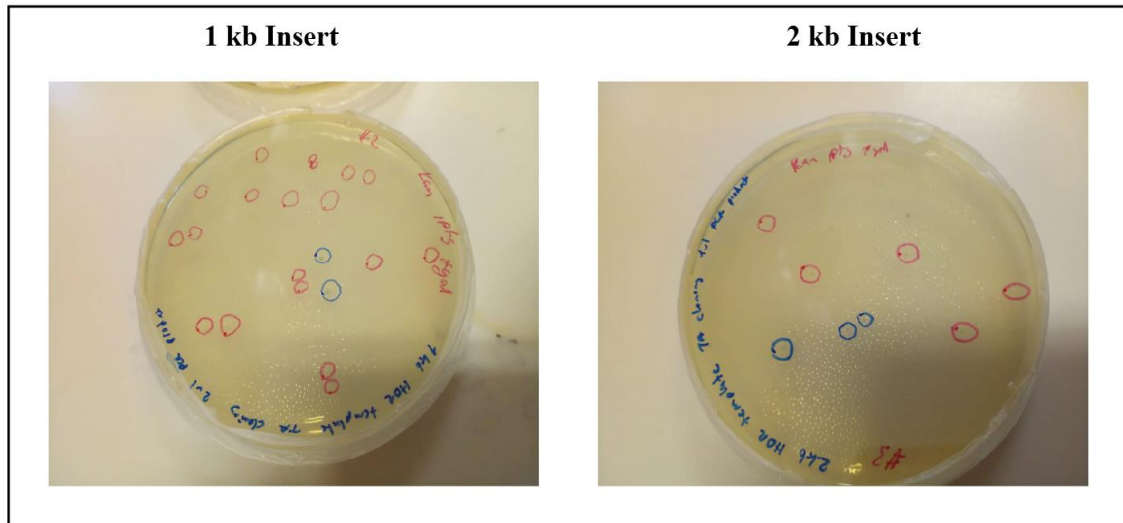
Figure 4.9. Blue-white screening of bacterial colonies transformed with pCR 2.1 plasmids containing 1 kb and 2 kb WT fragments. Blue colonies were circled with blue marker, and white colonies were circled with red marker

The white coloured colonies were the indicator of successful cloning of fragments into plasmids while blue coloured colonies were the results of self-ligated plasmids without insert. Thus, the single colonies with white colour were picked from agar plates to grow and retrieve the plasmids for analysis and site directed mutagenesis.

Two white colonies for each plasmid transformation were picked and plasmids were purified. Later, plasmids were run on 1% (w/v) agarose gel for size determination (Figure 4.10).
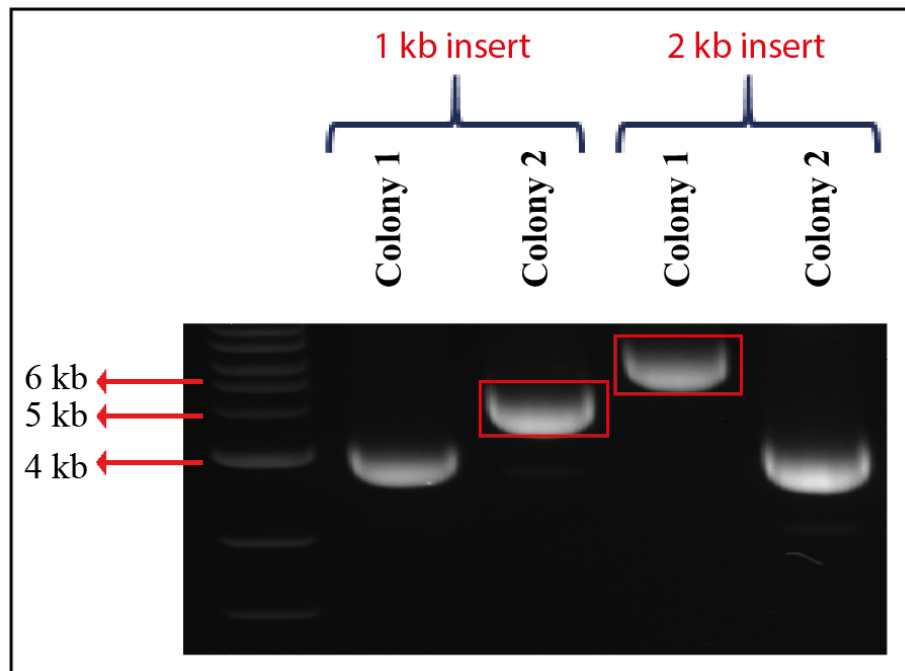
Figure 4.10. Size determination of the plasmids isolated.

The size of pCR 2.1 plasmid is 3.9 kb and 2 different fragments, 1kb and 2 kb, were cloned into these plasmids yielding 5 kb and 6 kb total sizes respectively. The gel image shows that the Colony 2 for 1 kb inserted plasmid transformation and the Colony 1 for the 2 kb inserted plasmid were correct in size. Although the Colony 1 for 1 kb insert and the Colony 2 for the 2 kb insert were white coloured, results were false positive. This might be caused by uneven spreading of IPTG/X-gal mixture on the plate.

To ensure the insert size in the selected plasmids, EcoRI digestion was performed. Referring the plasmids map in Figure 4.8, the fragments inserted were surrounded by two EcoRI restriction sites. Following the digestion reaction, samples were loaded to 1% (w/v) agarose gel and electrophoresis was performed. Gel was visualised for the easement of insert sizes (Figure 4.11).
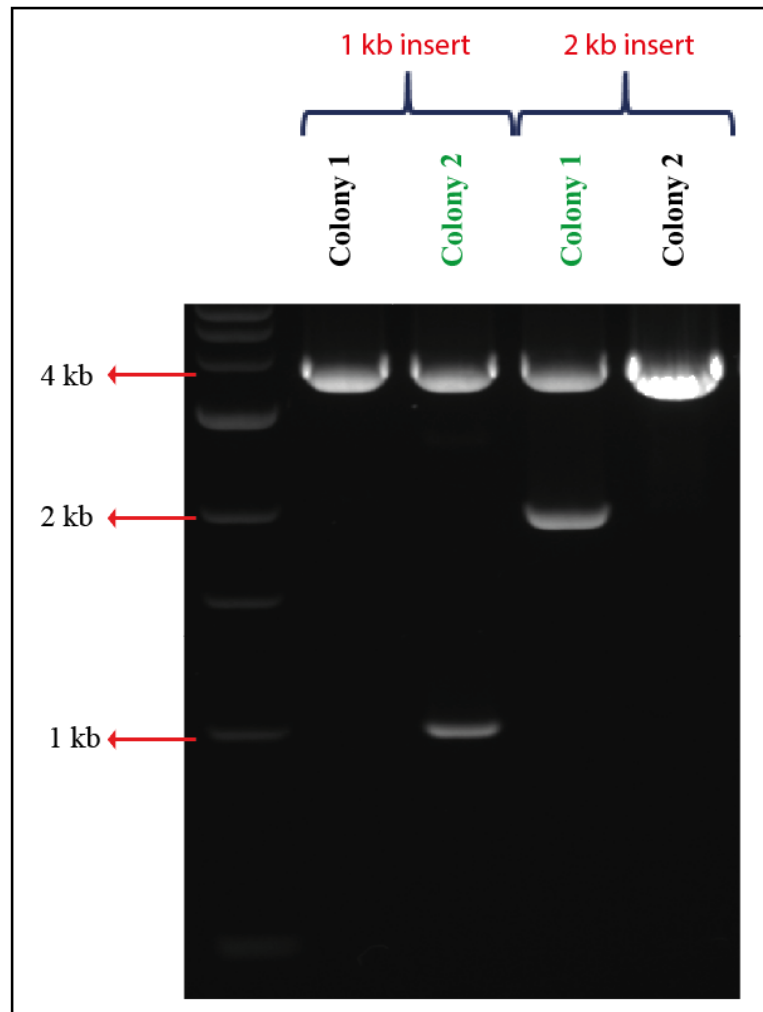
Figure 4.11. Agarose gel image of EcoRI digested plasmids to reveal insert size.

Digestion with EcoRI confirmed the correct insertions into pCR 2.1 plasmids. Colony 2 for 1 kb inserted plasmid transformation and the Colony 1 for the 2 kb inserted plasmid demonstrated successful TA cloning of wild type DNA fragments of target locus. In conclusion, plasmids with 5 kb and 6 kb total size were selected for the site directed mutagenesis step and the rest were discarded.

4.6. GNERATION OF HDR TEMPLATES

In this study, it was aimed to achieve targeted genomic modification via HDR mechanism. PAM sequence on the wild type sequence was aimed to be replaced with HindIII restriction

enzyme recognition sequence, which was absent in the wild sequence. In fusion cloning system used for the generation of this modification on template DNAs.

The method was based on inverse PCR using special primers containing the sequence changes. The first step of the site directed mutagenesis procedure was linearization of plasmid obtained from the previous section. Following the inverse PCR reaction using specific primers, products were run on agarose gel and size determination was done (Figure 4.12).
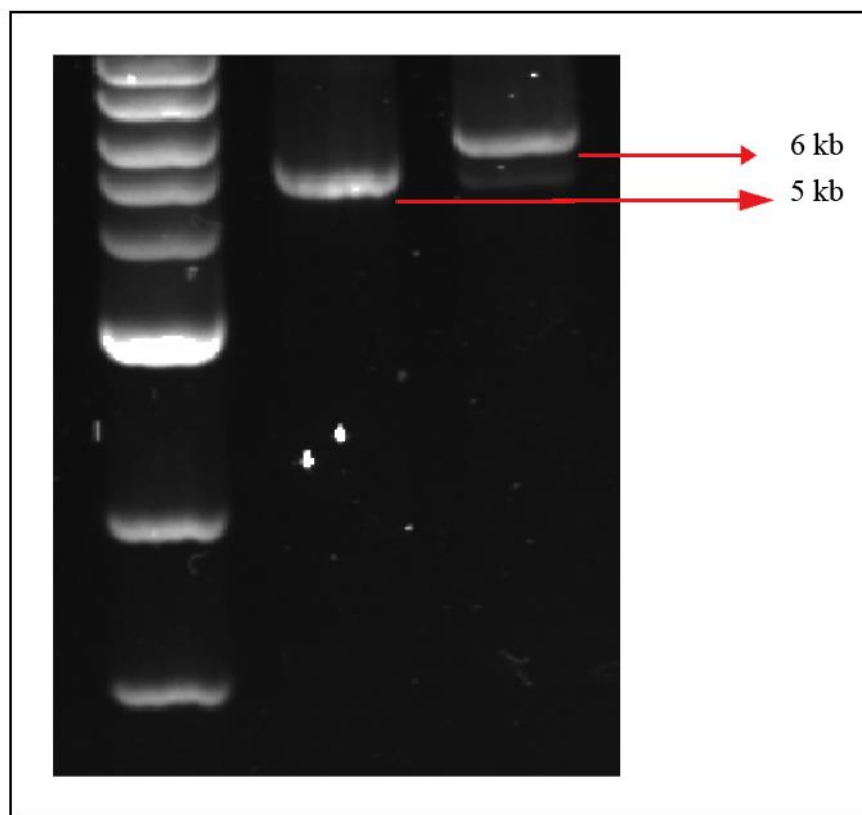
Figure 4.12. Agarose gel image of linearized plasmids with 1 and 2 kb inserts, 5 and 6 kb in total size

The image of agarose gel electrophoresis suggests that the inverse PCR reaction was performed efficiently, yielding linearized products with correct sizes. Moreover, the absence of unspecific bands and smears on the gel image indicates that PCR products could be used directly for the in fusion reaction, without any need for gel purification.

In fusion reaction was prepared and run to induce circularization of plasmids with targeted sequence changes. As the next step, chemically competent bacteria were transformed with reaction mixture. Colony selection and screening was performed, plasmid DNAs were isolated by midiprep protocol. In this part of study, it was aimed to introduce an extra HindIII recognition site erased the PAM sequence while the plasmid itself has only one HindIII site. Therefore, it was expected to obtain 2 strict bands after the digestion of mutagenized plasmids with HindIII enzyme (Figure 4.13).
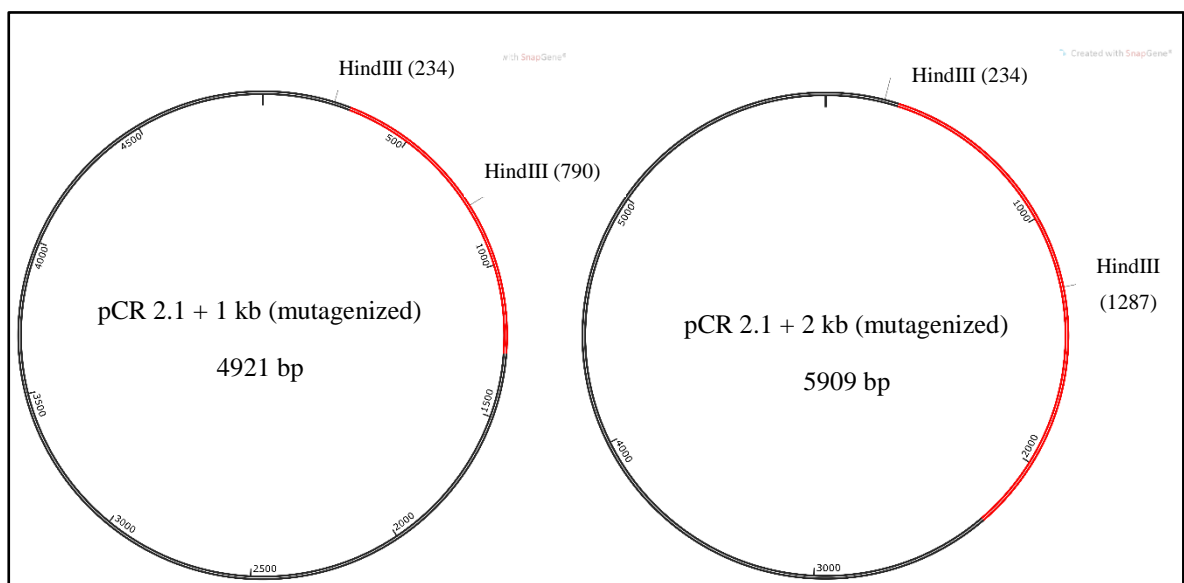
Figure 4.13. Plasmids maps of mutagenized pCR 2.1. Extra HindIII recognition sequences were introduced in the centre of inserts.

To confirm the sequence modification on HDR templates, restriction digestion with HindIII enzyme was performed. Digested products were run 1% (w/v) agarose gel (Figure 4.14).
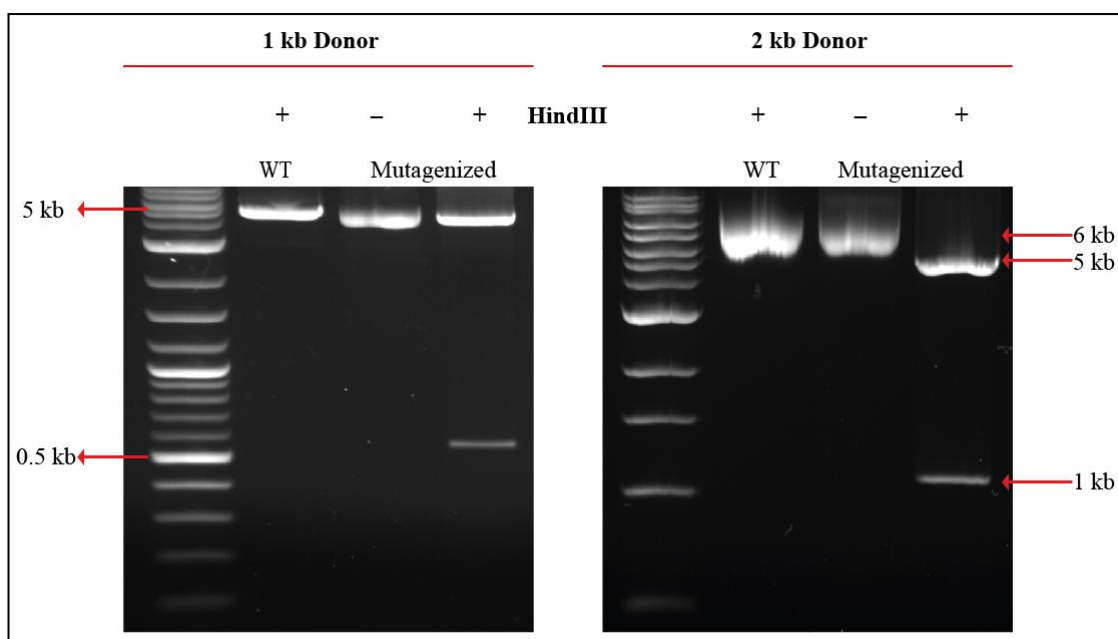
Figure 4.14. Site directed mutagenesis analysis

The agarose gel image indicates that the PAM sequence in wild type can was successfully replaced with HindIII recognition site, providing the addition of an extra restriction site downstream of the parental one. Result suggests that both plasmids can be used as donor templates for HDR mechanism to induce the desired modification on target locus of HBB gene.

4.7. CO-DELIVERY OF CAS9-SGRNA AND DONOR TEMPLATES

The final and most crucial step of this study was the delivery of Cas9, sgRNA and donor templates to induce site specific genomic modification. Cas9 and sgRNA was delivered in the form of plasmid. Based on the data presented at sections 4.2 and 4.3, it was decided to use guide-SCD-long for the gene targeting. Donor templates were delivered as circular plasmids carrying 1kb and 2 kb total homology lengths and the desired mutation for the screening of HDR events in target cells.

HEK293T cells were co-transfected with different amounts of Cas9/sgRNA and donor template plasmids. Transfection was done using 1:2 DNA to PEI protocol which was previously suggested as the most efficient method for this cell line. Additionally, cells were

treated with 1 µM SCR7 which is a small molecule reported as HDR enhancer for iPSCs. 24 hours after transfection, GFP expression of each sample well was assessed by flow cytometry (Figure 4.15).
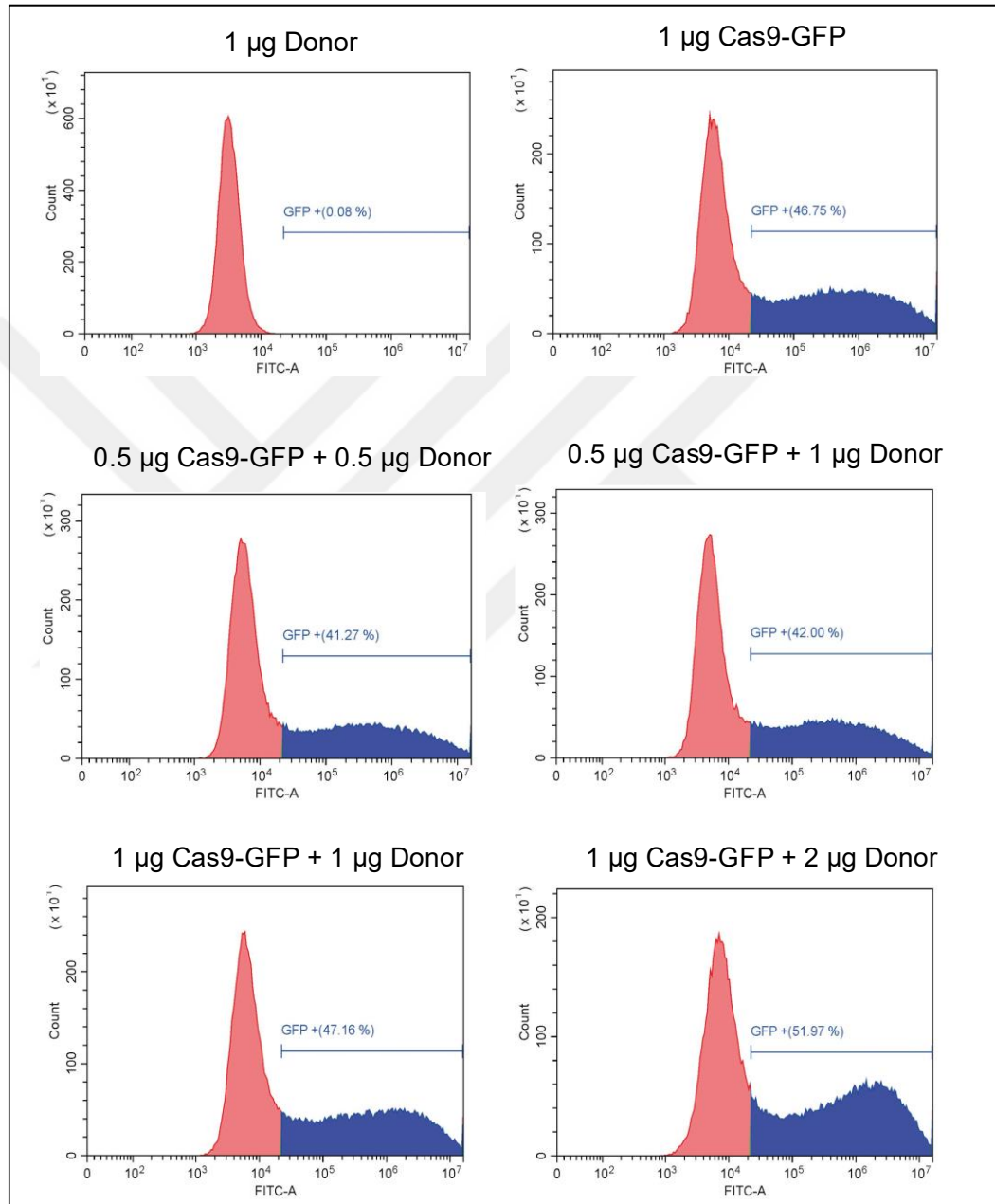
Figure 4.15. Flow cytometry analysis of GFP expression post-24 hours of transfection with pSpCas9(BB)-2A-GFP +guide-SCD-long and 1 kb donor template

Flow cytometry analysis of GFP expression for the cell transfected with pSpCas9(BB)-2A-GFP +guide-SCD-long and 1 kb donor template plasmids exhibited an efficient transfection yield after 24 hours. Increasing amount of pSpCas9(BB)-2A-GFP +guide-SCD-long resulted in increase in GFP positive population. Despite the increase in the percentage of GFP expressing cells, cytotoxicity was also increased due to the increased amount of PEI used for the transfection. Identical flow cytometry analysis was performed for HEK293T cells transfected with pSpCas9(BB)-2A-GFP +guide-SCD-long and 2 kb donor template plasmid (Figure 4.16).
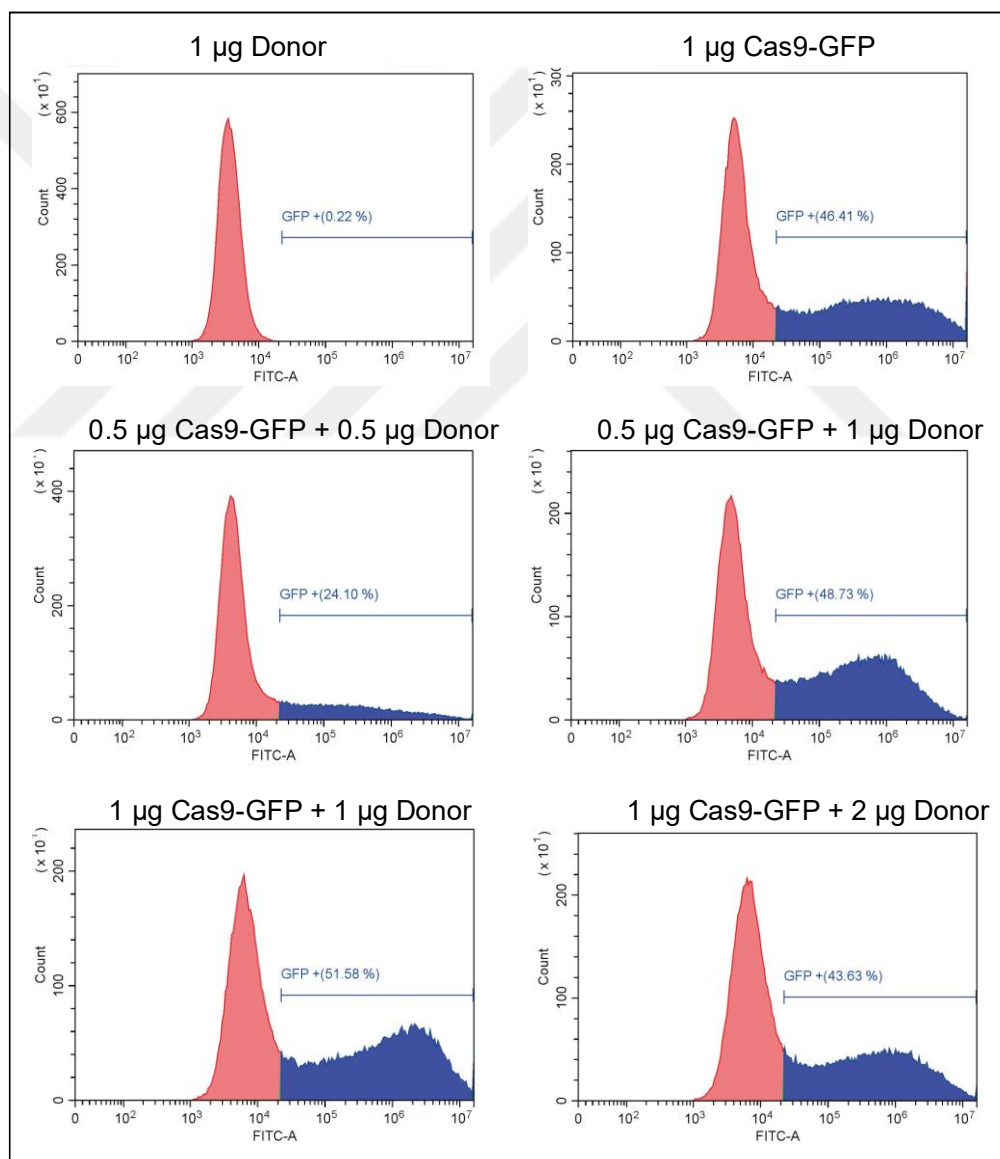
Figure 4.16. Flow cytometry analysis of GFP expression post-24 hours of transfection with pSpCas9(BB)-2A-GFP +guide-SCD-long and 2 kb donor template.

Flow cytometry results of the cells transfected with 2 kb donor templates showed a similar pattern of transfection efficiency and GFP expression. Similar to previous results, increased amount of pSpCas9(BB)-2A-GFP +guide-SCD-long resulted in higher percentage in GFP positive cell population, however an increase in cytotoxicity-based cell death was observed.

In conclusion, co-delivery of the components for site specific genome editing was achieved with the efficiency over 40%. Data suggests that the method of delivery for efficient genome editing is promising and suitable for HEK293T cells.

4.8. ON-TARGET GENE ANALYSIS OF HOMOLOGY DIRECTED REPAIR

The efficiency of gene editing on HBB gene in which the locus where point mutation causing SCD was examined 72 hours post-transfection and SCR7 treatment. Briefly genomic DNAs were isolated for each sample and the target locus was amplified centring the theoretical location of genomic modification. PCR amplification was followed by HindIII restriction enzyme digestion and the reaction mixture was run on agarose gel (Figure 4.17).
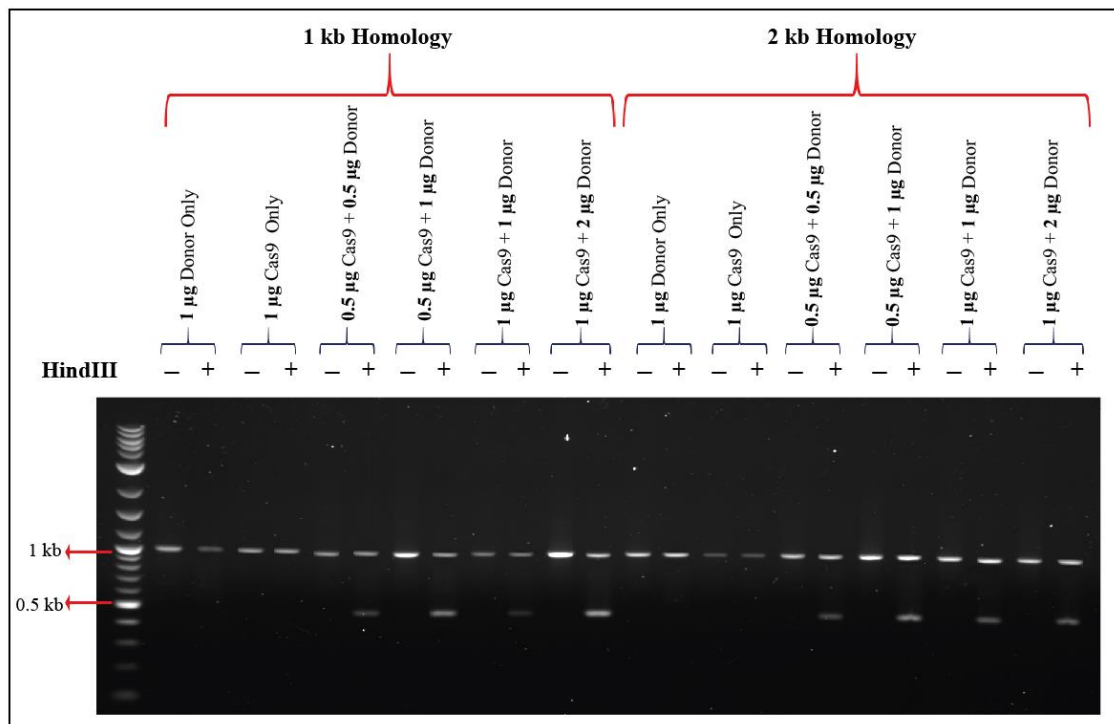
Figure 4.17. On-target analysis of HDR mediated gene editing on HBB gene

Data indicates that the desired genomic modification on HBB gene was achieved with varying efficiencies. According to on-target analysis, the required steps were successfully completed to fulfil the mission of sequence changes on HBB gene, with only a few base distances to the point mutation (rs334) causing SCD. The presence of 0.5 kb band on the gel is the indicator of the replacement of PAM sequence (GGG) with HindIII recognition sequence (AAGCTT) by the homology directed repair mechanism.

In this study, homology arms were designed symmetrically, 0.5 kb and 1 kb, 1kb and 2 kb in total. At the centre of each repair template derived from wilt type DNA sequence, the PAM sequence downstream to the sgRNA binding site was replaced with HindIII recognition site. Thus, the strategy has worked for both protection of donor DNA from sgRNA guided Cas9 activity and as the genomic marker for HDR event aimed to induce on target locus.

Different levels of efficiencies were observed on HDR analysis data. Using the data, it is possible to deduce which combination of Cas9/sgRNA plasmid, donor template plasmid and the transfection reagent PEI. In case of HDR with 1 kb total homology, 1:2 ratio of pSpCas9(BB)-2A-GFP +guide-SCD-long to the donor plasmid appeared to be the most efficient combination. Taking the cytotoxicity observed during previous steps into account, the mixture of 0.5 µg of Cas9 expressing plasmid and 1 µg of plasmid containing 1 kb homology would be the best choice for the target locus in HEK293T cells. Doubling the length of homology did not remarkably increased the efficiency of genome editing. For 2 kb homology scenario, 0.5 µg of Cas9 expressing plasmid and 1 µg of donor plasmid has shown the highest efficiency for HDR mediated modification. However, genome editing yield of 2 kb homology was not distinguishable from 1 kb total homology. Therefore, a homology length higher than 1 kb is not necessary to induce genomic modifications on HBB gene of HEK293T cells.

5.   DISCUSSION

The main objective of this study was to suggest an optimized protocol to induce direct genomic modifications on HBB gene. RNA mediated CRISPR/Cas9 technology was utilized to conduct sequence specific changes in mammalian cell lines, HEK293T and K562. Taking the advantage of HDR mechanism induced by DNA DSBs and homology donor DNA, it was achieved to generate planned base changes on the genome with high efficiencies and low cost.

CRISPR/Cas9 system has recently been a striking tool for site specific genome editing. RNA guided targeting nature of CRISPR/Cas9 system turns out to be the most efficient and target specific method for genome engineering in comparison to other endonuclease-based systems such as ZFNs and TALEN. The practical ease of design promised by CRISPR/Cas9 enables researchers to generate genomic modifications including knock-ins, knock-outs, translocations and base changes in a more robust and consistent fashion [205]. CRISPR/Cas9 is currently being widely used in tremendous number of molecular biology laboratories and it has improved the ability of generating disease models for drug targeting, cancer and stem cell researches [206].

The era of engineered nuclease-based genome editing has brought a new hope for inherited disorders. Using this technology, permanent cures can be provided to the patients. Sickle cell disease is one of the inherited disorders hoping for a cure from gene editing technology. Since it is caused by a point mutation on HBB gene, it is possible to efficiently induce the base change to restore functional beta subunit of haemoglobin. However, genome editing protocols still require improvements to reach 100% targeting efficiency and suggest lower cost, user friendly application and infinitesimal levels genotoxicity. In this study, it was aimed to target the locus of mutation causing SCD and introduce genomic changes using CRISPR/Cas9 induced DSB and HDR.

Sickle cell disease is caused by point mutation, rs334, causing amino acid valine to glutamic acid at sixth position of the beta globin subunit of haemoglobin [207]. To target the locus of mutation, 2 gRNAs were designed. The first gRNA named guide-SCD was standard 20 bp long sequence with 12 bp distance to mutation locus. To increase on-target efficiency and decrease off-target effects, a longer version of the first gRNA selected was designed. The

second gRNA named guide-SCD-long was 25 bp in size, obtained by extra 5 bases at the 5' end.

Initially, it was planned to use human hematopoietic cells collected from the donors suffering from SCD. Nevertheless, reaching a donor to use this model cells has not been possible during the time of this study. Thus, it was decided to use wild type human cell lines to test the ability to introduce base changes on the same locus. As the first step of the study, HEK293T cells were used for targeting since they are immortal, easy to handle and easy to transfect [208]. Non-viral delivery of Cas9 and gRNAs was attempted using 2 different transfection reagents, Lipofectamine 2000 and PEI. Lipofectamine 2000 is composed of cationic liposome that provides transfection in mammalian cells [209]. Polyethylenimine (PEI), is defined as a stable cationic polymer, condensing DNA into positively charged particles [210]. PEI is currently the cheapest product in market as the transfection reagent. However, PEI has higher toxicity over cells compare to Lipofectamine 2000 transfection. According to the results of this study, PEI demonstrated a better efficiency for transfection of HEK293T cells compare to Lipofectamine transfection method scaled to the same amount of plasmid used for transfection. Despite being more toxic, PEI was decided to be a better choice for transfecting HEK293T cell line. Furthermore, 2 different doses of PEI were tested in terms of toxicity and transfection efficiency. GFP expression analysis indicated that 1:2 DNA to PEI ratio has exhibited a better performance compare to 1:3 ratio. The percentage of GFP expressing cells were higher in 1:2 combination and the toxicity-based cell death was remarkably higher in 1:3 ratio. Average transfection efficiency was determined as 40%. However, K562 cell line did not show the same efficiency in terms of transfection with PEI, tested by GFP expression. The maximum yield of transfection was nearly 4%, applying the same protocol for K562 cell line. Since this level of transfection yield would not be suiTable for downstream applications of genome editing, it was decided to continue with HEK293T cells only. The results suggest that the transfection protocol for K562 must be further optimized or different transfection reagents must be tested to obtain higher yield.

The effect of sgRNA length was investigated with addition of 7 extra bases to the 5' of standard sgRNA designed for HBB gene targeting. Referring the on-target mutagenesis analysis using T7E assay, no significant difference was observed in terms of sgRNA activity on target locus. In previous studies, it was reported that longer sgRNA sequences would

decrease the off-target effects due to increased specificity of matching [211]. Despite off-target analysis was not performed in this study, it was determined to prefer longer gRNA, guide-SCD-long, over standard length.

In this study HDR donor templates were generated locally. Gene fragments with different lengths (1 kb, 2k and 4 kb) were attempted to be produced by traditional PCR method. Production of 4 kb fragment was failed due to the activity limitations of polymerase used in the lab. 1 kb and 2 fragments centring the target sequence was cloned into pCR 2.1 plasmids using TA cloning method, which is easy and convenient. In fusion cloning system was utilized to induce sequence changes, through a site directed mutagenesis approach, on wild type fragments cloned into plasmids. The PAM sequence downstream of gRNA binding site was replaced with HindIII recognition sequence for two reasons. The first reason was to avoid Cas9 cleavage activity on HDR templates since the endonuclease would not induce DSBs without the PAM sequence [212]. Besides protecting the HDR templates, HindIII sequence was used as a marker of gene editing on HEK293T cells. In case of successful gene editing via HDR mechanism, we would expect to observe cleavage on the PCR amplified target locus when treated with HindIII.

Following the generation of repair templates containing desired sequence modification, previously optimized protocol for transfection was applied for co-delivery of Cas9/sgRNA and donor templates for HDR. Different doses and combinations of Cas9/sgRNA and repair template plasmids were tested. Moreover, the effect of homolog length for HDR efficiency was tested using 1 kb and 2 kb total and symmetrical homologies. Initially it was planned to test different forms of homology templates such as circular plasmid, linearized plasmid and PCR product. It was required to efficiently generate linearized plasmids and PCR products by restriction digestion and gel purification protocols. Nevertheless, adequate amount of DNA and purities could not be obtained. Therefore, the repair templates were only used as circular plasmids.

To enhance the HDR efficiency, transfected cells were treated with SCR7, 12 hours after transfection. SCR7 is small molecule and inhibitor of DNA Ligase IV which has important role in NHEJ mechanism. Previous studies have reported that 1 µM SCR7 treatment to iPSCs resulted in significant increase in HDR mediated gene editing, inhibiting NHEJ mechanism [213]. After co-delivery of gene editing components into HEK293T cells, transfection

efficiencies were evaluated using flow cytometry. Transfection efficiencies were in a range between 40-50%. It was observed that increased total amount of plasmids resulted in increased cell death due to higher amounts of PEI used for transfection.

On-target analysis results demonstrated successful induction of base replacements on target locus. All combinations of Cas9/sgRNA and donor template plasmids have induced HDR based genome editing. As a direct deduction, the combinations in which the amount of donor template plasmid was the double of Cas9/sgRNA plasmid has shown better efficiency to generate the desired modification on HBB gene. There were not any significant differences between 1kb and 2 kb homology donors in terms of gene editing yield.

To conclude, the results of this study suggest that the protocol followed has been successful to generate base changes on HBB gene, 20-25 base pair distance to SCD causing mutation. The efficiencies achieved in this study may vary under different conditions. Type and passage number of the cells, target gene locus, and many other circumstances would influence HDR efficiency. Overall, the results of this study are only specific for HEK293T cells. The pipeline and protocols of this study can be modified for other cell lines and genes. Clinical translation of this approach requires more effort and further improvements.

6. FUTURE DIRECTIONS

Genome engineering has rapidly been developed and widely used in many diverse topics of molecular biology since its striking introduction in 2013. CRISPR/Cas9 has the highest capacity of gene manipulations in sequence specific fashion. In comparison to other endonuclease-based platforms, CRISPR/Cas9 is the most user friendly, robust and cost-efficient.

This study gives an insight that target specific sequence changes can be conducted in a basic molecular biology laboratory environment. Results and optimization steps are only specific for HEK293T cells and HBB gene, therefore these may not consistent and convenient for other cell types or genes. Further optimizations and off-target analysis should be done to improve the strategy to make it feasible for hard-to-transfect cells.

REFERENCES

1.   Auerbach C. Chemical mutagenesis. *Biological Reviews*. 1949;24(3):355-91.

2.   Muller HJ. Advances in radiation mutagenesis through studies on drosophila. *Progress in Nuclear Energy*. 1959;2:146-60.

3.   Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS. Insertion of DNA sequences into the human chromosomal β-globin locus by homologous recombination. *Nature*. 1985;317(6034):230-36.

4.   Thomas KR, Folger KR, Capecchi MR. High frequency targeting of genes to specific sites in the mammalian genome. *Cell*. 1986;44(3):419-28.

5.   Mansour SL, Thomas KR, Capecchi MR. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: A general strategy for targeting mutations to non-selectable genes. *Nature*. 1988;336(6197):348-57.

6.   Carroll D. Genome engineering with targetable nucleases. *Annual Review of Biochemistry*. 2014;83:409-39.

7.   Gaj T, Gersbach CA, Barbas III CF. Zfn, talen, and crispr/cas-based methods for genome engineering. *Trends in Biotechnology*. 2013;31(7):397-405.

8.   Carroll D. Genome engineering with zinc-finger nucleases. *Genetics*. 2011;188(4):773-82.

9.   Wyman C, Kanaar R. DNA double-strand break repair: All's well that ends well. *Annual Review Genetics*. 2006;40:363-83.

10.   Bhakta MS, Henry IM, Ousterout DG, Das KT, Lockwood SH, Meckler JF, et al. Highly active zinc-finger nucleases by extended modular assembly. *Genome Research*. 2013;23(3):530-8.

11.   Cathomen T, Keith Joung J. Zinc-finger nucleases: The next generation emerges. *Molecular Therapy*. 2008;16(7):1200-7.

12.   Kim S, Lee MJ, Kim H, Kang M, Kim J-S. Preassembled zinc-finger arrays for rapid construction of zfns. *Nature Methods*. 2011;8(1):7-18.

13.   Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, et al. High frequency modification of plant genes using engineered zinc finger nucleases. *Nature*. 2009;459(7245):442-5.

14.   Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to fok i cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*. 1996;93(3):1156-60.

15.   Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, et al. Efficient design and assembly of custom talen and other tal effector-based constructs for DNA targeting. *Nucleic Acids Research*. 2011;39(12):e82-e.

16.   Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. Flash assembly of talens enables high-throughput genome editing. *Nature Biotechnology*. 2012;30(5):460-5.

17.   Wood AJ, Lo T-W, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted genome editing across species using zfns and talens. *Science (New York, NY)*. 2011;333(6040):307-.

18.   Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, et al. Iterative capped assembly: Rapid and scalable synthesis of repeat-module DNA such as tal effectors from individual monomers. *Nucleic Acids Research*. 2012;40(15):e117-e.

19.   Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using crispr/cas systems. *Science (New York, NY)*. 2013;339(6121):819-23.

20.   Makarova S, Haft D, Barrangou R, Brouns S, Charpentier E, Horvath P, et al. Evolution and classification of the crispr-cas systems. *Nature Reviews Microbiology*. 2011;9(6):467-77.

21.   Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-rna–guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816.

22.   Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing talens with crisprs. *Cell Stem Cell*. 2013;12(4):393-4.

23.   Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9–crrna ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(39):E2579-E86.

24.   Zhang F, Wen Y, Guo X. Crispr/cas9 for genome editing: Progress, implications and challenges. *Human Molecular Genetics*. 2014;23(R1):R40-R6.

25.   Shah SA, Erdmann S, Mojica FJM, Garrett RA. Protospacer recognition motifs: Mixed identities and functional diversity. *RNA Biology*. 2013;10(5):891-9.

26.   Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic crispr defence system. *Microbiology*. 2009;155(3):733-40.

27.   Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the crispr rna-guided endonuclease cas9. *Nature*. 2014;507(7490):62-7.

28.   Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of pam-dependent target DNA recognition by the cas9 endonuclease. *Nature*. 2014;513(7519):569-73.

29.   Esvelt KM, Mali P, Braff JL, Moosburner M, Yaung SJ, Church GM. Orthogonal cas9 proteins for rna-guided gene regulation and editing. *Nature Methods*. 2013;10(11):1116-21.

30.   Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High frequency off-target mutagenesis induced by crispr-cas nucleases in human cells. *Nature Biotechnology*. 2013;31(9):822-6.

31.   Mali P, Aach J, Stranges PB, Esvelt KM, Moosburner M, Kosuri S, et al. Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*. 2013;31(9):833-8.

32.   Gratz SJ, Cummings AM, Nguyen JN, Hamm DC, Donohue LK, Harrison MM, et al. Genome engineering of drosophila with the crispr rna-guided cas9 nuclease. *Genetics*. 2013;194(4):1029-35.

33.   Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khayter C, Yeo DT, et al. In situ genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. *Stem Cells (Dayton, Ohio)*. 2011;29(11):1717-26.

34.   Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005;435(7042):646-51.

35.   Zou J, Mali P, Huang X, Dowey SN, Cheng L. Site-specific gene correction of a point mutation in human ips cells derived from an adult patient with sickle cell disease. *Blood*. 2011;118(17):4599-608.

36.   Torikai H, Reik A, Liu P-Q, Zhou Y, Zhang L, Maiti S, et al. A foundation for universal t-cell based immunotherapy: T cells engineered to express a cd19-specific chimeric-antigen-receptor and eliminate expression of endogenous tcr. *Blood*. 2012;119(24):5697-705.

37.   Voit RA, McMahon MA, Sawyer SL, Porteus MH. Generation of an hiv resistant t-cell line by targeted "stacking" of restriction factors. *Molecular Therapy*. 2013;21(4):786-95.

38.   Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, Hannenhalli S, et al. Genome-wide analysis of retroviral DNA integration. *Nat Rev Micro*. 2005;3(11):848-58.

39.   Ciccia A, Elledge SJ. The DNA damage response: Making it safe to play with knives. *Molecular Cell*. 2010;40(2):179-204.

40.   Skipper KA, Mikkelsen JG. Delivering the goods for genome engineering and editing. *Human Gene Therapy*. 2015;26(8):486-97.

41. Orlando SJ, Santiago Y, DeKelver RC, Freyvert Y, Boydston EA, Moehle EA, et al. Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Research*. 2010;38(15):e152-e.

42. Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssdna oligonucleotides with zinc-finger nucleases. *Nature Methods*. 2011;8(9):753-5.

43. Kim S, Kim D, Cho SW, Kim J, Kim J-S. Highly efficient rna-guided genome editing in human cells via delivery of purified cas9 ribonucleoproteins. *Genome Research*. 2014;24(6):1012-9.

44. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, et al. Rapid and highly efficient mammalian cell engineering via cas9 protein transfection. *Journal of Biotechnology*. 2015;208:44-53.

45. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of hiv-1 resistance in cd4(+) t cells by genome editing using zinc-finger nucleases. *Nature Biotechnology*. 2008;26(7):808-16.

46. Flotte TR. Size does matter: Overcoming the adeno-associated virus packaging limit. *Respiratory Research*. 2000;1(1):16-8.

47. Chira S, Jackson CS, Oprea I, Ozturk F, Pepper MS, Diaconu I, et al. Progresses towards safe and efficient gene therapy vectors. *Oncotarget*. 2015;6(31):30675.

48. Lombardo A, Genovese P, Beausejour C, Colleoni S, Lee Y-L, Kim K, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Blood Cells, Molecules, and Diseases*. 2008;40(2):278.

49. Naldini L. Ex vivo gene transfer and correction for cell-based therapies. *Nature Reviews Genetics*. 2011;12(5):301.

50. Rittié L, Perbal B. Enzymes used in molecular biology: A useful guide. *Journal of Cell Communication and Signaling*. 2008;2(1-2):25-45.

51.   Nishino T, Morikawa K. Structure and function of nucleases in DNA repair: Shape, grip and blade of the DNA scissors. *Oncogene*. 2002;21(58):9022.

52.   Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, et al. Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. *Current Gene Therapy*. 2011;11(1):11-27.

53.   Pingoud A, Silva GH. Precision genome surgery. *Nature Biotechnology*. 2007;25(7):743.

54.   San Filippo J, Sung P, Klein H. Mechanism of eukaryotic homologous recombination. *Annual Reviews Biochemistry*. 2008;77:229-57.

55.   Moore JK, Haber JE. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in saccharomyces cerevisiae. *Molecular and Cellular Biology*. 1996;16(5):2164-73.

56.   Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end joining pathway. *Annual Review of Biochemistry*. 2010;79:181-211.

57.   Blier PR, Griffith AJ, Craft J, Hardin JA. Binding of ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *Journal of Biological Chemistry*. 1993;268(10):7594-601.

58.   Lieber MR. The mechanism of human nonhomologous DNA end joining. *Journal of Biological Chemistry*. 2008;283(1):1-5.

59.   Ma Y, Lu H, Tippin B, Goodman MF, Shimazaki N, Koiwai O, et al. A biochemically defined system for mammalian nonhomologous DNA end joining. *Molecular Cell*. 2004;16(5):701-13.

60.   Yaneva M, Kowalewski T, Lieber MR. Interaction of DNA-dependent protein kinase with DNA and with ku: Biochemical and atomic-force microscopy studies. *The EMBO journal*. 1997;16(16):5098-112.

61.   Chen L, Trujillo K, Sung P, Tomkinson AE. Interactions of the DNA ligase iv/xrcc4 complex with DNA ends and the DNA-dependent protein kinase. *Journal of Biological Chemistry*. 2000;5(6):745-61.

62.   Ma Y, Schwarz K, Lieber MR. The artemis: DNA-pkcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair*. 2005;4(7):845-51.

63.   Gu J, Lu H, Tippin B, Shimazaki N, Goodman MF, Lieber MR. Xrcc4: DNA ligase iv can ligate incompatible DNA ends and can ligate across gaps. *The EMBO Journal*. 2007;26(4):1010-23.

64.   Kowalczykowski SC. An overview of the molecular mechanisms of recombinational DNA repair. *Cold Spring Harbor Perspectives in Biology*. 2015;7(11):a016410.

65.   Morrical SW. DNA-pairing and annealing processes in homologous recombination and homology-directed repair. *Cold Spring Harbor Perspectives in Biology*. 2015;7(2):a016444.

66.   Chen R, Wold MS. Replication protein a: Single-stranded dna's first responder: Dynamic DNA-interactions allow replication protein a to direct single-strand DNA intermediates into different pathways for synthesis or repair. *Bioessays*. 2014;36(12):1156-61.

67.   Heyer WD, Ehmsen KT, Liu J. Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics*. 2010;44:113-39.

68.   Schleifman EB, Chin JY, Glazer PM. Triplex-mediated gene modification. *Chromosomal Mutagenesis*: Springer; 2008. p. 175-90.

69.   Majumdar A, Muniandy PA, Liu J, Liu J-l, Liu S-t, Cuenoud B, et al. Targeted gene knock in and sequence modulation mediated by a psoralen-linked triplex-forming oligonucleotide. *Journal of Biological Chemistry*. 2008;283(17):11244-52.

70.   Arimondo PB, Thomas CJ, Oussedik K, Baldeyrou B, Mahieu C, Halby L, et al. Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates. *Molecular and Cellular Biology*. 2006;26(1):324-33.

71.   Simon P, Cannata F, Perrouault L, Halby L, Concordet J-P, Boutorine A, et al. Sequence-specific DNA cleavage mediated by bipyridine polyamide conjugates. *Nucleic Acids Research*. 2008;36(11):3531-8.

72.   Eisenschmidt K, Lanio T, Simoncsits A, Jeltsch A, Pingoud V, Wende W, et al. Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Research*. 2005;33(22):7039-47.

73.   Orlowski J, Boniecki M, Bujnicki JM. I-ssp6803i: The first homing endonuclease from the pd-(d/e) xk superfamily exhibits an unusual mode of DNA recognition. *Bioinformatics*. 2007;23(5):527-30.

74.   Zhao L, Bonocora RP, Shub DA, Stoddard BL. The restriction fold turns to the dark side: A bacterial homing endonuclease with a pd-(d/e)-xk motif. *The EMBO Journal*. 2007;26(9):2432-42.

75.   Grindl W, Wende W, Pingoud V, Pingoud A. The protein splicing domain of the homing endonuclease pi-sce i is responsible for specific DNA binding. *Nucleic Acids Research*. 1998;26(8):1857-62.

76.   Moure CM, Gimble FS, Quiocho FA. Crystal structure of the intein homing endonuclease pi-scei bound to its recognition sequence. *Nature Structural and Molecular Biology*. 2002;9(10):764.

77.   Jacquier A, Dujon B. An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell*. 1985;41(2):383-94.

78.   Choulika A, Perrin A, Dujon B, Nicolas J-F. Induction of homologous recombination in mammalian chromosomes by using the i-scei system of saccharomyces cerevisiae. *Molecular and Cellular Biology*. 1995;15(4):1968-73.

79.   Fenina M, Simon-Chazottes D, Vandormael-Pournin S, Soueid J, Langa F, Cohen-Tannoudji M, et al. I-scei-mediated double-strand break does not increase the frequency of homologous recombination at the dct locus in mouse embryonic stem cells. *PloS One*. 2012;7(6):e39895.

80.   Wang Y, Zhou X-Y, Xiang P-Y, Wang L-L, Tang H, Xie F, et al. The meganuclease i-scei containing nuclear localization signal (nls-i-scei) efficiently mediated mammalian germline transgenesis via embryo cytoplasmic microinjection. *PloS One*. 2014;9(9):e108347.

81.   Grizot S, Smith J, Daboussi F, Prieto J, Redondo P, Merino N, et al. Efficient targeting of a scid gene by an engineered single-chain homing endonuclease. *Nucleic Acids Research*. 2009;37(16):5405-19.

82.   Ménoret S, Fontanière S, Jantz D, Tesson L, Thinard R, Rémy S, et al. Generation of rag1-knockout immunodeficient rats and mice using engineered meganucleases. *The FASEB Journal*. 2013;27(2):703-11.

83.   Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim Y-G, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Molecular and Cellular Biology*. 2001;21(1):289-97.

84.   Cathomen T, Joung JK. Zinc-finger nucleases: The next generation emerges. *Molecular Therapy*. 2008;16(7):1200-7.

85.   Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I. Foki dimerization is required for DNA cleavage. *Proceedings of the National Academy of Sciences*. 1998;95(18):10570-5.

86.   Wolfe SA, Nekludova L, Pabo CO. DNA recognition by cys2his2 zinc finger proteins. *Annual Review of Biophysics and Biomolecular Structure*. 2000;29(1):183-212.

87.   Tupler R, Perini G, Green MR. Expressing the human genome. *Nature*. 2001;409(6822):832.

88.   Pavletich NP, Pabo CO. Zinc finger-DNA recognition: Crystal structure of a zif268-DNA complex at 2.1 a. *Science*. 1991;252(5007):809-17.

89.   Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Molecular Therapy*. 2008;16(2):352-8.

90. Kim HJ, Lee HJ, Kim H, Cho SW, Kim J-S. Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Research*. 2009;252(503):609-17.

91. Geurts AM, Cost GJ, Freyvert Y, Zeitler B, Miller JC, Choi VM, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*. 2009;325(5939):433-.

92. Urnov FD, Miller JC, Lee Y-L, Beausejour CM, Rock JM, Augustus S, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005;435(7042):646.

93. Carbery ID, Ji D, Harrington A, Brown V, Weinstein EJ, Liaw L, et al. Targeted genome modification in mice using zinc-finger nucleases. *Genetics*. 2010;186(2):451-9.

94. Buehr M, Meek S, Blair K, Yang J, Ure J, Silva J, et al. Capture of authentic embryonic stem cells from rat blastocysts. *Cell*. 2008;135(7):1287-98.

95. Hauschild J, Petersen B, Santiago Y, Queisser A-L, Carnwath JW, Lucas-Hahn A, et al. Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proceedings of the National Academy of Sciences*. 2011;108(29):12013-7.

96. Yu S, Luo J, Song Z, Ding F, Dai Y, Li N. Highly efficient modification of beta-lactoglobulin (blg) gene via zinc-finger nucleases in cattle. *Cell Research*. 2011;21(11):1638-47.

97. Radecke S, Radecke F, Cathomen T, Schwarz K. Zinc-finger nuclease-induced gene repair with oligodeoxynucleotides: Wanted and unwanted target locus modifications. *Molecular Therapy*. 2010;18(4):743-53.

98. Maier DA, Brennan AL, Jiang S, Binder-Scholl GK, Lee G, Plesa G, et al. Efficient clinical scale gene modification via zinc finger nuclease–targeted disruption of the hiv co-receptor ccr5. *Human Gene Therapy*. 2013;24(3):245-58.

99. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, et al. Gene editing of ccr5 in autologous cd4 t cells of persons infected with hiv. *New England Journal of Medicine*. 2014;370(10):901-10.

100. Swarthout JT, Raisinghani M, Cui X. Zinc finger nucleases: A new era for transgenic animals. *Annals of Neurosciences*. 2011;18(1):25.

101. Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S, et al. Knockout rats generated by embryo microinjection of talens. *Nature Biotechnology*. 2011;29(8):695.

102. Sung YH, Baek I-J, Kim DH, Jeon J, Lee J, Lee K, et al. Knockout mice created by talen-mediated gene targeting. *Nature Biotechnology*. 2013;31(1):23.

103. Panda SK, Wefers B, Ortiz O, Floss T, Schmid B, Haass C, et al. Highly efficient targeted mutagenesis in mice using talens. *Genetics*. 2013;31(9):437.

104. Wefers B, Panda SK, Ortiz O, Brandl C, Hensler S, Hansen J, et al. Generation of targeted mouse mutants by embryo microinjection of talen mrna. *Nature Protocols*. 2013;8(12):2355.

105. Xin J, Yang H, Fan N, Zhao B, Ouyang Z, Liu Z, et al. Highly efficient generation of ggta1 biallelic knockout inbred mini-pigs with talens. *PloS One*. 2013;8(12):e84250.

106. Wei J, Wagner S, Lu D, Maclean P, Carlson DF, Fahrenkrug SC, et al. Efficient introgression of allelic variants by embryo-mediated editing of the bovine genome. *Scientific Reports*. 2015;5:11735.

107. Liu H, Chen Y, Niu Y, Zhang K, Kang Y, Ge W, et al. Talen-mediated gene mutagenesis in rhesus and cynomolgus monkeys. *Cell Stem Cell*. 2014;14(3):323-8.

108. Li L, Atef A, Piatek A, Ali Z, Piatek M, Aouida M, et al. Characterization and DNA-binding specificities of ralstonia tal-like effectors. *Molecular Plant*. 2013;6(4):1318-30.

109. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu J-K, et al. Structural basis for sequence-specific recognition of DNA by tal effectors. *Science*. 2012;335(6069):720-3.

110. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of tal-type iii effectors. *Science*. 2009;326(5959):1509-12.

111. Kim H, Kim J-S. A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics*. 2014;15(5):321-34.

112.   Periwal V. A comprehensive overview of computational resources to aid in precision genome editing with engineered nucleases. *Briefings in Bioinformatics*. 2016;18(4):698-711.

113.   Cermak T, Starker CG, Voytas DF. Efficient design and assembly of custom talens using the golden gate platform. *Chromosomal Mutagenesis*: Springer; 2015. p. 133-59.

114.   Sakuma T, Yamamoto T. Engineering customized talens using the platinum gate talen kit. *Talens*: Springer; 2016. p. 61-70.

115.   Mak AN-S, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of tal effector pthxo1 bound to its DNA target. *Science*. 2012;12(1):6211.

116.   Kim Y, Kweon J, Kim A, Chon JK, Yoo JY, Kim HJ, et al. A library of tal effector nucleases spanning the human genome. *Nature Biotechnology*. 2013;31(3):251.

117.   Mattapallil MJ, Wawrousek EF, Chan C-C, Zhao H, Roychoudhury J, Ferguson TA, et al. The rd8 mutation of the crb1 gene is present in vendor lines of c57bl/6n mice and embryonic stem cells, and confounds ocular induced mutant phenotypes. *Investigative Ophthalmology & Visual Science*. 2012;53(6):2921-7.

118.   Low BE, Krebs MP, Joung JK, Tsai SQ, Nishina PM, Wiles MV. Correction of the crb1rd8 allele and retinal phenotype in c57bl/6n mice via talen-mediated homology-directed repair. *Investigative Ophthalmology & Visual Science*. 2014;55(1):387-95.

119.   Osborn MJ, Starker CG, McElroy AN, Webber BR, Riddle MJ, Xia L, et al. Talen-based gene correction for epidermolysis bullosa. *Molecular Therapy*. 2013;21(6):1151-9.

120.   Biffi A. Clinical translation of talens: Treating scid-x1 by gene editing in ipscs. *Cell Stem Cell*. 2015;16(4):348-9.

121.   Garate Z, Quintana-Bustamante O, Crane AM, Olivier E, Poirot L, Galetto R, et al. Generation of a high number of healthy erythroid cells from gene-edited pyruvate kinase deficiency patient-specific induced pluripotent stem cells. *Stem Cell Reports*. 2015;5(6):1053-66.

122.   Mojica FJ, Montoliu L. On the origin of crispr-cas technology: From prokaryotes to mammals. *Trends in Microbiology*. 2016;24(10):811-20.

123.   Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using crispr/cas systems. *Science*. 2013:1231143.

124.   Wang H, Yang H, Shivalila CS, Dawlaty MM, Cheng AW, Zhang F, et al. One-step generation of mice carrying mutations in multiple genes by crispr/cas-mediated genome engineering. *Cell*. 2013;153(4):910-8.

125.   Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, et al. An updated evolutionary classification of crispr–cas systems. *Nature Reviews Microbiology*. 2015;13(11):722.

126.   Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. Crispr provides acquired resistance against viruses in prokaryotes. *Science*. 2007;315(5819):1709-12.

127.   Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, et al. Crispr rna maturation by trans-encoded small rna and host factor rnase iii. *Nature*. 2011;471(7340):602.

128.   Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of rna-guided cas9 nucleases. *Nature Biotechnology*. 2013;31(9):827.

129.   Shah SA, Erdmann S, Mojica FJ, Garrett RA. Protospacer recognition motifs: Mixed identities and functional diversity. *RNA biology*. 2013;10(5):891-9.

130.   Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lecrivain A-L, Bzdrenga J, et al. Phylogeny of cas9 determines functional exchangeability of dual-rna and cas9 among orthologous type ii crispr-cas systems. *Nucleic Acids Research*. 2013;42(4):2577-90.

131.   Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-rna–guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012:1225829.

132.   Grissa I, Vergnaud G, Pourcel C. Crisprfinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*. 2007;35(suppl_2):W52-W7.

133.   Rath D, Amlinger L, Rath A, Lundgren M. The crispr-cas immune system: Biology, mechanisms and applications. *Biochimie*. 2015;117:119-28.

134.   Mojica FJ, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of Molecular Evolution*. 2005;60(2):174-82.

135.   Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (crisprs) have spacers of extrachromosomal origin. *Microbiology*. 2005;151(8):2551-61.

136.   Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, et al. Small crispr rnas guide antiviral defense in prokaryotes. *Science*. 2008;321(5891):960-4.

137.   Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. A DNA repair system specific for thermophilic archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research*. 2002;30(2):482-96.

138.   Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P, et al. Evolution and classification of the crispr-cas systems. *Nature Review Microbiology*. 2011;9(6):467-77.

139.   Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and atp-dependent helicase in the crispr/cas immune system. *The EMBO Journal*. 2011;30(7):1335-42.

140.   Staals RH, Agari Y, Maki-Yonekura S, Zhu Y, Taylor DW, Van Duijn E, et al. Structure and activity of the rna-targeting type iii-b crispr-cas complex of thermus thermophilus. *Molecular Cell*. 2013;52(1):135-45.

141.   Seruggia D, Montoliu L. The new crispr–cas system: Rna-guided genome engineering to efficiently produce any desired genetic alteration in animals. *Transgenic Research*. 2014;23(5):707-16.

142.   Harms DW, Quadros RM, Seruggia D, Ohtsuka M, Takahashi G, Montoliu L, et al. Mouse genome editing using the crispr/cas system. *Current Protocols in Human Genetics*. 2014;83(1):15-7.

143.   Han Y, Slivano OJ, Christie CK, Cheng AW, Miano JM. Crispr-cas9 genome editing of a single regulatory element nearly abolishes target gene expression in mice. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2014;114(3):517.

144.   Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by crispr/cas-mediated genome engineering. *Cell*. 2013;154(6):1370-9.

145.   Platt RJ, Chen S, Zhou Y, Yim MJ, Swiech L, Kempton HR, et al. Crispr-cas9 knockin mice for genome editing and cancer modeling. *Cell*. 2014;159(2):440-55.

146.   Peng J, Wang Y, Jiang J, Zhou X, Song L, Wang L, et al. Production of human albumin in pigs through crispr/cas9-mediated knockin of human cdna into swine albumin locus in the zygotes. *Scientific Reports*. 2015;5:16705.

147.   Seruggia D, Fernández A, Cantero M, Pelczar P, Montoliu L. Functional validation of mouse tyrosinase non-coding regulatory DNA elements by crispr–cas9-mediated mutagenesis. *Nucleic Acids Research*. 2015;43(10):4855-67.

148.   Birling M-C, Schaeffer L, André P, Lindner L, Maréchal D, Ayadi A, et al. Efficient and rapid generation of large genomic variants in rats and mice using crismere. *Scientific Reports*. 2017;7:43331.

149.   Zhang L, Jia R, Palange NJ, Satheka AC, Togo J, An Y, et al. Large genomic fragment deletions and insertions in mouse using crispr/cas9. *PloS One*. 2015;10(3):e0120396.

150.   Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han Y-C, et al. In vivo engineering of oncogenic chromosomal rearrangements with the crispr/cas9 system. *Nature*. 2014;516(7531):423.

151. Flynn R, Grundmann A, Renz P, Hänseler W, James WS, Cowley SA, et al. Crispr-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human ips cells. *Experimental Hematology*. 2015;43(10):838-48. e3.

152. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via crispr/cas9 mediated homology-independent targeted integration. *Nature*. 2016;540(7631):144.

153. Nelson CE, Hakim CH, Ousterout DG, Thakore PI, Moreb EA, Rivera RMC, et al. In vivo genome editing improves muscle function in a mouse model of duchenne muscular dystrophy. *Science*. 2016;351(6271):403-7.

154. Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z, et al. Crispr/cas9-mediated gene editing in human tripronuclear zygotes. *Protein & Cell*. 2015;6(5):363-72.

155. Kang X, He W, Huang Y, Yu Q, Chen Y, Gao X, et al. Introducing precise genetic modifications into human 3pn embryos by crispr/cas-mediated genome editing. *Journal of Assisted Reproduction And Genetics*. 2016;33(5):581-8.

156. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, et al. Rna-guided human genome engineering via cas9. *Science*. 2013;339(6121):823-6.

157. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the crispr-cas9 system. *Nature Protocols*. 2013;8(11):2281.

158. Wang H, La Russa M, Qi LS. Crispr/cas9 in genome editing and beyond. *Annual Review of Biochemistry*. 2016;85:227-64.

159. Zhang Z, Zhang Y, Gao F, Han S, Cheah KS, Tse H-F, et al. Crispr/cas9 genome-editing system in human stem cells: Current status and future prospects. *Molecular Therapy-Nucleic Acids*. 2017;252(5007):809-17.

160. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, et al. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nature Methods*. 2008;5(5):374-5.

161. Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. Flash assembly of talens for high-throughput genome editing. *Nature Biotechnology*. 2012;30(5):460-5.

162.   Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. Rna-programmed genome editing in human cells. *Elife*. 2013;2:e00471.

163.   Pattanayak V, Ramirez CL, Joung JK, Liu DR. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nature Methods*. 2011;8(9):765-70.

164.   Mussolino C, Morbitzer R, Lutge F, Dannemann N, Lahaye T, Cathomen T. A novel tale nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Research*. 2011;39(21):9283-93.

165.   Kim YK, Wee G, Park J, Kim J, Baek D, Kim JS, et al. Talen-based knockout library for human micrornas. *Nature Structural Molecular Biology*. 2013;20(12):1458-64.

166.   Cho SW, Kim S, Kim JM, Kim J-S. Targeted genome engineering in human cells with the cas9 rna-guided endonuclease. *Nature Biotechnology*. 2013;31:230.

167.   Cradick TJ, Fine EJ, Antico CJ, Bao G. Crispr/cas9 systems targeting beta-globin and ccr5 genes have substantial off-target activity. *Nucleic Acids Research*. 2013;41(20):9584-92.

168.   Pattanayak V, Lin S, Guilinger JP, Ma E, Doudna JA, Liu DR. High-throughput profiling of off-target DNA cleavage reveals rna-programmed cas9 nuclease specificity. *Nature Biotechnology*. 2013;31(9):839-43.

169.   Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of crispr/cas-derived rna-guided endonucleases and nickases. *Genome Research*. 2014;24(1):132-41.

170.   Şöllü C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK, et al. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Research*. 2010;38(22):8269-76.

171.   Li W, Teng F, Li T, Zhou Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using crispr-cas systems. *Nature Biotechnology*. 2013;31(8):684-6.

172.   Niu Y, Shen B, Cui Y, Chen Y, Wang J, Wang L, et al. Generation of gene-modified cynomolgus monkey via cas9/rna-mediated gene targeting in one-cell embryos. *Cell*. 2014;156(4):836-43.

173.   Kim Y, Kweon J, Kim JS. Talens and zfns are associated with different mutation signatures. *Nature Methods*. 2013;10(3):185.

174.   Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crrna ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Science of the USA*. 2012;109(39):E2579-86.

175.   Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science*. 2003;300(5620):763.

176.   Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a crispr-cas system. *Nature Biotechnology*. 2013;31(3):227-9.

177.   Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nature Biotechnology*. 2008;26(6):695-701.

178.   Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotechnology*. 2007;25(11):1298-306.

179.   Holkers M, Maggio I, Liu J, Janssen JM, Miselli F, Mussolino C, et al. Differential integrity of tale nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Research*. 2013;41(5):e63.

180.   Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, et al. Establishment of hiv-1 resistance in cd4+ t cells by genome editing using zinc-finger nucleases. *Nature Biotechnology*. 2008;26(7):808-16.

181.   Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the crispr-cas9 system. *Science*. 2014;343(6166):80-4.

182. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF, 3rd. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nature Methods*. 2012;9(8):805-7.

183. Chen Z, Jaafar L, Agyekum DG, Xiao H, Wade MF, Kumaran RI, et al. Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Research*. 2013;41(19):e182.

184. Cho SW, Lee J, Carroll D, Kim J-S, Lee J. Heritable gene knockout in caenorhabditis elegans by direct injection of cas9–sgrna ribonucleoproteins. *Genetics*. 2013;195(3):1177-80.

185. Sung YH, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, et al. Highly efficient gene knockout in mice and zebrafish with rna-guided endonucleases. *Genome Research*. 2014;24(1):125-31.

186. Thom CS, Dickson CF, Gell DA, Weiss MJ. Hemoglobin variants: Biochemical properties and clinical correlates. *Cold Spring Harbour Perspective Medicine*. 2013;3(3):a011858.

187. Berg JM, Tymoczko JL, Stryer L. *Biochemistry*. Basingstoke: W.H. Freeman; 2012.

188. Frenette PS, Atweh GF. Sickle cell disease: Old discoveries, new concepts, and future promise. *Journal of Clinical Investigation*. 2007;117(4):850-8.

189. Ashley-Koch A, Yang Q, Olney RS. Sickle hemoglobin (hb s) allele and sickle cell disease: A huge review. *American Journal of Epidemiology*. 2000;151(9):839-45.

190. Azar S, Wong TE. Sickle cell disease. *Medical Clinics*. 2017;101(2):375-93.

191. Hassell KL. Population estimates of sickle cell disease in the u.S. *American Journal of Preventive Medicine*.38(4):S512-S21.

192. Piel FB. The present and future global burden of the inherited disorders of hemoglobin. *Hematology/Oncology Clinics of North America*. 2016;30(2):327-41.

193. Aliyu ZY, Tumblin AR, Kato GJ. Current therapy of sickle cell disease. *Haematologica*. 2006;91(1):7.

194.   Shenoy S. Hematopoietic stem cell transplantation for sickle cell disease: Current practice and emerging trends. *ASH Education Program Book*. 2011;2011(1):273-9.

195.   Locatelli F, Pagliara D. Allogeneic hematopoietic stem cell transplantation in children with sickle cell disease. *Pediatric Blood & Cancer*. 2012;59(2):372-6.

196.   Abil Z, Xiong X, Zhao H. Synthetic biology for therapeutic applications. *Molecular Pharmaceutics*. 2014;12(2):322-31.

197.   Hoban MD, Cost GJ, Mendel MC, Romero Z, Kaufman ML, Joglekar AV, et al. Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood*. 2015;125(17):2597-604.

198.   Sun N, Liang J, Abil Z, Zhao H. Optimized tal effector nucleases (talens) for use in treatment of sickle cell disease. *Molecular BioSystems*. 2012;8(4):1255-63.

199.   Sun N, Zhao H. Seamless correction of the sickle cell disease mutation of the hbb gene in human induced pluripotent stem cells using talens. *Biotechnology and Bioengineering*. 2014;111(5):1048-53.

200.   Ramalingam S, Annaluru N, Kandavelou K, Chandrasegaran S. Talen-mediated generation and genetic correction of disease-specific human induced pluripotent stem cells. *Current Gene Therapy*. 2014;14(6):461-72.

201.   Wright AV, Nuñez JK, Doudna JA. Biology and applications of crispr systems: Harnessing nature's toolbox for genome engineering. *Cell*. 2016;164(1):29-44.

202.   Huang X, Wang Y, Yan W, Smith C, Ye Z, Wang J, et al. Production of gene-corrected adult beta globin protein in human erythrocytes differentiated from patient ipscs after genome editing of the sickle point mutation. *Stem Cells*. 2015;33(5):1470-9.

203.   Cottle RN, Lee CM, Archer D, Bao G. Controlled delivery of β-globin-targeting talens and crispr/cas9 into mammalian cells for genome editing using microinjection. *Scientific Reports*. 2015;5(1):182-191.

204.   DeWitt M, Magis W, Bray NL, Wang T, Berman JR, Urbinati F, et al. Efficient correction of the sickle mutation in human hematopoietic stem cells using a cas9 ribonucleoprotein complex. *BioRxiv*. 2016:036236.

205. M. T-SD, Te-Wen L. Molecular biology at the cutting edge: A review on crispr/cas9 gene editing for undergraduates. *Biochemistry and Molecular Biology Education*. 2018;46(2):195-205.

206. Musunuru K. The hope and hype of crispr-cas9 genome editing: A review. *JAMA Cardiol*. 2017;2(8):914-9.

207. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. *Lancet*. 2010;376(9757):2018-31.

208. Thomas P, Smart TG. Hek293 cell line: A vehicle for the expression of recombinant proteins. *Journal of Pharmacol Toxicol Methods*. 2005;51(3):187-200.

209. Dalby B, Cates S, Harris A, Ohki EC, Tilkins ML, Price PJ, et al. Advanced transfection with lipofectamine 2000 reagent: Primary neurons, sirna, and high-throughput applications. *Methods*. 2004;33(2):95-103.

210. Boussif O, Lezoualc'h F, Zanta MA, Mergny MD, Scherman D, Demeneix B, et al. A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: Polyethylenimine. *Proceedings of the National Academy of Science USA*. 1995;92(16):7297-301.

211. Zhang JP, Li XL, Neises A, Chen W, Hu LP, Ji GZ, et al. Different effects of sgrna length on crispr-mediated gene knockout efficiency. *Science Reports*. 2016;6:285-96.

212. Zhang F, Wen Y, Guo X. Crispr/cas9 for genome editing: Progress, implications and challenges. *Human Molecular Genetics*. 2014;23(R1):R40-6.

213. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for crispr-cas9-induced precise gene editing in mammalian cells. *Nature Biotechnology*. 2015;33(5):543-8.