NUCLEASE MEDIATED GENE TARGETING OF THE SICKLE CELL DISEASE MUTATION UTILIZING NON-VIRAL DELIVERY STRATEGIES

A Dissertation
Presented to
The Academic Faculty

By

Renée N. Cottle

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Biomedical Engineering

Georgia Institute of Technology

December 2015

NUCLEASE MEDIATED GENE TARGETING OF THE SICKLE CELL DISEASE MUTATION UTILIZING NON-VIRAL DELIVERY STRATEGIES

Approved by:

Dr. Gang Bao, Advisor
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University
&
Department of Bioengineering
Rice University

Dr. David Archer
Department of Pediatrics
Emory School of Medicine

Dr. Kevin Bunting
Department of Pediatrics
Emory School of Medicine

Dr. Wilbur Lam
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Dr. Manu Platt
Department of Biomedical Engineering
Georgia Institute of Technology and Emory University

Date Approved: August 04, 2015

To Brandon Johnson; Faith and Desmond Burnett;

and Ioline and Herman Broomfield

ACKNOWLEDGEMENTS

Firstly, I would like to thank my Lord and Savior Jesus for opening the doors of opportunity to pursue my education and endowing me with the fortitude and intelligence to accomplish my academic goals. I am truly amazed by God's unwavering grace and look forward to the wonderful things that are to come in my life. I want to thank my parents, Faith and Desmond Burnett, for their ardent prayers, strong support, and for never letting me give up on my dreams. My parent's exemplary example of hard work, determination, and dedication to excellence was the inspiration that never fails to motivate me. I want to thank my husband, Brandon Johnson, for being a strong rock of support, my cheer leader and my best friend. Brandon is the best thing that happened to me, and I very grateful for the opportunity to share my life with him. I also want to thank my brothers and sister, Elijah, Jordan and Danielle, for always encouraging me. To the rest of my family and friends, thank you from the bottom of my heart for your support and prayers.

I would like to thank my advisor, Dr. Gang Bao, for providing me with the opportunity to train in his laboratory and advocating for my acceptance into the PhD program. I could not have made it through Georgia Tech without Dr. Bao's strong mentorship and support. I thank the members of the Bao Lab, particularly Dr. Ciaran Lee and Harsha Deshmukh, for helping me with my project and preparing me for a high standard of research. I also want to thank my thesis committee members: Drs. Manu Platt, David Archer, Kevin Bunting, and Wilbur Lam for their guidance and collaborative research efforts on my project. I wish to thank my collaborators: Reginald Tran and Dr. David Myers from the Lam Lab for making the microfluidic cell capturing devices that I tested while optimizing the microinjection methodology used for my project; Aaron Rae

from the Flow Cytometry Facility in Emory University and Children's Pediatric Research Center for helping me with FACS even during the weekends; Prithiviraj Jothkiumar from the Zhu Lab for helping me to prepare glass microcapillaries for microinjection; Seonhee Park from the Prausnitz Lab for helpful discussions about microinjection, Dr. Matthew Porteus for providing the plasmid constructs encoding for TALENs and donor templates used in my study; and Dr. Charles Gersbach for generously providing purified Cas9 endonuclease proteins.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

BM	Bone marrow
BSA	Bovine serum albumin
BCNU	1,3 bis(2-choloroethyl)-1-nitrosourea
BG	O6 benzylguanine
Cas9	CRISPR-associated protein 9
Cas9n	Cas9 nickase
CB	Umbilical cord blood
CFC	colony forming cell assay
CRISPR	Clustered regularly interspaced short palindromic repeats
DSBs	Double stranded breaks
ESCs	Embryonic stem cells
FACS	Fluorescence activated cell sorting
G-CSF	Granulocyte-colony stimulating factor
GRIN3A	Glutamate receptor
HbA	Adult hemoglobin
HBB	β-globin
HBD	δ-globin
HbF	Fetal hemoglobin
HBG1	γ-globin
HbS	Sickle hemoglobin
HDR	Homology directed repair

HIV	Human immunodeficiency virus
HSCs	Hematopoietic stem cells
HSCT	Hematopoietic stem cell transplantation
IDLV	Integrase defective lentivirus
indels	Insertions and deletions
iPSCs	Induced pluripotent stem cells
NHEJ	Non-homologous end joining
NOD	Non-obese diabetic
P140K-MGMT	P140K mutant O6-methylguanine methyltransferase
PAM	Proto-spacer adjacent motif
RBCs	Red blood cells
RFLP	Restriction fragment length polymorphism
RNP	Ribonucleoprotein
RVDs	Repeat-variable di-residues
SCD	Sickle cell disease
SCID	Severe combined immunodeficiency
sgRNA	Single guide RNA
ssOGN	Single stranded oligonucleotides
T7E1	T7 Endonuclease I
TALENs	Transcription activator-like effector nucleases
TIDE	Tracking of Indels by Decomposition
ZFNs	Zinc finger nucleases
β-LCR	β-globin cluster locus control region

SUMMARY

Sickle cell disease (SCD) is a painful and debilitating autosomal recessive blood disorder. A novel approach to treating SCD involves the simultaneous delivery of nucleases specifically targeting the mutant β-globin (HBB) gene and a donor template to induce homology directed repair (HDR)-mediated correction of the SCD mutation in hematopoietic stem cells (HSCs). Examples of site-specific nuclease platforms are zinc-finger nucleases (ZFNs), Transcription activator-like effector nucleases (TALENs), and recently developed clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) systems. One major hurdle for advancing endonuclease-based therapeutic strategies is achieving their efficient expression levels in clinically relevant cell types using currently available delivery approaches. The objective of this thesis is to investigate microinjection, traditionally applied for *in vitro* fertilization of oocytes, for direct, controlled delivery of HBB-aiming nucleases and donor template into human cultured cells. We thoroughly characterized a microinjection system, investigated the effects of microinjection on cell functionality, and demonstrated proof-of-principle of gene editing in human hematopoietic K562 cells microinjected with TALENs and CRISPR/Cas9 and donor template. We found that injection negligibly affects the cell proliferation potential and provides high cell viability, and can be used to control the exposure of nucleases in injected cells. In addition, we performed a study to optimize the delivery of HBB-targeting CRISPR/Cas9 nucleases and nickases and quantify their effects on the hematopoietic potential and cell viability in human HSCs. This work provides insight into the clinical applicability of gene targeting reagents for correcting SCD in HSCs.

CHAPTER 1: REVIEW OF GENE CORRECTION APPROACHES FOR HEMOGLOBINAPTHIES

1.1 Hemoglobinopathies and available therapies

Hemoglobinopathies are a group of genetic blood disorders caused by aberrant hemoglobin expression or structure and are major health burdens with severe mortality worldwide[1,2]. Such disorders, include sickle cell disease (SCD) and β-thalassemia, occurring in 330,000 births annually worldwide according to recent estimates[1,2]. In the United States, SCD alone affects between 90,000 to 100,000 people[3] and is associated with an economic burden exceeding $1.1 billion annually in medical care costs[4]. The molecular basis for SCD is an A to T transversion resulting in the substitution of a non-polar valine for polar glutamic acid in the sixth amino acid position of the β-globin (*HBB*) gene[5], resulting in the formation of hemoglobin S (HbS). This mutation is associated with protection from *Plasmodium falciparum* malaria in people with the sickle cell trait, characterized by heterozygosity for HbS and wild type adult hemoglobin (HbA) typically without SCD symptoms[6,7]. In patients homozygous for HbS, aberrant polymerization of hemoglobin occurs in deoxygenated red blood cells (RBCs), causing them to deform into a sickle-shape and become rigid and adhesive. Repetitive cycles of sickling as HbS molecules switch from oxygenated to deoxygenated states causes RBC fragility and promotes vasocclusions, painful crises, chronic anemia, acute chest syndrome, organ failure, stroke, and death[6,7]. Symptoms indistinguishable from the homozygous HbS version of SCD occur in compound heterozygous patients that co-inherent the sickle cell mutation with other types of hemoglobinopathies, such as β-thalassemia. The severity of the hemoglobinopathy has dependence on levels of persistent fetal hemoglobin F (HbF)

and the underlying mutation[8]. In β-thalassemia, point mutations or small deletions in *HBB* cause a reduction or complete elimination of HBB chains and HbA. Similar to SCD, homozygosity results in a disease phenotype that ranges in severity, from mild forms of anemia, associated with thalassemia intermedia to severe anemia, associated with thalassemia major[6,9]. In addition to anemia, patients homozygous for β-thalassemia have hyperstimulated erythropoiesis and bone marrow cavity expansion by the erythroid tissues, resulting in cosmetic deformities[6].

The available therapies for SCD and β-thalassemia ameliorate the disease symptoms, but are each associated with undesirable side effects and limitations. Erythrocyte transfusion therapy is typically used to treat severe anemia in affected patients, more commonly in treating β-thalassemia major, but is limited by risks of potentially deadly iron overload complications and is often addressed by administering iron chelators[6]. Alternatively, hydroxyurea therapy is administered for inducing the expression of HbF, alleviating clinical features, including episodes of painful crises[10,11]. Although the only FDA approved drug for SCD, hydroxyurea is not effective in treating all patients and is further limited by its myelosuppressive activity in the marrow with unclear risks from long-term treatment[11,12]. Allogeneic hematopoietic stem cell transplantation (HSCT) is the only curative modality for both disorders[13-15]. HSCT involves the replacement of bone marrow, the source of defective RBCs, in affected patients with marrow from a healthy donor **(Figure 1)**. Transplant procedures are typically performed in symptomatic patients less than 16 years of age that have an HLA-identical family member donor[14-16]. Excluded from transplant eligibility are patients with extensive end-organ dysfunction, including severe renal and neurologic impairments or stage II/IV sickle cell lung disease[14,16]. In a multicenter investigation performed in the US and Europe, HSCT resulted in event-free survival in 84% of SCD patients with resolution of vasocclusive crises and anemia as well as improvement in organ function in stably

engrafted patients[16]. One interesting observation reported in HSCT clinical trials is that stable mixed donor-host chimerism was sufficient in reversing the disease phenotype for SCD[14,16] and β-thalassemia[17]. Despite the curative effects, HSCT is associated with therapeutic challenges. Only 14% of SCD patients have histocompatible donors for transplantation[18]. The therapy requires myeloablative conditioning regiments to reduce the risks of graft rejection, which is not well tolerated in many patients, particularly those with severe irreversible complications[19]. As a consequence of higher mortality risks resulting from myeloablative conditioning, transplantation is not an option for adults[19]. Furthermore, HSCT is further limited by graft-verse-host disease, disease re-occurrence, and mortality risks[16] that must be weighed against the severity of the disease symptoms, making decisions regarding HSCT difficult for care-providers and the patients' families.

Gene therapy and emerging precision gene-editing technologies have opened new therapeutic possibilities for SCD and β-thalassemia. This review will describe the application of gene therapy and gene-editing tools for curing SCD and β-thalassemia in animal models and clinical trials (in the case of gene therapy). The clinical challenges and emerging technical opportunities for advancing a novel therapy using each approach will be addressed.

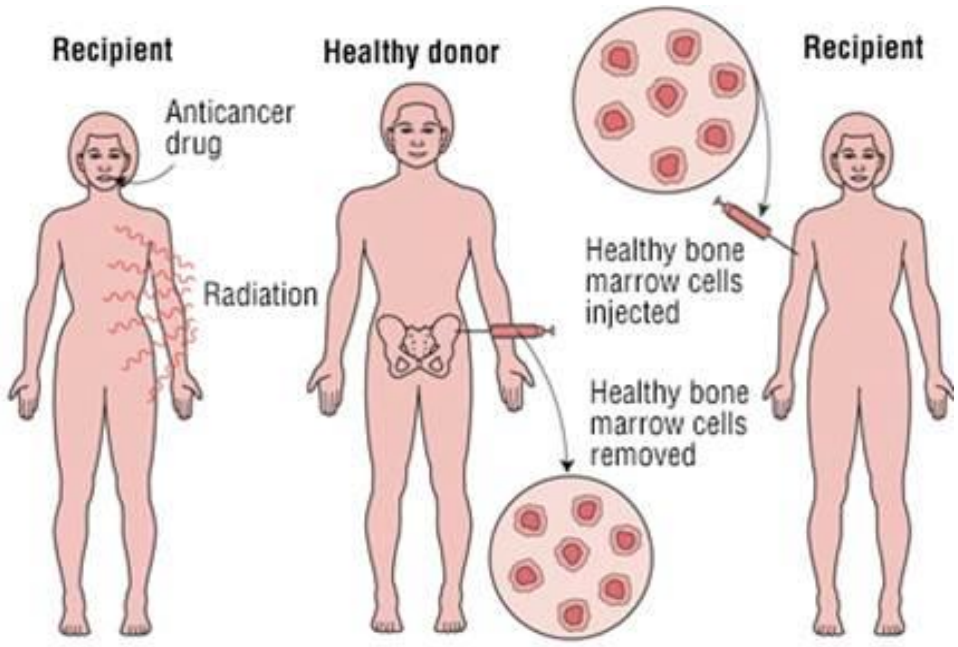


Figure 1: Schematic of allogeneic HSCT

The patient is first treated with chemotherapeutic drugs to ablate the bone marrow. Bone marrow from a healthy donor is harvested and enriched for $CD34^+$ cells using immunolabeling and cell separation techniques. The donor HSCs are infused into the patient to repopulate the entire blood system, thereby replacing the defective RBCs with healthy RBCs. The image is taken from[20].

1.2 Gene therapy for hemoglobinopathies

The goal of gene therapy is transferring the normal *HBB* into hematopoietic stem cells (HSCs) harvested from the patient for subsequent transplantation as a long-term autologous therapy **(Figure 2)**. HSCs can be enriched and isolated from the patient's bone marrow aspirates or peripheral blood by immunolabeling and separating the $CD34^+$ cells. The enriched HSCs can then be transduced using viral vectors containing the correct *HBB* and transfused back into the patient[21]. Alternatively, the patient's whole bone marrow can be transduced and transplanted. Because autologous HSCs have very

low potential to activate graft-verse-host-disease in the patient, immunosuppression may not be necessary post-transplantation.

Establishing proof-of-principle application of the gene therapy approach for correcting hemoglobinopathies has been made possible by the availability of transgenic mouse models for SCD[22,23] and β-thalassemia[24,25]. Additional advances leading to seminal studies have been the genomic mapping of the human *HBB* cluster locus control region (β-LCR)[26-29] and the identification of regulatory elements within the *HBB* cluster required for high levels of erythroid-specific, and vector position independent *HBB* expression[30-32]. Extensive studies demonstrating correction of *HBB* in murine models applied human immunodeficiency virus (HIV)-1-derived lentiviral vectors containing the *HBB* locus and regulatory elements. Lentiviral vectors efficiently infect non-dividing HSCs, provide long-term expression that depends on the number of copies integrated per cell, and have relatively high RNA processing efficiency[33]. In a pioneering study by May et al., mouse bone marrow transduced with *HBB* lentiviral vectors and transplanted into lethally irradiated recipient β-thalassemia mice resulted in 13% of total hemoglobin incorporating the wild type HBB chain at 24-weeks after transplantation[34]. Although there was incomplete transduction into the hematopoietic compartment, as indicated by a proviral copy number of 0.75 per transduced cell, recipient mice transplanted with transduced marrow showed correction in phenotypic features of thalassemia as assessed by improvements in hematocrit level, RBC count, reticulocyte count, and hemoglobin levels relative to control mice[34]. Furthermore, a study by Imren et al. showed that enhanced viral titer preparations for transduction resulted in 3 proviral copies per transduced cell leading to human wild type *HBB* expression in 95% of RBCs[35]. Complete transduction of bone marrow transplanted into severe β-thalassemia mice coincided with 32% of all HBB chains consisting of the wild type HBB, which was sufficient to correct indices of anemia and ineffective erythropoiesis[35]. Additional studies verified that

complete reconstitution of the hematopoietic compartment with donor marrow similarly transduced with human *HBB* or ɣ-globin (*HBG1*) lentiviral vectors provides permanent human hemoglobin expression and ameliorates the thalassemia phenotype in murine models[36-38].

Transplantation of bone marrow transduced with anti-sickling hemoglobin variants and β-LCR resulted in correction of the disease pathology in SCD mouse models. In the study by Pawliuk et al., the $\beta^{A\text{-}T87Q}$ -globin variant, a potent inhibitor of HbS polymerization, having a threonine to glutamine mutation at codon 87, was inserted into a lentiviral vector structurally optimized for erythroid-specific expression and stable gene transfer into HSCs, yielding high viral titers and multiple chromosomal integration events. Transplantation of $\beta^{A\text{-}T87Q}$ lentivirus-transduced bone marrow from SCD mice into lethally irradiated mouse recipients resulted in up to 52% of total hemoglobin consisting of the $\beta^{A\text{-}T87Q}$ variant. High levels of $\beta^{A\text{-}T87Q}$ hemoglobin containing erythrocytes coincided with corrected RBC and reticulocyte counts and amelioration of SCD-associated splenomegaly and urine concentration defect[39]. Comparable correction of a SCD mouse model was demonstrated by Levasseur et al. Minimal amounts of lentiviral vector was used to transduce purified murine HSCs with anti-sickling human β^{AS2}-globin, having substitution of alanine and glutamine at positions 22 and 87 respectively in *HBB*[40]. Primary and secondary transplant mouse recipients showed correction of SCD associated characteristics, including a significant increase in RBC counts, hematocrit values, hemoglobin levels, and restored urine concentration capacity[40].

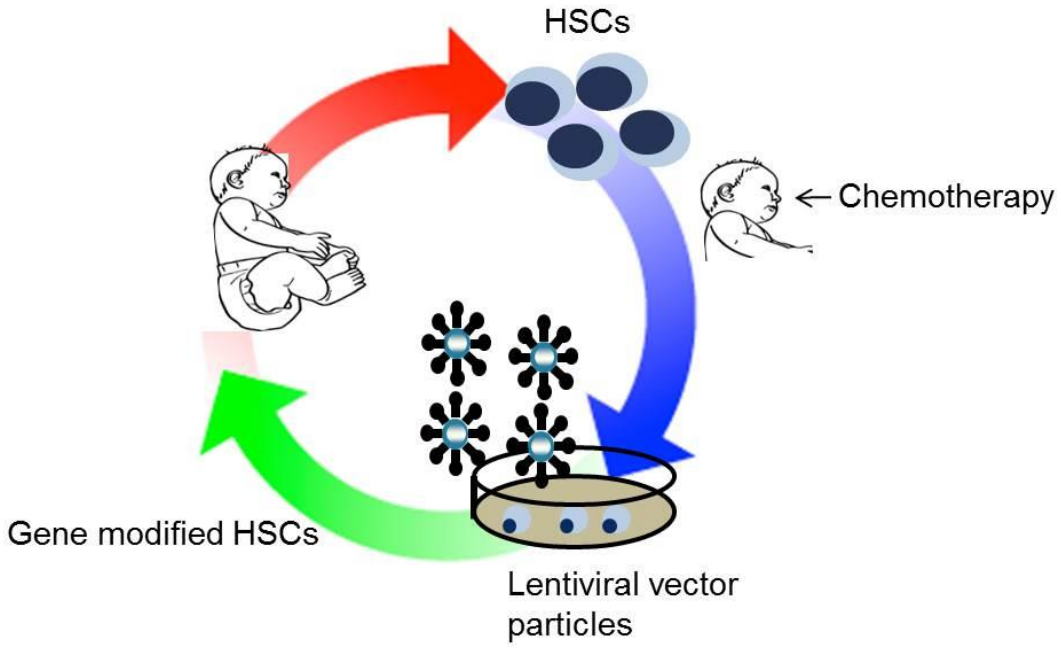


Figure 2: Schematic of gene therapy for hemoglobinopathies

CD34+ cells are collected from the patient with β-thalassemia or SCD and treated with lentiviral vector particles containing the normal *HBB* and LCR elements. The patient is given chemotherapy to ablate the bone marrow and make room for the engraftment by gene modified HSCs. Following viral transduction, the gene modified HSCs are infused back into the patient.

1.3 Challenges with gene therapy

Although gene therapy studies in murine SCD and β-thalassemia models were promising, there are several safety risks that have impeded its clinical application. A major concern is the possibility for insertional mutagenesis by gene transfer vectors. The X-linked severe combined immunodeficiency (SCID) gene therapy human trial in France was halted after two patients developed lymphoblastic leukemia within 30 months after therapy as a result of proviral integration within the *LMO-2* locus leading to aberrant LMO-2 expression in T cells[41-43]. The absence of competent retrovirus suggests that an

insertional mutagenesis event in a transduced $CD34^+$ cell was the basis for malignancy[41]. Furthermore, a study by Imren et al., showed that human cord blood-derived $CD34^+$ cells stably transduced with clinically relevant levels of anti-sickling β^{A-T87Q}-globin gene using a lentiviral vector, yielded sustained production of β^{A-T87Q}-globin in transplanted erythroid progenitor cells *in vitro*. However, a thorough analysis of the proviral integration events revealed that multiple different clones had proviral integrations, although preferring introns, occurring in a number of genes involved in signaling transduction pathways or tumor suppressor genes associated with leukemia[44]. Furthermore, changing patterns of viral integration was observed in a single transplanted mouse recipient over a two month period, highlighting the possibility of tumor-suppressor gene inhibition or activation of an oncogene following proviral gene disruption, potentially resulting in malignancy[44].

There are concerns about the enrichment of sufficient numbers of HSCs for viral transduction. Because the corrected HSCs must compete with non-corrected cells for occupancy in the bone marrow and lack selective advantage, higher doses of treated HSCs is desirable, increasing the chances for success in an autologous transplant in non- or partial myeloablative settings. Furthermore, the requirement for emergency HSCs in the event that the autologous transplant fails to engraft provides even more justification for high-yield collections of HSCs. A non-invasively acquired source of human HSCs for gene therapy is granulocyte-colony stimulating factor (G-CSF)-mobilized peripheral blood. G-CSF has been shown to egress high numbers of $CD34^+$ cells into circulation, however, it is unsafe for SCD patients due to severe adverse side effects, including acute crises and death[45,46]. β-thalassemia patients having splenomegaly as a result of their disease symptoms are at risk for splenic rupture[47] as well as hyperleukocytosis[48] when receiving G-CSF. Therefore, great caution should be used when administrating G-CSF in SCD and β-thalassemia patients. These risks are

further complicated by the fact that G-CSF at safer reduced doses has been associated with poor mobilization of $CD34^+$ cells in β-thalassemia patients[48]. A small clinical trial has shown that combining treatment with G-CSF with plerixafor improves HSC mobilization by several-fold in thalassemia patients with splenomegaly without inducing hyperleukocytosis[48]. However, larger studies are necessary to confirm the safety of the G-CSF and plerixafor combination.

1.4 β-globin gene therapy clinical trials

The successful demonstration of gene therapy in mouse models of SCD and β-thalassemia prompted the initiation of human clinical trials in France. The lentiviral vector used in the first trials, HPV569 LentiGlobin, encodes for the therapeutic $\beta^{A\text{-}T87Q}$ gene and was designed with safety features to lower the risks of mutagenic integration observed in the SCID gene therapy trial[41-43]. Modifications in the lentiviral vector were made to render it a self-inactivating vector, such that the only promoter-enhancer activity was in the β-LCR elements for erythroid-specific expression. Furthermore, insulators containing core elements of the chicken chromatin HS4 was added to prevent activation of neighboring genes. The HPV569 LentiGlobin vector was deemed safe, not showing evidence of activating malignancy in long-term survival studies in mice. The patients selected for the clinical study had severe forms of transfusion-dependent β-thalassemia and SCD and did not have available HLA-matched donors for HSCT. The trial procedure involves isolating $CD34^+$ cells from the patient bone marrow, a portion of which is cryopreserved in the event of graft rejection and the remainder transduced with LentiGlobin. Following *ex vivo* transduction, the treated HSCs are intravenously infused into the myeloablated patient. A sample of transduced HSCs is subjected to replication-competent lentivirus testing to verify safety and quantification of $\beta^{A\text{-}T87Q}$ gene expression. The clinical protocol

was designed with an endpoint analysis of transplant success at 2 years post-transplant and requires recipients to be monitored over 15 years for adverse event analysis, replication-competent lentivirus testing, and insertional mutagenesis testing[49].

The first patient having engraftment of autologous $CD34^+$ cells transduced with HPV569 LentiGlobin was recently reported to be in his 5^{th} year of transfusion independence and without any adverse events since gene therapy[50]. The patient had severe β- thalassemia and previously required transfusions once per month and did not respond to hydroxyurea therapy. The *ex vivo* transduction efficiency measured at 1 week in culture following transduction was 0.6 vector copies integrated per cell[51]. Transplantation of transduced cells (3.9×10^6 $CD34^+$ cells per kg) into the patient, pre-conditioned by Busulfex treatment, resulted in complete hematopoietic reconstitution. The levels of nucleated blood cells containing the $\beta^{A\text{-}T87Q}$-globin vector stabilized at 11% while 36.2% of HBB chains consisted of $\beta^{A\text{-}T87Q}$ by 30 months after transplantation[51]. The peripheral blood cells containing the $\beta^{A\text{-}T87Q}$ vector gradually increased and stabilized at 3.5 g/dL by 2 years after transplantation[50]. The patient had corrected mean corpuscular hemoglobin content at 28.4 pg and a normal range of hemoglobin between 8.5-10 g/dL, one third of the hemoglobin comprised of the therapeutic hemoglobin $\beta^{A\text{-}T87Q}$, which eliminated the need for blood transfusion[50,51]. Although expression from the integrated $\beta^{A\text{-}T87Q}$ vector was erythroid specific, DNA pyrosequencing analysis revealed multilineage chromosomal integration of the HPV569 LentiGlobin vector into multiple sites, with the most abundant integration sites found in the *RFX3, ZZEF1, and HMGA2 loci*. Interestingly, over time there was a growing dominance of cells with integration within the *HMGA2* locus, specifically found in granulocytes-monocytes and erythroblasts, but not lymphocytes. This increase in myeloid cells positive for disruption in the *HMGA2* suggests the clonal expansion of a transduced HSC having a myeloid bias. The *HMGA2* integration site resulted in a 10,000-fold increase in HMGA2 expression compared to

pre-transplant levels due to combined enhancement from the β-LCR and vector-induced truncation causing the *HMGA2* mRNA to become insensitive to microRNA let-7 regulation[51]. The vector-induced enhancement of *HMGA2* mRNA is evidence that the core insulators in the vector likely failed to protect against activation of neighboring genes, thus reinforcing concerns about viral vector safety for therapeutic gene transfer applications.

Viral vector disruption in the *HMGA2* site was implicated in conferring clonal growth advantage in other gene transfer studies[52,53]. Although the effects of *HMGA2* activation on malignancy are unknown in the patient described in[51], one can speculate that there is a lingering risk for transformation. In a study by Ikeda et al., transgenic mice with a similar truncation in the 3' untranslated region of *HMGA2* resulted in microRNA let-7 insensitivity, HMGA2 overexpression, increased peripheral blood cells from all blood lineages, splenomegaly, and EPO-independent erythroid colony formation in bone marrow cells[54]. Competitive and serial bone marrow transplant assays in mice revealed that HMGA2 overexpression conferred clonal growth advantage and self-renewal capacity in HSCs reminiscent of the observations made in the study by Cavazzana-Calvo et al. The results of the Ikeda study also showed upregulation in Jak2 mRNA, pSTAT3, and pAKT proteins with a contaminant decrease in pSTAT5 expression; implicating the involvement of JAK-STAT5 and P13K-AKT signaling pathways in the proliferation of myeloid cells and hematopoiesis. The transgenic mice did not develop lymphoma, but showed splenomegaly and growth advantage in B- and T-cells as a result of enhanced HMGA2 expression[54].

Because of the HMGA2 activation observed in the β-thalassemia clinical trial, the HPV569 LentiGlobin lentivirus vector was improved for safety, and transduction and manufacturing efficiency. The insulator domains found to be unstable in the HPV569 LentiGlobin were removed in the second generation BB305 LentiGlobin vector to

enhance safety. In addition, the promoter/enhancer was changed from the 5' HIV U3 LTR to CMV, resulting in further enhancement in the viral titers and yields. *In vitro* comparison of the two vector designs revealed that the BB305 LentiGlobin provided 3-4 fold higher viral titers and a 2-3 fold higher vector copy number in transduced $CD34^+$ cells relative to the HPV569 vector[55]. In a murine bone marrow transplant study, peripheral blood from mice engrafted with transduced BB305 LentiGlobin had a 1.1-1.5 fold higher vector copy number relative to the HPV569 vector. Both HPV596 and BB305 LentiGlobin vectors displayed preferred integration into gene coding regions with 51.9% of integrations found in common insertion site regions, which have not previously been associated with adverse events in patients. Preferred insertional integration was not observed in high risk genes LMO2 or MDS1-EVI1, associated with transformation, for either HPV596 or BB305 LentiGlobin vectors in primary and secondary transplanted mouse recipients[55], thus providing evidence of vector safety. Clinical trials for treating β-thalassemia and SCD patients using the BB305 LentiGlobin lentiviral vector is currently in progress in France and the United States (ClinicalTrials.gov, identifier NCT02151526).

1.5 Homologous recombination mediated gene correction

Gene replacement of the HBB gene containing mutations causing SCD or β-thalassemia with a normal copy of HBB via homologous recombination is an alternative approach that may diminish the risks associated with gene therapy. Proof-of-principle for was established in the study by Wu et al.[56]. The study derived embryonic stem cells (ESCs) from blastocysts from a SCD mouse model, and subsequently transfected them with a targeting vector containing the human γ-β^A-globin gene fragment flanked by 5' and 3' homologous mouse sequences, which resulted in successful homologous recombination-induced HBB gene replacement in 14.2% of isolated colonies. Mice

derived from the corrected sickle ESCs had high levels of HbA expression in differentiated RBCs. Furthermore, anemia and organ pathology was ameliorated in the mice derived from the corrected ESCs, suggesting that HDR induced gene replacement was sufficient to correct the SCD genotype and symptoms[56]. A similar study demonstrated that HDR induced gene corrected ESCs derived from sickle cell mice can be differentiated into hematopoietic cells[57]. However, the majority of the hematopoietic cell clones derived from the targeted ESCs expressed both HbS and HbA[57]. Clinical application of this approach for treating SCD or β-thalassemia would involve a multiple step process, including generating patient derived ESCs by somatic cell nuclear transfer, correction of the mutated HBB gene, and differentiation of ESCs into HSCs[57]. ESCs are highly desirable to use as substrates for this therapeutic approach because they have the potential to differentiate into all types of cells. However, the major weaknesses of this therapeutic approach is the ethical and religious implications associated with the clinical use of human ESCs derived from donor oocytes, and potential immune rejection risks resulting from immunological incompatibility of derived HSCs[58].

To overcome some of the complexities associated with deriving patient customized ESCs, correction in induced pluripotent stem cells (iPSCs) was proposed as an alternative strategy. iPSCs are derived by direct reprogramming of mouse or human somatic cells using retroviral transduction of several transcription factors and have genetic and epigenetic similarities with ESCs[59-61]. In the study by Hanna et al, mouse iPSCs were derived by transducing fibroblasts isolated from an adult humanized mouse model of sickle disease with Oct4, Sox2, Klf4, and c-Myc transcription factors[62]. The resulting ESC-like cells expressing pluripotent markers were used to generate cell lines corrected through homologous recombination using targeting constructs containing the normal human HBB gene, differentiated into hematopoietic progenitors *in vitro*, and subsequently transplanted into irradiated SCD recipient mice[62]. The recipient mice

injected with the corrected iPSC derived hematopoietic progenitor cells had stable engraftment and corrected genotype, expressed HbA proteins, and had ameliorated pathological features of SCD[62]. Human iPSCs have been derived from various cell sources, such as skin fibroblasts, amniotic fluid cells, and chronic villus samples harvested from a homozygous β-thalassemia patient, and were shown to differentiate into hematopoietic progenitor cells; raising the possibility for utilizing patient derived iPSCs to treat patients with β-thalassemia or SCD at early neonatal stages[63]. The study by Wang et al. demonstrated proof-of principle for correcting the β-thalassemia mutation in iPSC derived from fibroblasts harvested from a homozygous β-thalassemia patient[64]. The patient-derived iPSCs were corrected by transfecting a homologous targeting vector encoding for the normal HBB gene and drug resistance marker, retained expression of pluripotent marker genes, and were capable of differentiation into hematopoietic progenitor cells[64].

Although pluripotent stem cells hold promise as tools for treating hemoglobinopathies and other genetic diseases, the main challenge for their clinical application is safety. ESCs and iPSCs have uncontrolled proliferation and tissue differentiation potential and are capable of forming teratomas at the implantation site[65]. Stringent cell separation methods are necessary to produce pure populations of differentiated cell progeny and reduce the risks of teratoma formation[65]. Furthermore, the protocols for differentiating iPSCs or ESCs have low and variable efficiencies, and many require the use of animal feeder layers that can potentially introduce animal pathogens[65]. An inherent limitation of the gene targeting strategy using transgenes encoding for the correct HBB gene is the low frequency of homologous recombination events in mammalian cells, occurring in roughly 10^{-3} to 10^{-7} transfected cells[66]. In the study demonstrating homologous recombination-mediated gene correction of the β-

thalassemia mutation in patient-derived iPSCs, the targeting efficiency measured was 0.81% with drug selection[64], thus, underscoring the low efficiency for this approach.

1.6 Precision gene-editing tools

Mammalian cells have complex machinery to repair double stranded breaks (DSBs) in DNA and maintain genomic integrity. A DSB can be repaired by two main pathways: non-homologous end-joining (NHEJ) and homology directed repair (HDR)[67]. Repair by the NHEJ pathway results in small deletions and/or insertions at the break site, generated to obtain alignment of complementary bases[67]. The HDR pathway involves high fidelity repair of the broken ends using homologous sequences found in sister chromatids, homologous chromosomes or exogenous DNA containing homologous sequences[67]. In these studies, the homing I-SceI endonuclease (also called meganucleases) from *Saccharomyces cerevisiae*, which generates DSBs at a targeted recognition sequence of 18-bp[68], was co-transfected along with a homologous donor plasmid DNA into human COS-1[69] and mouse ESCs[70] to induce HDR mediated knock-in of the donor sequence into the I-SceI target site. Additional studies confirmed that the induction of a DSB by meganucleases can dramatically increase the frequency of HDR by 3-5 orders of magnitude when a homologous exogenous donor plasmid DNA is provided[71]. Although meganucleases can catalyze DSBs critical for genetic alterations, modifying the recognition specificity of meganucleases such as I-SceI limits their use as a gene-editing tool for stimulating site-specific alterations[72]. This challenge inspired the design of customizable designer nucleases capable of targeting pre-existing genomic sequences that can mediate different types of gene modifications through activation of NHEJ and HDR **(Figure 3)**. In the following paragraphs, we will explore different precision gene-editing tools.

One class of versatile targeting nucleases is zinc finger nucleases (ZFNs) **(Figure 4a)**. The ZFN structure consists of a zinc finger protein (ZFP) DNA-binding domain composed of zinc finger motifs[73] fused to the nuclease domain of the *FokI* restriction enzyme, conferring the hybrid protein with DNA binding capability and robust cleavage activity[74]. The ZFPs contain a tandem array of Cys_2-His_2 units, each roughly 30 amino acids in length, bound to a zinc atom, and recognizing 3 base pairs of DNA[75]. DNA cleavage activity requires dimerization of the *FokI* domains in a pair of ZFNs, each targeting neighboring recognition sequences, arranged in an inverted orientation on the genome, and separated by a spacer region[76]. The dimerization requirement can be exploited to prevent cleavage events at single binding sites, increasing the overall specificity[77]. When delivered into cells, ZFNs catalyze DSBs repaired by NHEJ or HDR pathways, resulting in several different genetic alteration outcomes, such as insertions, deletions, inversions, duplications, and translocations[78]. In an early study, delivery of a ZFN pair, each composed of three ZFPs specifying 9-bp, along with homologous donor template DNA was used to stimulate gene targeting in the X-chromosome of the *Drosophila melanogaster* germ line[79]. Since then, additional studies demonstrated ZFN mediated gene modification in many different organisms[80-82] and human cell lines[83,84]. The versatility of potential genomic modifications makes ZFNs an attractive tool for therapeutic purposes. ZFNs were designed to disrupt the HIV host co-receptor CCR5 gene [85] as an HIV therapy currently in Phase 2 clinical trials[86]. However, the major weakness of the ZFN nuclease platform is its limited targeting range using publically available methods, making it difficult to design a ZFN at a desired target site[87].

Transcription activator-like effector nucleases (TALENs) are a second class of hybrid tailorable nuclease for gene-editing **(Figure 4b)**. TALENs consist of the DNA binding domain from TALE proteins isolated from bacteria of the *Xanthomonas* genus, fused to the DNA cleavage domain from the *FokI* restriction enzyme, conferring TALENs with

DNA binding capability and high cleavage activity[88]. A DSB is formed at the target locus by a TALEN pair, each binding adjacent elements on separate strands of the DNA in a tail-to-tail orientation, separated by optimized DNA spacing required for *FokI* dimer formation[88]. The TALE DNA binding domain is modular, consisting of 15.5-19.5 repeats roughly 34 residues in length that each recognizes a specific nucleotide determined by the repeat-variable di-residues (RVDs) at positions 12 and 13 within the module[88]. The Golden Gate cloning strategy is used to assemble libraries of unique arrangements of RVD repeats from which it is possible to screen using a cognate target genomic sequence[89]. An attractive feature that makes TALENs more desirable as a tool for genome editing compared to ZFNs is the availability of the RVD-DNA recognition code that identifies the range of specificities of RVD sequences, including RVDs that specify unique nucleotides, enabling design of customized binding domains[90,91]. Although TALENs have been shown to have comparable cleavage activity to ZFNs in studies targeting the same genes, its simpler design and broader targeting range makes TALENs more attractive than ZFNs[78]. TALENs have been applied for targeted genomic editing in various animal species[92,93] and human pluripotent cells [94].

The third class of gene editing tools are the recently developed clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) system **(Figure 4c)**[95]. CRISPR systems play a role in the adaptive immune systems for bacteria, providing protection from invading nucleic acids[96]. The Cas9 endonuclease is directed to a target DNA site by a single guide RNA (sgRNA) through Watson-Crick base-pairing rules[97]. In the type II CRISPR system of *Streptococcus pyogenes*, Cas9 endonuclease cleaves at the complementary 20 nucleotide target site specified by the sgRNA immediately 5' of the NGG proto-spacer adjacent motif (PAM) sequence[97]. The commonly used CRISPR/Cas9 system for gene-editing involves the co-delivery of the Cas9 endonuclease and sgRNA often encoded on a single plasmid DNA.

Redirecting the CRISPR/Cas9 nuclease to desired targets in the genome requires modification of the targeting sgRNA sequence while the other components remain fixed, making the process of developing CRISPR/Cas9 nucleases simpler compared to ZFN and TALEN platforms. Furthermore, the Cas9 endonuclease can be co-delivered with two or more sgRNAs targeting multiple sites simultaneously, enabling multiplex disruption within the genome[98]. The main disadvantage of the CRISPR/Cas9 nuclease platform is its high frequency of cleavage activity at off-target sites in the genome, resulting in gross chromosomal deletions and other types of chromosomal rearrangements[99-101]. Tolerated mismatches and bulges between the sgRNA and gene sequences were implicated for non-specific cleavage by the CRISPR-Cas9 system[102], providing evidence for the need to carefully screen potential off-target sites during the design stage. The CRISPR/Cas9 system has been applied for efficient generation of a wide range of transgenic animal models[103-107] and engineered plants[108-110]. We will explore therapeutic applications for precision gene-editing tools in the next section.

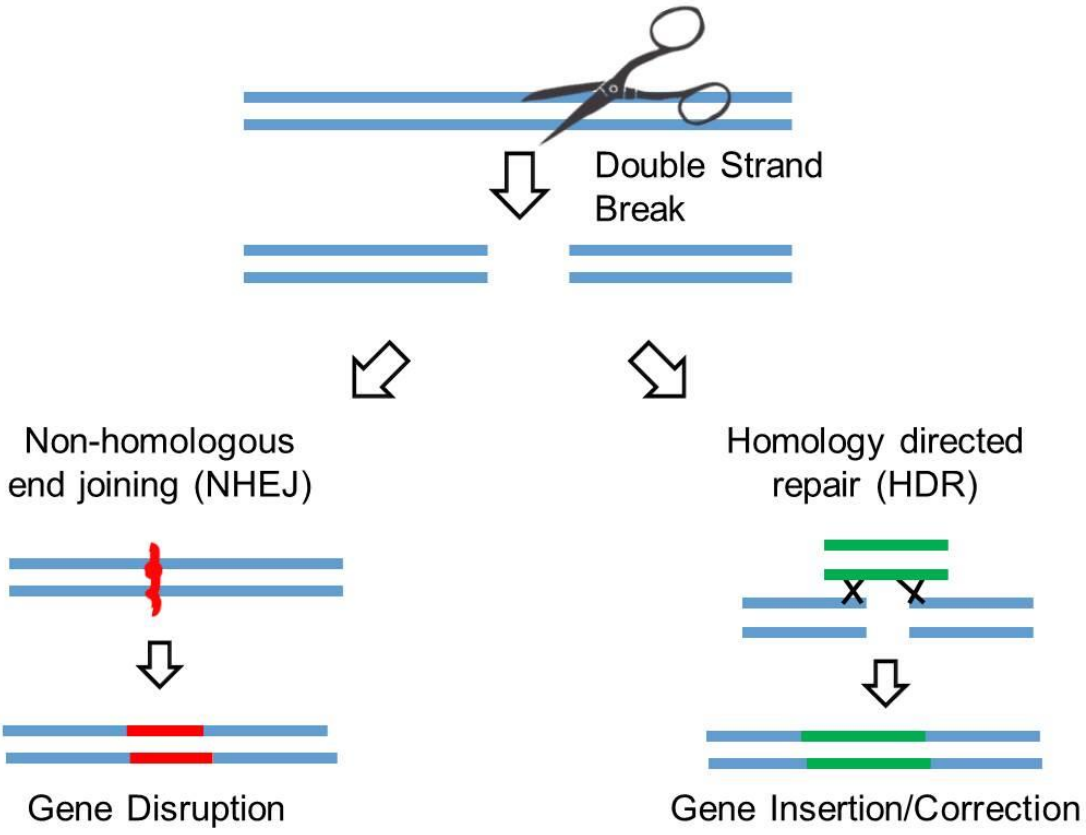


Figure 3: Gene modification outcomes following the induction of a DSB by nucleases

The DSB is resolved by endogenous repair machinery involved in the NHEJ and HDR pathways. Activation of the NHEJ repair pathway leads to insertions and/or deletions of sequences at the DSB site to enable resecting of both strands for repair, resulting in gene disruption. The simultaneous delivery of nucleases and donor repair template DNA activates repair using the HDR pathway, whereby the information encoded on the donor template is transferred to the target site for gene insertion or gene correction.

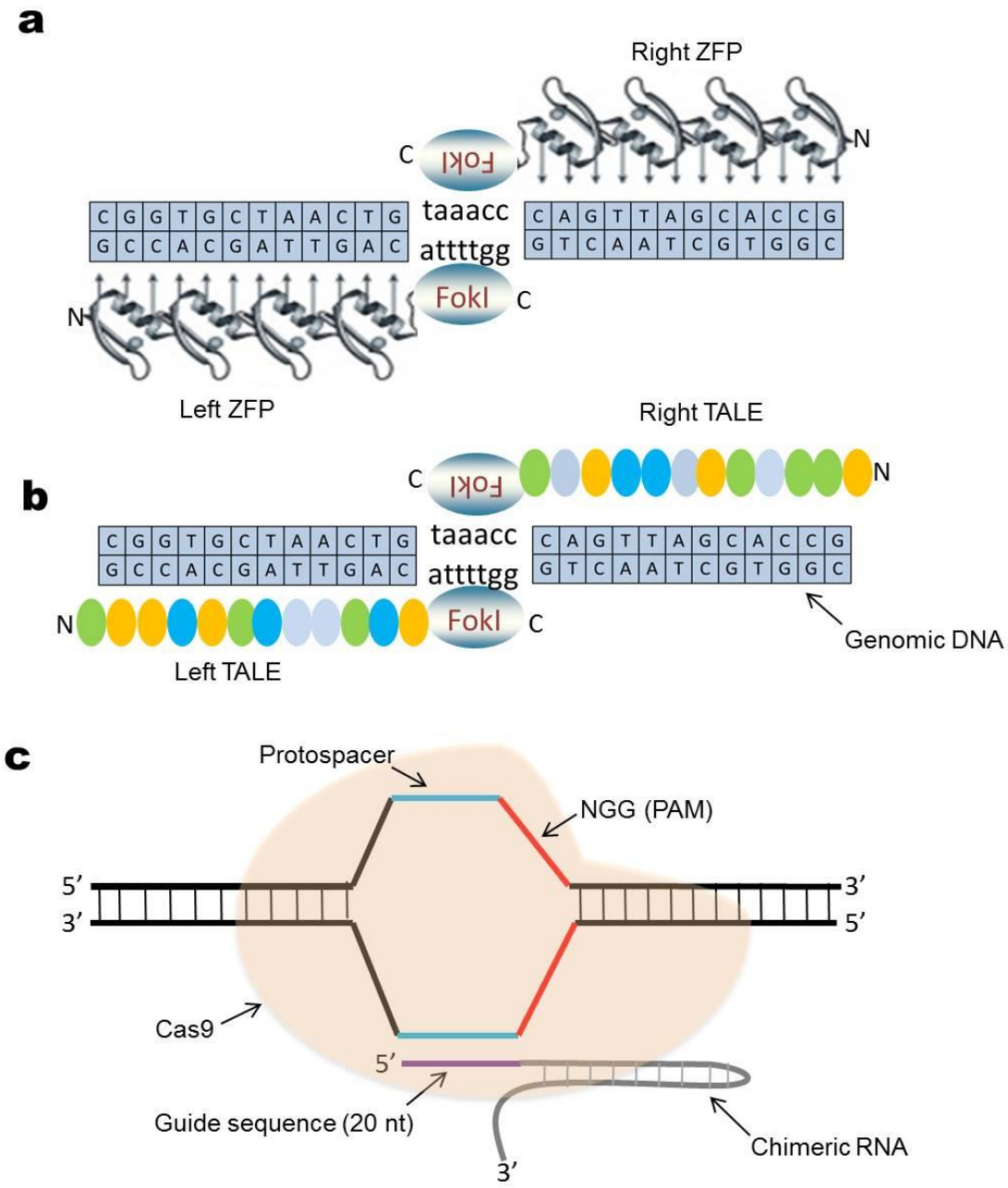


Figure 4: Schematic of precision gene-editing tools

(a) ZFNs consist of a DNA binding ZFP fused to a *FokI* restriction enzyme. Each ZFP recognizes 3 nucleotides and are linked together to mediate binding to a desired target site. A ZNF pair binds to the DNA in a head to tail orientation. **(b)** TALENs consist of TALE repeat domains, each recognizing a single nucleotide, fused to FokI. The effector domains from a TALEN pair bind to adjacent effector elements in a tail to tail orientation with optimized spacing. The generation of a DSB within the spacing between TALE and

ZFN pairs requires the dimization of the *FokI* domains. **(c)** In the CRISPR/Cas9 systems, a Cas9 endonuclease is guided by a chimeric sgRNA containing the complementary sequence (purple) to the target site immediately 5' of the PAM sequence (red).

1.7 Gene correction using precision gene editing tools

Therapeutic gene correction for hemoglobinopathies involves the replacement of mutant sequences within the *HBB* locus with a wild type donor sequence through activation of the HDR repair pathway **(Figure 5)**[56,57]. As discussed earlier, the frequency of HDR can be dramatically increased by inducing a DSB using site-specific nucleases[111,112]. Early work by Urnov et al. demonstrated gene correction of a disease mutation in human cell lines and primary cells treated with ZFNs and donor vector [84]. The study observed up to 20% gene targeting in K562 cells, a human leukemia cell line, treated with *IL2RG* ZFNs and donor vector without selection[84]. Comparable levels of HDR was observed in human $CD4^+$ T cells treated with ZFNs[84]. Interestingly, ZFN induced HDR of the target site occurred at the highest frequency in G2 arrested cells, suggesting that gene targeting is favored in the S/G2 stage of the cell cycle[84]. The work by Lombardo et al. confirmed that gene targeting in a disease related gene downstream of its own promoter is feasible using ZFNs and donor vector co-delivered into a variety of human cell types, including ESCs and HSCs[113]. Studies have shown that TALENs[114,115] and the recent CRISPR/Cas9 nucleases[116] provide efficient gene targeting of the SCD mutation in the *HBB* locus in K562 cells.

Early studies demonstrating efficient gene targeting by ZFNs[117] and TALENs[94] combined with donor template in human ESCs and iPSCs provided the motivation for developing therapeutic strategies in patient-specific pluripotent stem cells. Pioneering work by Yusa et al. demonstrated gene correction of the point mutation in *A1AT* causing

α1-antitrypsin deficiency in iPSCs, resulting in the rescued A1AT function in derived liver cells[118]. A1AT ZFNs and donor vector, designed with a drug selection marker flanked by piggyBAC repeats, were transfected into human iPSCs derived from patients with α1-antitrypsin deficiency[118]. The drug-resistant iPSC clones obtained showed a targeting efficiency of 54% and 4% in one allele and two alleles respectively[118]. The piggBAC repeats mediated seamless removal of the drug cassette, providing biallelic excision in 11% of colonies that also showed corrected *A1AT* sequences in both alleles[118]. ZFN induced gene correction did not alter the pluripotency of corrected cells, whereby differentiation into hepatocyte-like cells *in vitro* resulted in the functional correction of the disease phenotype[118]. Similarly, phenotypic correction was observed in mice transplanted with corrected iPSC-derived hepatocytes[118]. A similar approach was used for correcting the SCD mutation in work by Sebastiano et al.[87]. Human iPSCs derived from individuals with SCD were transfected with *HBB* ZFNs and donor vector, containing the correct *HBB* and a drug selection cassette flanked by loxP sites[87]. HDR mediated gene targeting was achieved in up to 37% of drug resistant clones, each having retained their pluripotency and normal karyotype[87]. A limited investigation of the ZFN specificity showed that there were no mutations generated in cognate off-target sites[87]. The floxed drug resistant gene and reprogramming cassette was excised by transient expression of Cre recombinase in corrected iPSCs, demonstrating efficient generation of transgene-free corrected iPSC cell lines from SCD patients[87]. A similar study confirmed the feasibility of generating gene corrected iPSCs derived from SCD patients using *HBB* ZFNs and donor vector[119]. However, even after excision of the drug selection cassette using the Cre-LoxP system, the endogenous gene expression of the corrected allele was only partially restored in erythroid-differentiated cells, likely because of interfering “scar” sequences remaining after excision[119]. Challenges excising the drug cassette and

differentiating iPSCs after treatment using nucleases and donor vector may limit the clinical applicability of this approach.

HSCs are more clinically relevant compared to pluripotent stem cells for targeted gene editing strategies for treating hemoglobinopathies because they can differentiate into healthy RBCs *in vivo*. The work by Genovese et al. was the first to demonstrate successful HDR-mediated gene targeting in human HSCs using site-specific nucleases[120]. In this study, cord blood $CD34^+$ cells were treated with ZFN mRNA and donor vector designed to transfer GFP into the mutational hotspot of *IL2RG*, associated with X-linked SCID, and the AAVS1 'safe harbor' site[120]. The ZFN stimulated HDR frequency observed was 5% in bulk cultured $CD34^+$ cells[120]. In a repopulation assay using non-obese diabetic (NOD)/SCID mice, 100% of mice had on average 3% GFP positive, targeted cells over the long-term[120]. $CD34^+$ cells were also treated in the Genovese study with *IL2RG* ZFNs and a targeting vector spanning exons 5-8 of *IL2RG* to stimulate gene correction of SCID-XI deficiency[120]. Transplantation of treated $CD34^+$ cells into NOD/SCID mice resulted in the functional reconstitution of *IL2RG*, as evident by T cells that were physiologically similar to healthy controls [120]. Evaluation of off-target activity by the *IL2RG* ZFNs revealed negligible off-target mutation at previously identified cognate sites[120]. In work by Hoban et al., the ZFN platform was applied for targeting the SCD mutation in human HSCs[121]. $CD34^+$ cells treated with *HBB* ZFNs along with donor vector containing a silent restriction fragment length polymorphism showed 18.5% average gene targeting[121]. Similar levels of gene conversion was measured in healthy $CD34^+$ cells treated with ZFNs and donor encoding the SCD mutation, and in $CD34^+$ cells isolated from SCD patients treated with ZFNs and donor encoding wild type *HBB*[121]. The corresponding HbA protein level was 5.3% in erythrocytes derived from patient $CD34^+$ cells treated with ZFNs and wild type *HBB* donor, providing evidence that precision gene-editing tools can functionally correct hemoglobinopathies[121]. Although

bulk CD34$^+$ cells treated with *HBB* ZFNs and donor template engrafted equally as efficiently as control non-treated cells and differentiated into all expected lineages, there was up to a 50-fold reduction in the presence of gene modified cells engrafted in these mice compared to the input values [121]. High throughput sequencing revealed only 0.85% of the bone marrow in mice transplanted with ZFN- and oligonucleotide-treated cells were gene modified, whereas the bulk population prior to transplant had 17.3% gene modification[121]. This result suggests that gene modified cells have a competitive disadvantage for engraftment compared to unmodified CD34$^+$ cells. An assessment of off-target activity revealed that the *HBB* ZFNs had high levels of specificity with off-target modification occurring only in the homologous δ-globin gene (*HBD*)[121]. It is unclear whether gene modification within the *HBD* locus will have adverse consequences in a clinical setting. These seminal studies demonstrate efficient gene targeting by ZFNs in human CD34$^+$ cells, however, further research addressing low levels of engraftment by gene modified cells in the xenograft model will be critical for advancing a nuclease-based therapeutic strategy.

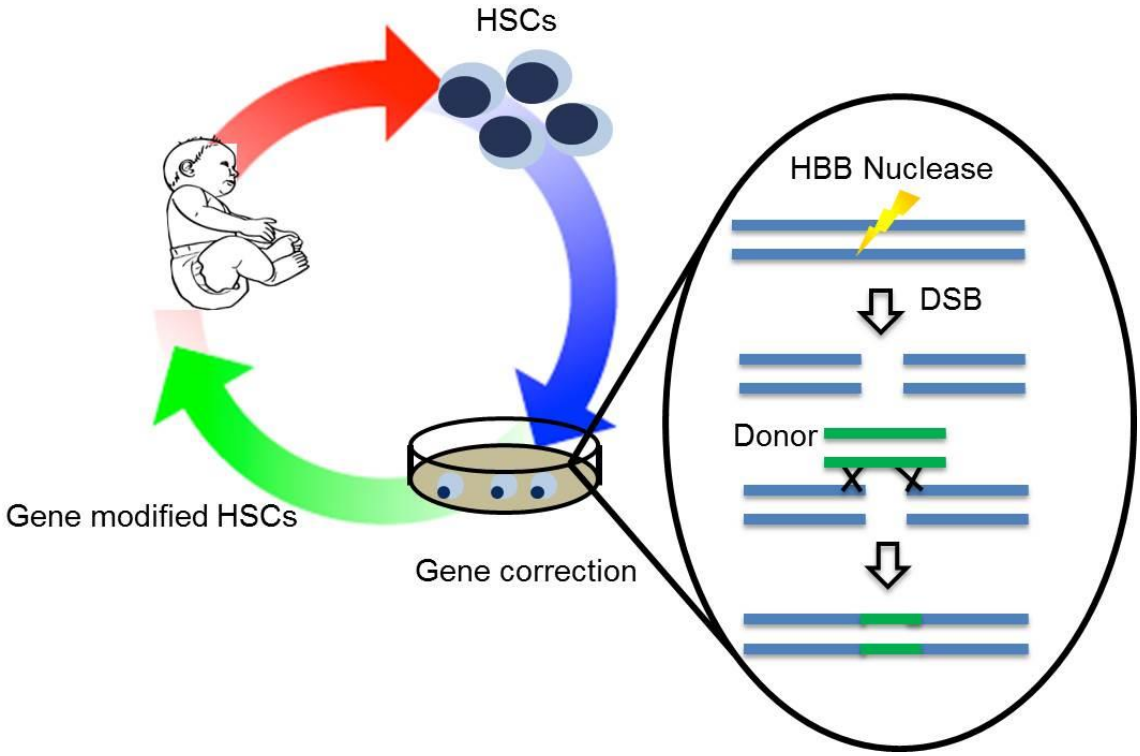


Figure 5: Schematic of gene correction approach for treating hemoglobinopathies

$CD34^+$ cells collected from a β-thalassemia or SCD patient and treated with HBB-targeting nucleases and donor template DNA. HDR mediated repair of the DSB stimulated by the nucleases will enable transfer of the normal copy of HBB gene into the genome. Transfusion of the nuclease induced gene modified HSCs will provide long-term replacement of the defective RBCs with healthy RBCs for the lifetime of the patient.

1.8 Future directions

Nuclease mediated gene correction is an attractive approach as an autologous therapy for hemoglobinopathies because it eliminates insertional mutagenesis risks associated with gene therapy[42,43]. However, there are several barriers with gene correction using precision gene editing tools that must be addressed to advance its clinical application. Firstly, optimizing the design of nucleases and donor template to enable high levels of HDR mediated transfer of the normal copy of *HBB* into the mutation

site is a major challenge. Although the donor template can be simplified through single-stranded oligodeoxynucleotides, their design involves optimizing the length of homologous sequences, which may not be straightforward for high levels of gene correction. The process for constructing TALENs and ZFNs are difficult to perform, even for experienced researchers, and involves complicated, low throughput cloning steps for assembling multiple plasmids encoding a single nuclease pair[78,122]. TALEN RVDs have multiple specificities for different nucleotides, leading to cleavage activity in off-target sites[78]. Likewise, zinc finger domains can also tolerate binding to off-target sites and display text dependency for high cleavage activity[122]. The specificity for both platforms can be improved by increasing the lengths of the DNA binding domains or through the use of heterodimeric *FokI* nuclease domains[77,123]. There are several online design tools available to help guide the design of TALENs[124,125], ZFNs[126,127], and CRISPRs[126,128-130] using reagents accessible to the research community and enable accurate prediction of their off-target cleavage sites *in silico*[101,131,132]. The CRISPR/Cas9 system is easier to design and only involves modifying the 20 nucleotide sequence within the sgRNA to bind to a new target sequence, but they have been shown to have substantial off-target activity[99-101], which may limit the applicability of the system for gene therapies. Alternatively, the specificity can be improved through the Cas9 nickase system, shown to facilitate efficient gene editing with 50-1500-fold lower off-target activity[133,134]. Future studies should investigate gene targeting of the *HBB* using CRISPR/Cas9 nickases as a strategy to enhance HDR activity without compromising the specificity. A challenge shared by all precision gene editing tools is accurately quantifying nuclease activity. Most studies rely on gel assays that quantify nuclease activity with limited accuracy and sensitivity. Alternatively, the traffic light reporter system can be used for measuring both NHEJ and HDR in single cells treated with gene editing reagents[135], but this system is time consuming and difficult to construct. Deep sequencing provides higher accuracy for

quantifying gene modification at on- and off-target sites[116], but is limited by the high cost and time required for processing and analysis. Recently, a novel method called Tracking of Indels by Decomposition (TIDE) was developed a faster and cost-effective approach for accurate quantification of activity[136]. TIDE analysis may enable more efficient optimization of *HBB* targeting nucleases.

Secondly, gene targeting in clinically relevant cells, such as human $CD34^+$ cells, requires high levels of nuclease protein expression and sufficient donor template delivered into cells. Therefore, the development of delivery methodologies for nucleases and donor template is critical for advancing a nuclease-based therapy for hemoglobinopathies. The requirement for the donor and nucleases to be encoded by separate constructs makes it more challenging to deliver all reagents needed for gene targeting into single cells with high efficiency. Integrase deficient lentivirus (IDLV) has been shown to provide high levels of nuclear delivery and have significantly reduced capacity to integrate into the genome, thus, they are progressively lost during cell expansion[137]. Seminal studies showed that ZFNs and donor template transduced using IDLVs results in low levels of gene targeting in $CD34^+$ cells[113]. A limitation of this approach is low permissiveness of human HSCs to multiple infections by separate IDLVs encoding nucleases and donor template[113]. In contrast, delivering the donor using IDLV vector transduction and the ZFN mRNA using electroporation resulted in considerably higher levels of gene targeting[120,121]. Interestingly, it was found that cell cycle activation through culture stimulation is necessary to obtain optimized levels of lentiviral delivery and gene targeting, particularly, in primitive HSCs[120]. The addition of aryl hydrocarbon receptor antagonist and 16,16-dimethyl-prostaglandin E2 to the culture media was found to reduce differentiation and may contribute to enhanced gene targeting[120]. Although mRNA delivery of nucleases is less toxic in cells compared to plasmid DNA, they still provide high levels of off-target gene modification[121]. Future

studies can investigate whether nuclease specificity and toxicity can be improved through the direct delivery of their purified proteins into cells[138].

Thirdly, the therapeutic benefit of functionally corrected HSCs depends on their capacity to engraft and provide long-term production of healthy RBCs in a patient. Therefore, a critical topic for investigation is the engraftment potential of human $CD34^+$ cells treated with nucleases and donor template. The study by Hoban et al. show high levels of gene conversion in human HSCS treated with ZFNs and donor template from a SCD patient, however, the engraftment potential of the patient HSCs was not characterized[121]. Thus, it is unclear whether gene corrected HSCs from a SCD patient would engraft, particularly, within the context of an autologous transplant. Addressing this problem will require the development of xenograft models of hemoglobinopathies that better recapitulate the disease phenotype.

The amount of HbF persisting into adulthood influences the severity of hemoglobinopathies[139]. There are a number of different single nucleotide polymorphisms identified in patients with hemoglobinopathies associated with persistent HbF production and mild to non-existent disease symptoms[140,141]. The largest therapeutic effect comes from mutations in the *BCL11A* locus leading to lower levels of BCL11A transcripts encoding for a multi-zinc finger transcription factor that binds to discrete regions in the *HBB* cluster in adult erythroid progenitors, resulting in reduced transcriptional repression of HbF[142]. *BCL11A* knock down led to a modest reduction in the levels of *BCL11A* transcripts with concomitant enhanced HbF expression without affecting erythroid differentiation in human erythroid progenitor cells[142]. Therefore, *BCL11A* represents a therapeutic target for treating hemoglobinopathies. It was postulated that directed knock down of *BCL11A* in SCD or β-thalassemia patients would elevate the HbF levels and ameliorate the disease severity[142]. The study by Bauer et al. identified specific hypersensitive sites contained within *BCL11A* intron-2, representing an erythroid specific

enhancer element, having a strong association with hereditary persistent HbF[143]. This study demonstrates the application of TALENs for deleting regions flanking hypersensitive sites of *BCL11A,* resulting in the loss of the *BCL11A* enhancer and absence of BCL11A expression in the erythroid lineage, while not affecting BCL11A expression in non-erythroid cells[143]. In recent phase I clinical studies by Sangamo, erythrocytes differentiated from human HSCs treated with *BCL11A* ZFNs showed elevated levels of HbF, consisting of over 40% of all hemoglobin[144]. Comparable levels of HbF were observed in HSCs isolated from β-thalassemia patients[144]. Furthermore, the HSCs modified with *BCL11A* ZFNs were capable of engraftment in the xenograft mouse model and showed low levels of off-target activity[144].

Overall, we look forward to the development of a gene therapy for hemoglobinopathies. Targeted correction of the *HBB* locus or *BCL11A* down regulation through the application of precision gene-editing tools may enable the development of a novel curative autologous therapy. Further research efforts addressing design, specificity, and delivery of gene targeting reagents into HSCs and the development of mouse models for accurately quantifying the repopulation efficiency by gene corrected human HSCs is critical for the clinical translation of the strategy.

CHAPTER 2: THESIS DISSERTATION OVERVIEW

2.1 Introduction

An effective therapy for SCD is urgently needed. Currently, the only FDA approved drug for ameliorating the clinical features of SCD is hydroxyurea therapy, shown to be effective in some patients and is associated with unclear risks from long-term use[11]. HSCT is the only curative therapy for SCD, but is not available for the vast majority of patients due to stringent eligibility requirements, including an HLA-matched donor[18]. Gene therapy involving the non-targeted insertion and forced expression of normal or anti-sickling HBB variants has been proposed, showing promising results for functionally correcting the disease phenotype in murine models[34,35] and in human clinical trials[51]. The major issue with gene therapy is the risk of insertional mutagenesis as observed in the clinical trial for X-linked SCID[41,43]. A novel curative approach for SCD involves the application of precision gene-editing tools to stimulate targeted gene correction of the disease causing mutation in HSCs. One major hurdle is delivering nucleases along with donor template into HSCs *ex vivo* with high delivery efficiency[120,145] while maintaining the cell viability and engraftment potential. This dissertation explores and compares different *ex vivo* delivery methods, namely, microinjection and nucleofection, for nuclease-mediated gene modification and determines the nuclease targeting efficiency and specificity.

2.2 Specific Aim 1

Microinjection has many advantages, namely, the ability to precisely control the amount of molecules delivered into single cells with theoretically 100% efficiency and accuracy[146]. For difficult to transfect cells, microinjection is an attractive option because it obviates the need to use antibiotic selection for cell enrichment, making injection less toxic and stressful in cells[146]. In this aim, we will develop a system for microinjecting human suspension cells using glass microcapillaries. The innovative feature of this aim will be adapting a microinjection technique, traditionally used for adherent cells, and applying it for injecting somatic cells grown in suspension. Because human $CD34^+$ cells and culturing media are expensive, it will not be feasible to use them during the development and characterization of the microinjection system. In lieu of HSCs, we will use leukemia K562 cells as a model cell line because they have similar morphology to human $CD34^+$ cells, including a high nucleus to cytoplasm ratio, and less expensive to culture. We will explore techniques for artificially immobilizing K562 cells to facilitate injection using a single glass microcapillary, optimize the rate and efficiency of injection, and investigate parameters to accurately control the injection volume. In addition, we will quantify the effects of injection on the cell viability and proliferation potential in single injected cells.

The second component of this aim is to develop *in vitro* culturing and nucleofection protocols for human $CD34^+$ cells. Nucleofection is a modified electroporation method involving the application of optimized electrical pulses and cell line-specific buffers to permeablize the cell membrane and directly deliver molecules into the cell nucleus. We will use a 4D-nuclefector and Amaxa nucleofection kits (Lonza) containing cuvettes and buffers specific for the K562 cell line and human $CD34^+$ cells. K562 cells will be cultured and nucleofected with protocols that are well established by

the manufacturer and provide high cell viability and transfection efficiency, but there are not well established protocols for culturing and nucleofecting $CD34^+$ cells. Therefore, the development of *in vitro* culturing and nucleofection protocols for $CD34^+$ cells will represent a meaningful contribution to the field. We will compare the effects of nucleofection on the cell viability and gene expression efficiency in K562 cells and $CD34^+$ cells.

2.3 Specific Aim 2

In this aim, we will compare microinjection and nucleofection for the delivery of HBB nucleases and donor template for gene modification in K562 cells. We will quantify the nuclease-mediated on- and off-target insertions and deletions (indels) in K562 cells injected with HBB TALENs and CRISPR/Cas9 nucleases using the microinjection system developed in Specific Aim 1. The indels in K562 cells microinjected with nucleases will be compared to nucleofected cells. We will apply single cell analysis to quantify and compare the types of allelic disruptions and indel spectra by the TALEN and CRISPR/Cas9 nucleases. In addition to investigating nuclease-induced indels, we will quantify the frequency of HDR-mediated gene modification in single K562 cells injected with HBB TALENs and CRISPR/Cas9 nucleases along with donor template DNA. To validate the microinjection delivery approach, the frequency of HDR-mediated gene modification measured will be compared to nucleofected cells. The completion of this aim will provide insight into the feasibility of applying microinjection for delivering precision gene-editing tools for gene correction of the SCD mutation as a potential therapeutic delivery approach.

2.4 Specific Aim 3

The genomic toxicity of nucleases is caused by its off-target cleavage activity and exposure in cells. Nucleases are typically delivered as plasmid DNA, which is inherently toxic in cells[116]. Because plasmid DNA can remain stable in cells for long periods of time, it is difficult to control the level of nuclease expression and duration of active nucleases in the cells. Alternatively, nucleases can be delivered as mRNA or purified proteins[138,147] to eliminate the toxicity associated with plasmid DNA and reduce the levels of off-target cleavage activity. In this aim, we will for the first time, as far as our knowledge, directly compare gene modification by RNA-guided nucleases delivered as plasmid DNA, mRNA, and ribonucleoprotein (RNP) complexes. We will separately optimize the delivery of HBB targeting sgRNA and Cas9 nucleofected as plasmid DNA, mRNA, and RNP complexes into K562 cells. We will identify the amount of reagents that provides the highest on-target activity, and subsequently compare the ratio of off-target to on-target indels to determine the strategy that provides the highest specificity. In addition, we will compare the cleavage activity by HBB CRISPR/Cas9 in $CD34^+$ cells nucleofected as plasmid DNA, mRNA, and RNPs. We will also compare cord blood to bone marrow $CD34^+$ cells to determine if the source of HSCs is a critical factor for nuclease activity. Furthermore, we will quantify the levels of HDR-mediated gene-targeting in $CD34^+$ cells nucleofected with HBB CRISPR/Cas9 as plasmid DNA, mRNA, and RNP complexes along with donor template DNA. In addition, we will use methylcellulose colony forming cell (CFC) assay to determine the effects of CRISPR/Cas9 and donor on the proliferation and myeloid differentiation potential of $CD34^+$ cells *in vitro*. Completion of this aim will provide insight into the optimized strategy for delivering gene editing reagents to obtain high specificity and HDR-mediated gene modification.

CHAPTER 3: CHACTERIZATION OF MICROINJECTION AND THE COMPARISON OF MICROINJECTION AND NUCLEOFECTION FOR GENE EDITING IN K562 CELLS

3.1 Introduction

Site-specific modification of endogenous genomic loci mediated by engineered nucleases has unprecedented potential for a wide array of applications, such as engineering model organisms[103,148-150] and developing new therapeutic strategies[151,152]. Examples of site-specific nuclease platforms include ZFNs, TALENs and CRISPR/Cas9 systems. DNA DSB induced by engineered nucleases can be repaired by the NHEJ or HDR pathways, leading to genome alterations, such as gene knockout or reconstitution at a desired target site[153]. HDR guided by exogenous donor template DNA having homologous sequences on both sides of the break site can be exploited for gene correction of mutations causing diseases, such as sickle cell anemia[152]. The potential benefits of nuclease-mediated HDR are the precise control of gene correction instead of uncontrollable random gene integration, and enhanced levels of gene correction compared to delivering homologous donor template DNA alone into cells. Recently, modification of the *HBB* locus was achieved using TALENs[114,115] and the CRISPR/Cas9 system[154], demonstrating the potential for a nuclease-based gene correction approach for treating sickle cell anemia.

One major challenge for advancing nuclease-based therapeutic strategies is to deliver optimal levels of nucleases and donor DNA into clinically relevant cell types[145]. Specifically, if the amount of nuclease-encoding plasmid and donor DNA is too low, the HDR rate would be insufficient to have a reasonable level of gene correction. On the

other hand, if the plasmid and donor DNA levels are too high, a large amount of cell death could occur due to cytotoxicity. Thus, it is critical to optimize the delivery protocol so that both a high level of nuclease activity and low level of cytotoxicity could be achieved. Although viral-based methods have been used for the delivery of gene targeting reagents into cell lines[155] and stem cells[113,156], there are many concerns, including random vector insertion, immunogenicity, integrity of packaged vectors, and effects of extensive stem cell culturing[145,155,157]. The most deleterious safety issue associated with viral-mediated delivery is potential activation of proto-oncogenes leading to tumors as a result of random vector insertion[157]. Transfection-based methods, such as nucleofection, has been used as a nuclease delivery method into primary cells[151,156], but cell loss due to cytotoxicity[156] remains an issue. Furthermore, with transfection it is difficult to control the amount of nucleases and donor template delivered into cells.

As an alternative, microinjection can be used for the direct delivery of nucleases and donor template into cells by penetrating the cell membrane using glass microcapillaries with fine tips, which has shown to successfully deliver macromolecules into human cells[158,159]. There have been multiple studies demonstrating successful microinjection-based delivery of nucleases into embryos for direct production of animal models with targeted mutations[160-162]. Although with low throughput, microinjection allows for precise control of the amount delivered into single cells, and can achieve high (~100%) delivery efficiency[146]. For cells that are difficult to transfect, microinjection is an attractive alternative and potentially less toxic and stressful to cells[146]. It may also provide a useful tool for characterization and optimization of other delivery methods, such as transfection.

Here we demonstrate microinjection based delivery of nucleases into human K562 cells with resulting gene modifications. We characterized a glass microcapillary injection system and quantified the effects of microinjection on cell viability and

proliferation. Microinjection of plasmids encoding a *HBB*-targeting TALEN pair and CRISPR/Cas9 system respectively into single K562 cells resulted in moderate to high levels of cleavage activities as quantified by the T7 Endonuclease I (T7E1) mutation detection assay and Sanger sequencing. We also quantified HDR in single K562 cells co-injected with *HBB*-targeting TALENs or CRISPR/Cas9 along with a GFP donor template, and determined the off-target cleavage in the *HBD* and glutamate receptor (*GRIN3A*) genes respectively. Our results suggest that microinjection of engineered nucleases, such as TALENs and CRISPR/Cas9 systems, lead to high levels of gene editing while minimally affecting the cell proliferation potential and viability. It is expected that the same microinjection based approach can be applied to primary cells that grow in suspension, including HSCs, and the successful demonstration of nuclease based gene editing in K562 cells using glass-needle microinjection may facilitate the development of high-throughput microinjection systems for genome editing in primary cells.

3.2 Results

Early methods tested for glass microcapillary-mediated microinjection

We tested several different approaches for performing microinjection. Successful injection was assessed by supplementing the injection solution with dextran conjugated to a fluorescent marker dye, such as FITC- and TRITC-dextran (10 kDa). The first microinjection system involved the application of two glass microcapillaries, one microcapillary for capturing single cells and a second one for injection **(Figure 6)**. The capture capillaries, prepared in house using a micropipette puller, had an outer tip diameter of 3-4 μm. We found that using two microcapillaries for injection was limited by: 1) low throughput injection into cells, 2) deformation of the cells after release from the capture microcapillary **(Figure 6c)**, 3) fouling of the capture microcapillary that caused

cells to adhere tightly, making it difficult to release cells after injection, and 4) cells with successful injection were mixed with non-injected cells. On average we injected roughly 16 single K562 cells within an hour using the two microcapillary injection approach. After 6-7 injections, we found that the capture capillary could not release cells due to fouling, making it necessary to change the capture microcapillary multiple times during a single injection session.

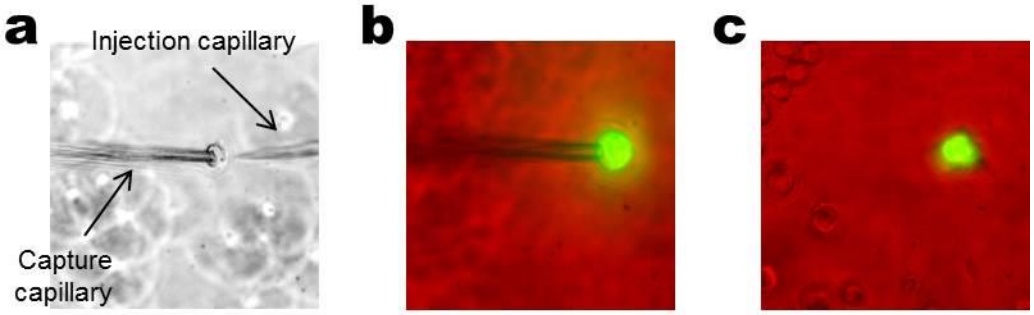


Figure 6: K562 cells microinjected using two glass microcapillaries

(a) Phase contrast image of a single K562 cell captured with a glass microcapillary and simultaneously injected with a second microcapillary. Merged phase contrast and fluorescent images of **(b)** single cells successfully injected with FITC-dextran and **(c)** after release from the capture microcapillary.

We attempted to address the challenges associated with releasing and isolating successfully injected cells by utilizing a cell capture and release microfluidic device **(Figure 7)**. The device was designed to assemble to an inverted Delta vision microscopy system and consisted of a microinjection chamber for injecting cells using two microcapillaries **(Figure 7a).** The device also contained a built in microfluidic channel to facilitate the movement and isolation of injected cells into a viewing chamber embedded in the microinjection chamber (**Figure 7b)**. There was also a release port connected to the chamber designed to enable the isolation of cells through an external channel

assembled to the device. We hypothesized that the flow of buffer through the microfluidic channel would allow for cells to become released from the capture microcapillary more efficiently **(Figure 8a)**. We also hypothesized that the viewing chamber would enable monitoring of injected cells for studies on the effects of injection. Testing of the microfluidic device revealed that the flow of buffer through the microfluidic channel did not drastically improve the efficiency of releasing cells from the capture microcapillary. With the device, we could only inject up to 10 K562 cells before having to replace the capture microcapillary due to fouling. We also found that it was difficult to isolate the cells into the viewing chamber without inflicting cell damage. Many of the cells that we attempted to isolate into the viewing chamber became wedged between the filters **(Figure 8b)** that were designed to prevent cells from flowing out of the device through the outlet port. During testing, we injected the cells with TRITC-dextran and pmaxGFP plasmid DNA. In the first experiment, a flow rate of 1 μL min^{-1} applied to capture cells resulted in only 20% (2 out of 10) of the injected cells captured into the viewing chamber. The remaining 80% of the injected cells (8 out of 10) were trapped between the filters. Contamination prevented quantification of the percentage of cells with gene expression. In the second experiment, a flow rate of 2 μL min^{-1} resulted in 60% (3 out of 5) of injected cells captured into the viewing chamber, which was a higher capture efficiency compared to the first experiment. At 24 hours after injection, all three injected cells were positive for GFP expression **(Figure 8c)**. When we applied a flow rate below 1 μL min^{-1}, we were unable to trap any cells into the viewing region. Thus, the efficiency for successfully isolating cells into the viewing chamber after release from the capture capillary depended on the flow rate of buffer in the microfluidic channel. The low throughput injection and cell damage issues were not resolvable.

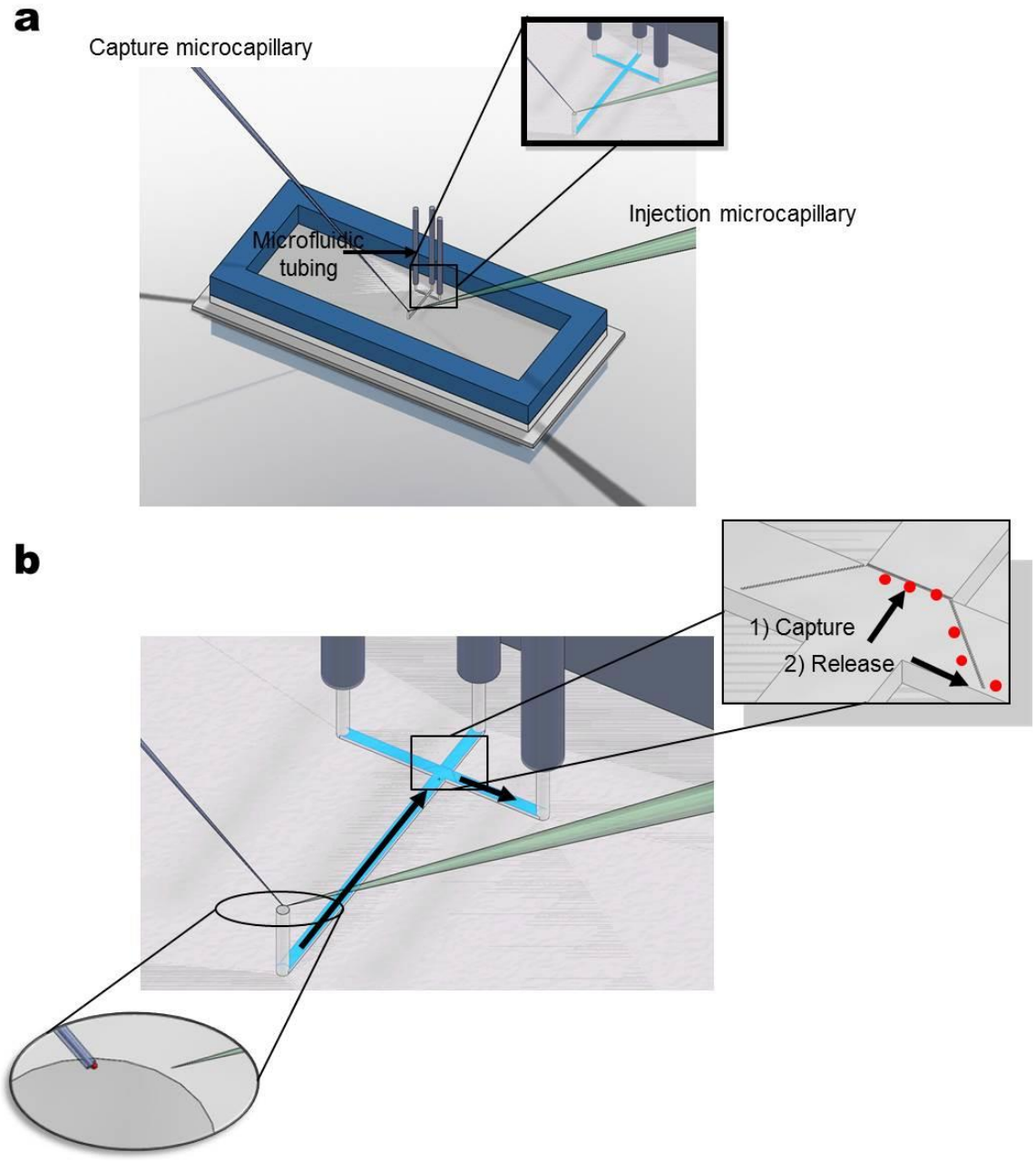


Figure 7: Schematic of cell release and capture microfluidic device

The microfluidic device assembles with an inverted microscopy system to facilitate the isolation and monitoring of cells injected using two microcapillaries. **(a)** Exterior view of the microinjection chamber and interior view of the microfluidic channel that captures cells after injection. **(b)** Microfluidic channel that connects to the viewing chamber for monitoring injected cells and release port for isolating the injected cells. The drawings were prepared by Dr. David Myers at the Georgia Institute of Technology.

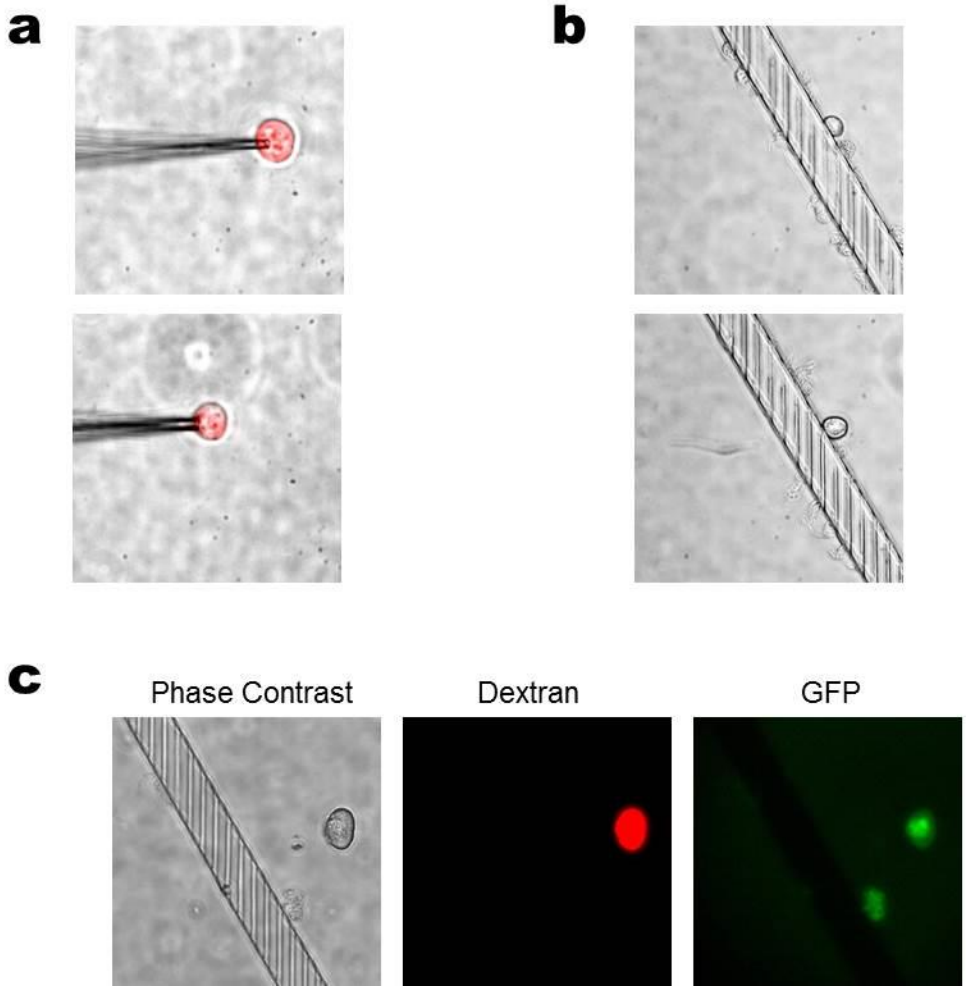


Figure 8: K562 cells monitored after injection using two microcapillaries

Cells were injected with solution supplemented with 100 uM TRITC-dextran and 2 ng uL^{-1} pmaxGFP plasmid DNA. **(a)** Merged phase contrast and fluorescent microscopy images of 2 different cells after injection using the two microcapillary method. **(b)** Two separate phase contrast images of cells after release from the capture microcapillary and found wedged between filters. **(c)** Phase contrast and fluorescent images of a cell injected at 24 hours after injection. The injected cell was positive for dextran and GFP fluorescence.

The third approach tested involved capturing the cells using a microfluidic cell holder device to facilitate microinjection using a single glass microcapillary. The major benefit of using one microcapillary for injection is it eliminates the issues associated with

using a capture microcapillary, including microcapillary fouling and low efficiency for releasing cells after injection, and the approach potentially increases the throughput of injection. The microfluidic device was designed with C-shaped holders to trap cells using the flow of media through a microfluidic channel. In the first experiment, roughly 2.4 x 10^5 K562 cells were deposited on one side of the chamber and a flow rate of 400 μL min^{-1} was applied for 10 minutes to trap cells into the holders. Only 21% of the holders were occupied by a cell **(Figure 9)**. Adjusting the flow rate did not improve the efficiency of trapping cells. One contributor to the low efficiency for trapping is the process of initiating and maintaining flow in the microfluidic channel. Manual generation of flow through the outlet was necessary due to clogging of cells at the outlet/inlet ports, reducing consistent flow through the chamber. Another challenge was microinjecting the cells trapped in the holder without causing damage to the injection microcapillary.

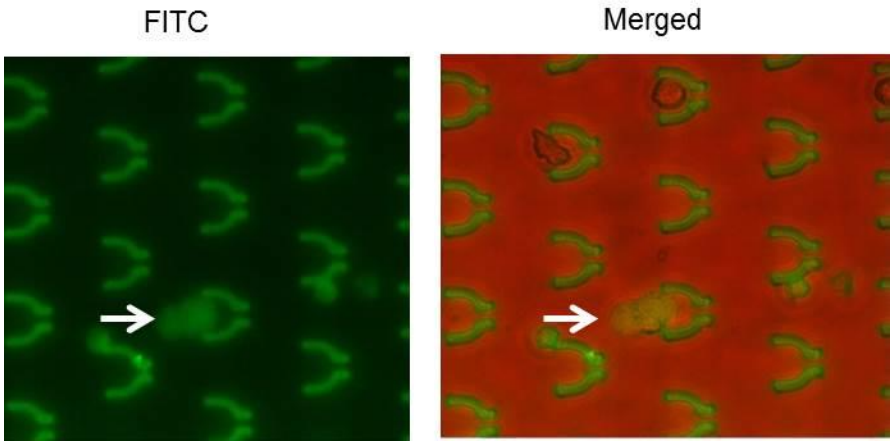


Figure 9: K562 cells in cell holder microfluidic device

Fluorescent and merged phase contrast and fluorescent images of K562 cells trapped in cell holders. One cell was injected with FITC-dextran, shown by white arrow.

Microinjection on retronectin-coated polystyrene dishes

The fourth approach tested for microinjection into K562 cells involved the immobilizing of cells on polystyrene surfaces coated with retronectin **(Figure 10a)**. The success of injection was assessed by supplementing the injection solution with FITC-dextran as a fluorescent marker. Retronectin is recombinant human fibronectin consisting of a CS-1 site and RGDS domain that interact with the $\alpha_4\beta_1$ and $\alpha_5\beta_1$ integrins expressed on K562 cell membranes[163]. Injections were performed at ambient conditions at a rate of ~4 cells per minute with injection sessions limited to 40 minutes per plate to ensure high cell viability. Cells adhering to retronectin-coated surfaces had a flattened, spherical morphology and remained firmly bound following injection **(Figure 10a)**. We evaluated the effects of the cell-matrix interaction on cell viability by seeding and detaching cells on polystyrene surfaces coated with solutions containing different initial retronectin concentrations. Consistent with a previous study[159], detachment of cells from retronectin surfaces after incubation for 1.5 hours and attachment did not significantly affect the cell viability, even at increasing concentrations of retronectin **(Figure 10b)**. However, the retronectin concentration had an effect on the attachment of cells under the force applied by the glass microcapillary during injection **(Figure 11)**. We found that the highest level of successful injection (82%) was achieved with an initial retronectin concentration of 50 $\mu g\ mL^{-1}$ or greater in the coating solution **(Figure 11)**. Cells immobilized on retronectin-coated surface had a cell viability of ~63-72% after injection **(Figure 11)**, which was estimated by the percentage of cells that retained the dextran fluorescence and maintained membrane integrity after a 2 hour incubation period following injection. We applied the retronectin-mediated microinjection approach for all of the following experiments.

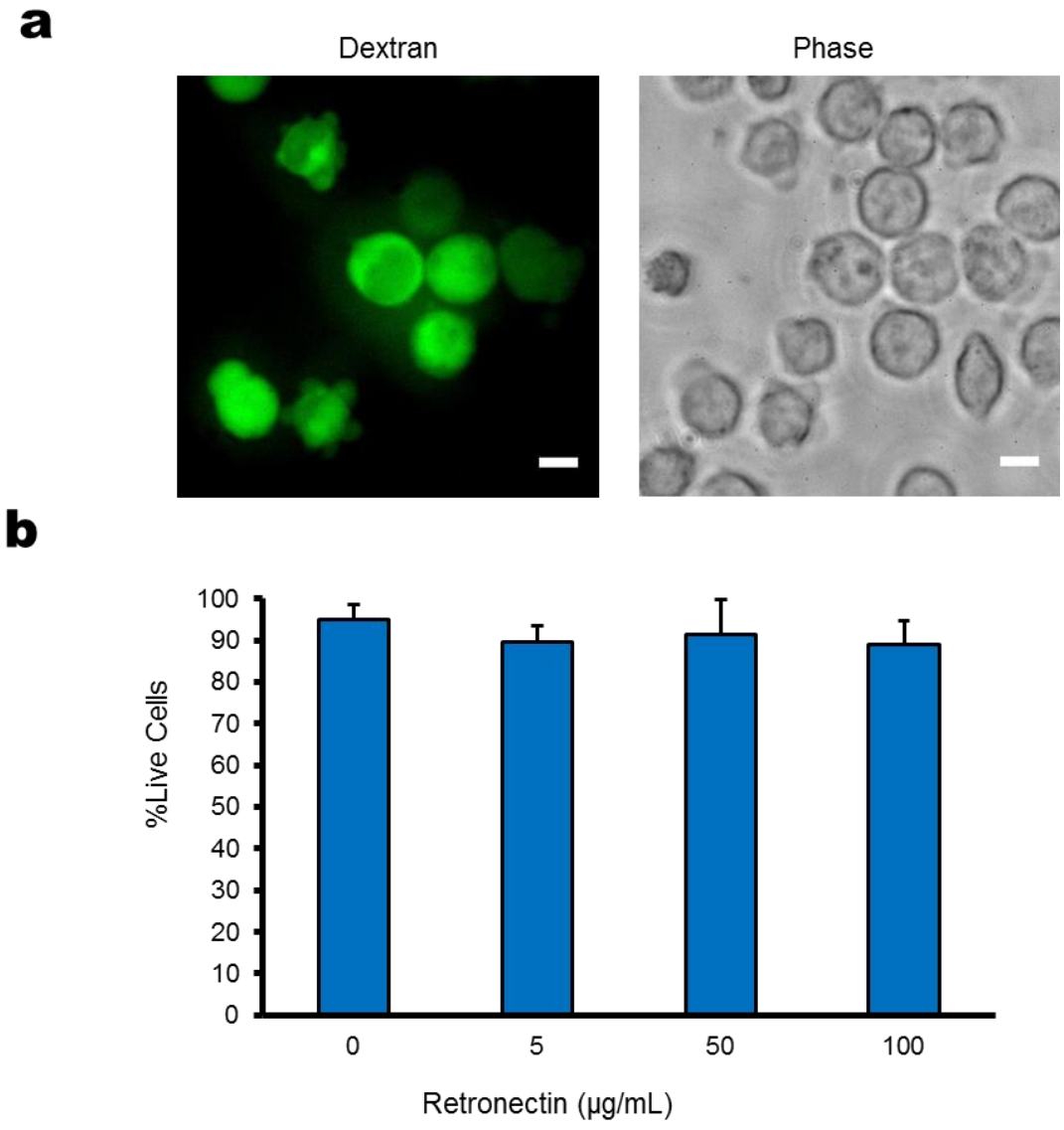

a
Dextran
Phase
b
%Live Cells
100
90
80
70
60
50
40
30
20
10
0
0
5
50
100
Retronectin (μg/mL)
c
Dextran
pmaxGFP

Figure 10: Microinjected K562 cells on retronectin coated dishes

(a) Fluorescence and phase contrast microscopy images of successfully injected cells immediately following injection. **(b)** Cell viability was not affected by retronectin, as shown by the percentage of live cells detached from dishes coated with solutions containing different Retronectin concentrations. For each sample, cells (3 X 10^4) attached to coated dishes were detached using pipetting. Bars represent statistical mean for 3 replicates ± standard deviation. **(c)** Fluorescence microscopy images of K562 cells injected with dextran-TRITC and pmaxGFP plasmid at 24 hours after injection. Cells with successful injections (red) were analyzed for GFP expression (green). Scale bar width corresponds to 10 µm.

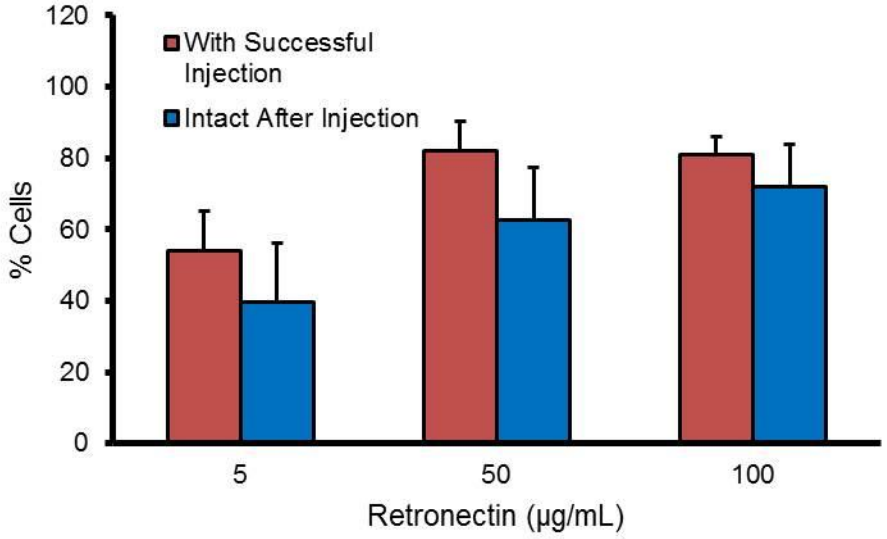


Figure 11: Microinjection efficiency and cell viability after injection

Percentage of successfully injected (red bars) and intact cells (blue bars) on surfaces coated with solutions containing different concentrations of retronectin. Intact cells are defined as retaining dextran fluorescence 2 hours after injection. Each experiment consisted of 30-40 cells injected at each initial coating concentration. The experiments were repeated 3 times. Error bars represent standard deviation.

Injection parameter optimization

A clear advantage of microinjection is that the injection volume is precisely controlled[146]. We used a FemtoJet (Eppendorf) microinjector to precisely control the volume released by adjusting the applied injection pressure and duration of injection. To generate a standard curve correlating injection pressure and volume, we injected dye-labeled dextran into mineral oil using different injection pressures and fixed injection time of 0.1 seconds and measured the diameter of the injected spheres using a micrometer[164]. The results obtained **(Figure 12a)** are consistent with the volume vs pressure measurements reported in a study[165] that used a different approach for estimating the injection volume. To determine the optimal injection pressure, cells immobilized on a retronectin-coated surface were co-injected with TRITC-dextran, a fluorescence reporter for injection, and pmaxGFP plasmid DNA with different injection pressures ranging from 30 to 120 hPa (corresponding to injection volumes between 3 to 30 pL). At 24 hours after injection **(Figure 10c)**, signal from dye-labeled dextran was used to identify the injected cells[166] and the percentage of injected cells that exhibited GFP fluorescence was quantified[167] **(Figure 12b)**. We found that an injection volume of 7 pL gave the highest amount of GFP-expressing cells **(Figure 12b)**. Injection volumes greater than 30 pL resulted in noticeable cell damage and cell death (data not shown). Thus, for all subsequent microinjection experiments an injection volume of 7 pL was used.

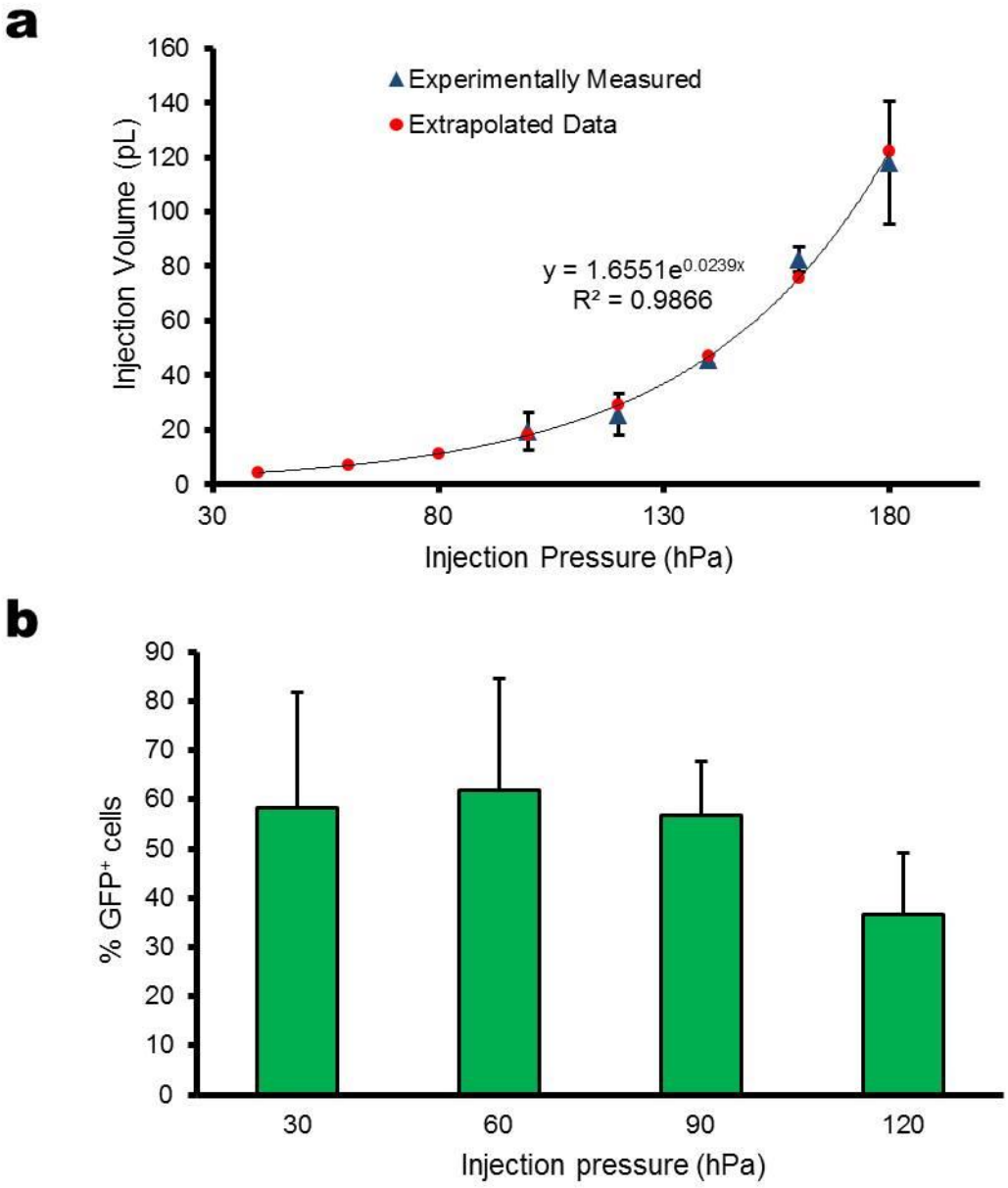


Figure 12: The effects of injection pressure on injection volume and gene expression

(a) Plot of injection volume for different injection pressure settings. The injection volume was estimated experimentally by performing injections of dye labeled dextran into a droplet of mineral oil. The injected sphere was estimated using the equation $\frac{4}{3}r^3\pi$ for pressure settings between 100-180 hPa, where r is the radius of the injected solution measured using a micrometer. The equation for the best fit exponential curve was used to estimate the volume for injection pressure settings below 100 hPa. **(b)** Plot of percentage of injected cells with observable GFP expression at 24 hours after injection with different pressure settings.

Bars represent mean percentage of GFP expressing cells out of 35-56 cells injected ± standard deviation for 4 independent experiments.

The effects of microinjection on cell doubling time

The effect of glass microcapillary injection on cell doubling time was evaluated. Cells successfully injected with FITC-dextran or nucleofected with pmaxGFP plasmid (as positive control) at 24 hours after delivery were deposited into Terasaki MicroWell plates at 1 cell per well using fluorescence activated cell sorting (FACS) **(Figure 13)** and subsequently cultured for two days. Non-treated cells detached from retronectin-coated or uncoated plates were used as controls. The cell number in each well was counted at 24 and 48-hours after sorting and the cell doubling time was quantified. We estimated the doubling time for control cells to be 20-22 hours, which is similar to the doubling rate reported in the literature[168]. Consistent with that shown previously[159], microinjection had no adverse effect on the cell doubling time **(Figure 14)**, and nucleofection did not have much effect on the cell doubling rate either.

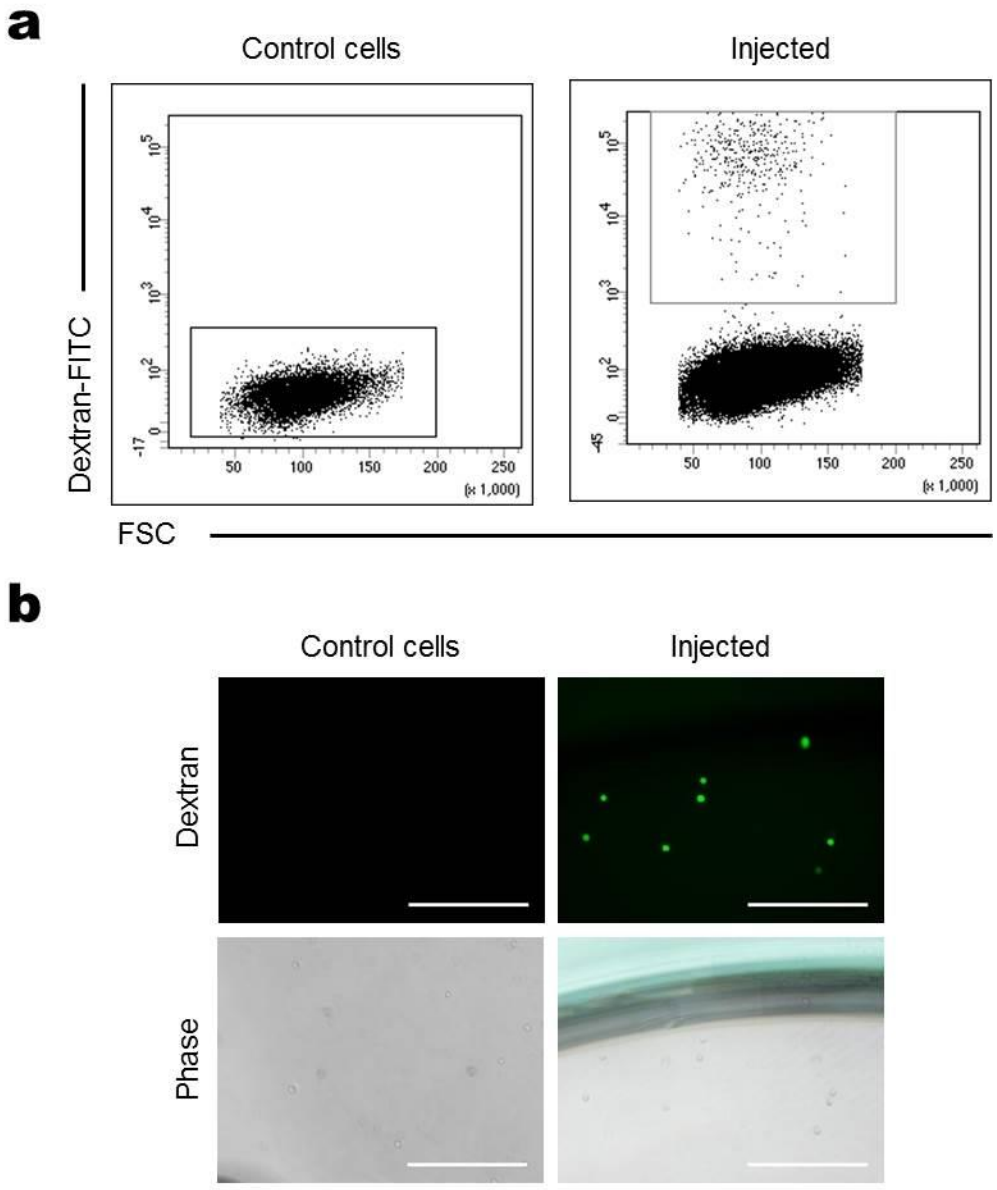


Figure 13: Separation of microinjected K562 cells using FACS

Cells injected with dextran-FITC were detached from retronectin coated dishes after injection and subjected to fluorescence activated cell sorting. **(a)** Panels show stable dextran fluorescence in cells gated for viability. For clarity, the control and injected K562 cells are shown in black and gray boxes respectively. **(b)** Fluorescence and phase microscopy images of injected cells after FACS. Scale bar width corresponds to 500 µm.

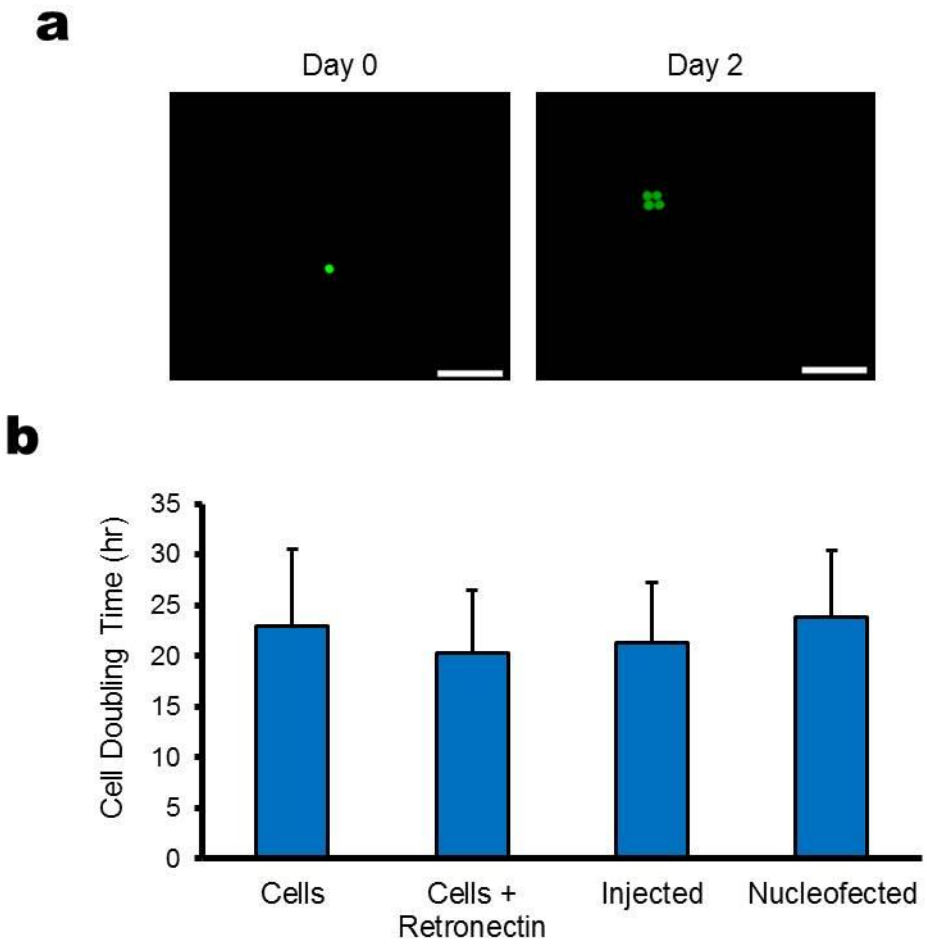


Figure 14: K562 cell doubling time after microinjection and nucleofection

Viable K562 cells were deposited as single cells into Terasaki MicroWell plates using FACS at 24 hours after injection of dextran-FITC or nucleofection of pmaxGFP. **(a)** Fluorescence microscopy images of injected cells at 24 hours and 48 hours after FACS**. (b)** Plot of the cell doubling time for viable control and treated cells. The doubling time was calculated using the equation: duration times log (2) divided by the cell growth in 24 hours. As controls, we show the cell doubling time for viable untreated cells in suspension and detached from retronectin coated plates. One-way ANOVA indicates absence of significant difference between different conditions ($p > \alpha$, $\alpha = 0.05$). Scale bar width corresponds to 100 µm. Bars represent mean cell doubling time ± standard deviation (n = 3).

Controlled delivery of *HBB*-targeting nucleases using microinjection

Microinjection enables simultaneous delivery of multiple genome editing reagents, including nucleases and donor template DNA, into cells. An added advantage of microinjection over other delivery methods is the precise control of the amount delivered into cells. To demonstrate the ability of microinjection based delivery for genome editing, K562 cells were injected with plasmids expressing TALENs or CRISPR/cas9 respectively, and the nuclease on- and off-target activities were quantified. The TALEN pair L4 (left TALEN) and R4 (right TALEN) used in this study was designed to target the *HBB* gene[114], the target sites shown in **Figure 15a**. The L4-TALEN overlaps the sickle mutation in codon 6 of the *HBB* locus, separated by a 15-base spacer from the R4-TALEN binding site **(Figure 15a)**. The CRISPR RNA R02, a 20-base guide sequence, was designed to target *HBB* as well[154], near the sickle mutation **(Figure 15a)** adjacent to a PAM sequence containing the trinucleotide NGG. To label injected cells, in addition to plasmids encoding TALENs or CRISPR/Cas9, K562 cells were co-injected with FITC-dextran as a fluorescence marker.

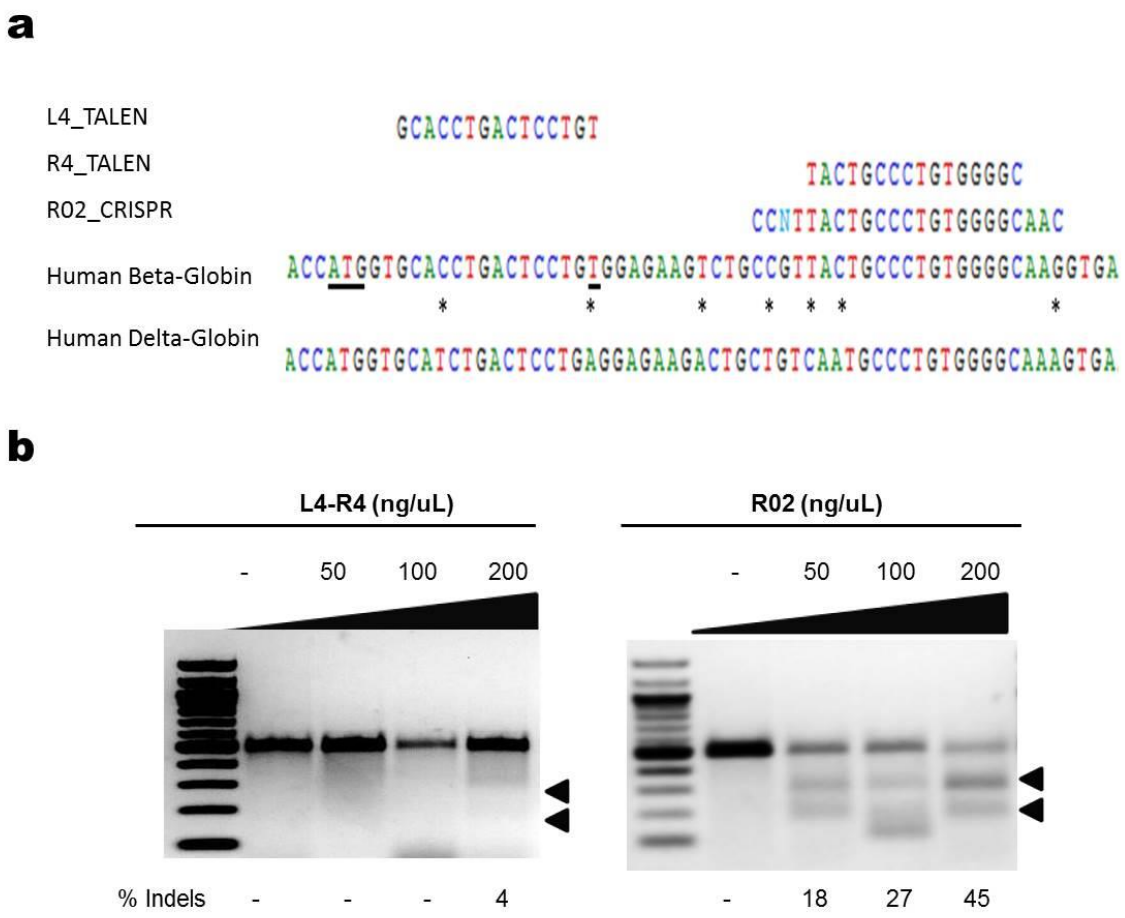


Figure 15: Gene editing by *HBB*-targeting nucleases using microinjection

(a) Schematic showing targeting sequences of L4-R4 TALEN pair and R02 CRISPR guide RNA aligned to the *HBB* and *HBD* loci. The CRISPR guide RNA is shown complementary to the reverse stand and is listed to the right of the PAM sequence. The ATG start codon and the sickle cell mutation are underlined. Asterisks between *HBB* and *HBD* indicate mismatches. The A, T, C, and G nucleotides are shown in green, red, blue, and black respectively for clarity. **(b)** Nuclease-induced indel rate as a function of plasmid concentration. Plasmids encoding L4-R4 TALENs or R02 CRISPR/Cas9 were microinjected into K562 cells with an injection volume of 7 pL and the nuclease-induced cleavage at the HBB locus was analyzed using the T7E1 assay. Shown is a comparison of the indel rates by L4-R4 TALENs and R02 CRISPR/Cas9 system at plasmid concentrations of 50, 100 and 200 ng/µL.

Successfully injected cells were deposited into 96-well plates with 1 cell per well on average using FACS. The clonal colonies derived from the single microinjected cells after 14-16 days of culturing were pooled together. T7E1 mutation detection assay **(Figure 16)** was performed to quantify the rate of cleavage-induced indels. We found that the on-target cleavage rate is dose-dependent and, for the L4-R4 TALEN pair tested, the indel rate was 4% at a concentration of 200 ng μL^{-1} total TALEN plasmid, while no measurable activity at concentrations of 50 and 100 ng μL^{-1} was observed **(Figure 15b).** In contrast, for the CRISPR/Cas9 system tested, much higher indel rates were obtained **(Figure 15b)**. Specifically, with plasmid encoding R02 CRISPR/Cas9, indel rates of 18%, 27%, 45% were obtained at plasmid concentrations of 50, 100 and 200 ng μL^{-1}.

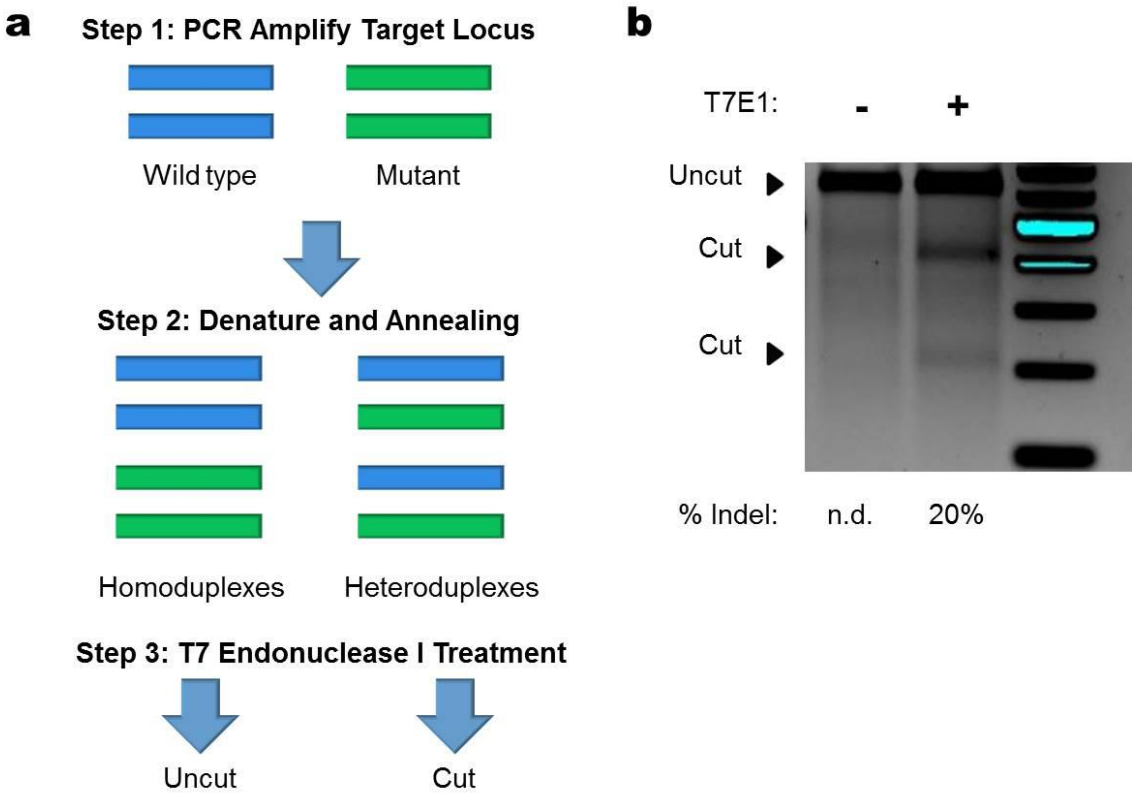


Figure 16: T7E1 assay for detecting nuclease induced mutations

(a) The target locus is amplified using primers. Resulting in amplicons containing the wild type sequence or indels induced by NHEJ repair of the nuclease induced DSB. The PCR product is subjected to denaturing and annealing to form random duplexes. The addition of T7E1 enzyme will cut the duplexes containing wild type and mutant sequences. **(b)** Gel of PCR amplicons treated with T7E1 enzyme. The addition of the enzyme resulted in the formation of two smaller sized, cut bands.

To benchmark the cleavage activity measured in the microinjection studies, we compared the on- and off-target activity in K562 cells nucleofected with L4-R4 TALENs. Cells were nucleofected with plasmids encoding L4-R4 TALENs using a 4D-nucleofector system (Lonza) and cultured for 3-days following nucleofection. The T7E1 assay was performed to measure the L4-R4 TALEN induced indels in the *HBB* locus of bulk nucleofected and microinjected cells **(Figure 17)**. Off-target activity was measured in the *HBD* locus, which has a sequence similar to the *HBB* gene **(Figure 15a)**. Interestingly, the mean cleavage activity in microinjected cells was slightly higher compared to nucleofected cells, but the difference was not significant **(Figure 17).** This indicates that the L4-R4 TALENs expressed in cells following microinjection are highly active, providing further evidence that microinjection works well for delivering genome editing reagents into human somatic cells.

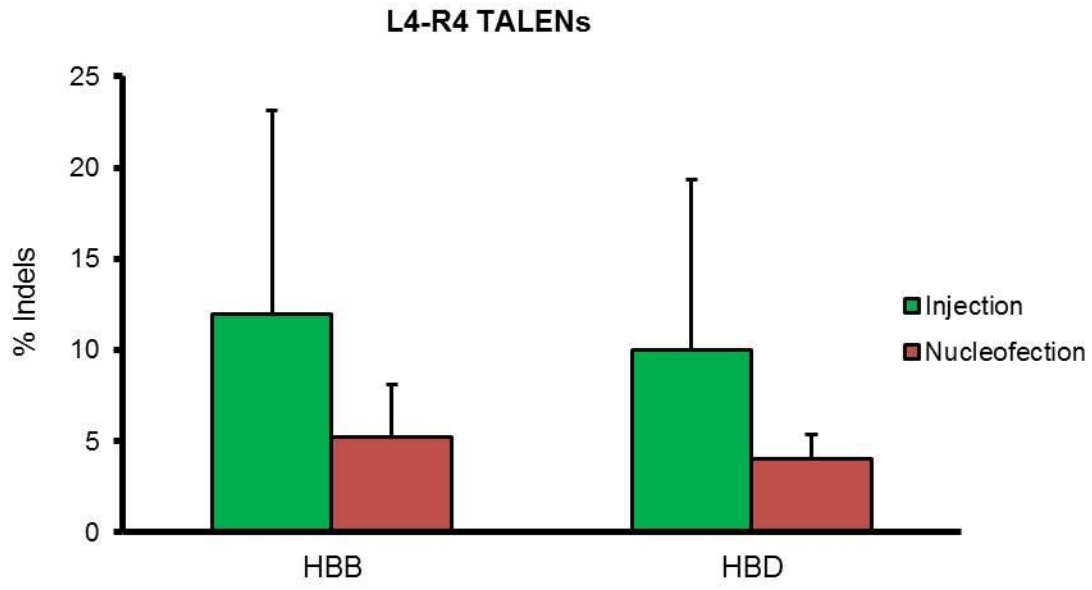


Figure 17: Comparison of indel frequency induced by the L4-R4 TALEN pair delivered using microinjection and nucleofection

Green and red bars represent mean percent indels in microinjected and nucleofected cells respectively. The indels shown for microinjected cells represent the average of 58 single cell clones pooled together.

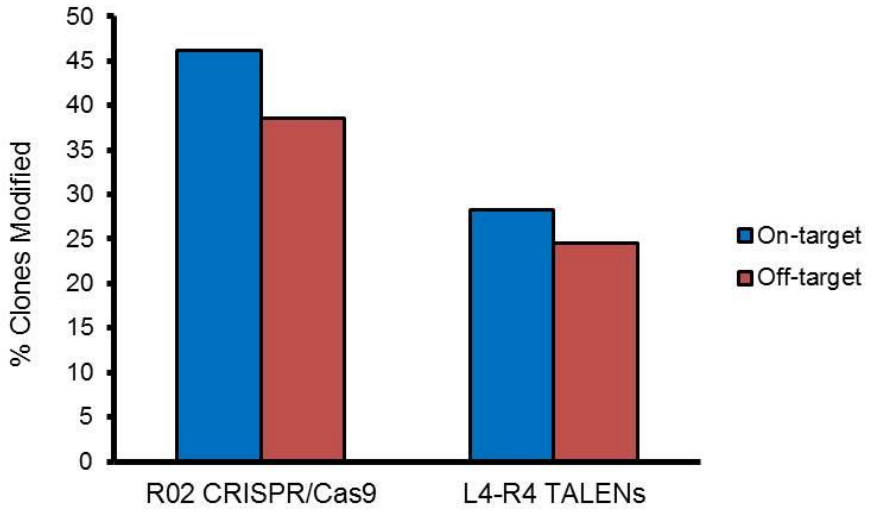
Single cell analysis of on- and off-target indels

We performed single-cell analysis of indels in cells microinjected with R02 CRISPR/Cas9 system and L4-R4 TALENs respectively. Clonal colonies were generated from K562 cells microinjected with nucleases by depositing injected single cells into multi-well plates (one cell per well) using FACS, followed by culturing for 14 to 16-days. On- and off-target activities of clones derived from single microinjected cells were measured using the T7E1 assay and the percentage of clones with indels was determined. For the R02 CRISPR/Cas9 system, off-target cleavage was evaluated at the *GRIN3A* locus, which was shown to have a high level of off-target cleavage[154]. We found that with the R02 CRISPR/Cas9 system, 36 out of 78 clones (46%) had the *HBB* locus modified while 30 clones (38.5%) showed indels in *GRIN3A* **(Figure 18a)**. In contrast, for

the 53 clones derived from single cells injected with L4-R4 TALENs, 28% on-target (*HBB*) and 24.5% off-target (*HBD*) indels were found **(Figure 18a)**. Compared with bulk measurements, these results may give a more accurate quantitation of the cleavage efficiencies of CRISPR/Cas9 and TALENs, since the amount of nuclease-encoding plasmids delivered is more uniform among the cells.

To gain additional insight into nuclease-induced DNA cleavage, the clones having measurable indel rates at *HBB* were further analyzed using Sanger sequencing. PCR primers used to amplify the *HBB* gene for sequencing were designed with a 5' 4-base barcode **(Supplementary Table 3)** so that each clone has a unique sequence identifier. The percentage of clones that had 1, 2, 3 or > 3 mutations in the *HBB* gene was quantified. We found that, of the 36 clones having R02 CRISPR/Cas9 induced cleavage at *HBB*, 22 (61.1%) had 3 *HBB* mutations **(Figure 18b)**, indicating a high level of Cas9 activity with *HBB* cleaved in all three copies of chromosome 11 **(Figure 18a)**. Given that K562 cells have trisomy in chromosome 11[169], clones having more than 3 *HBB* mutations is likely a result of cleavage activity by the CRISPR/Cas9 system in daughter cells following a cell division. Consistent with results shown in Figures 2b and 3a, L4-R4 TALENs had lower cleavage activity compared to the R02 CRISPR/Cas9 system, with only 20% of the clones having 3 *HBB* mutations **(Figure 18b)**.

a



b

HBB Analysis	R02	L4-R4
Total modified clones analyzed by sequencing	36	15
% Clones with 1 mutation	13.9	40.0
% Clones with 2 mutations	19.4	40.0
% Clones with 3 mutations	61.1	20.0
% Clones with > 3 mutations	5.6	0

Figure 18: On- and off-target indels in single K562 cells microinjected with *HBB*-targeting nucleases

Cells microinjected with 200 ng/μL plasmids encoding L4-R4 TALENs or R02 CRISPR/Cas9 were expanded from single cells into clonal colonies in a 14 to16-day culture. T7E1 assay and Sanger sequencing were performed in individual clones and the number of clones having measurable indels was determined. **(a)** The percentages of clones with measurable on- and off-target indels. The off-target indels were detected in the *GRIN3A* and *HBD* loci for R02 CRISPR/Cas9 and L4-R4 TALENs respectively. N = 78 (R02) and 53 (L4-R4). **(b)** The percentage of clones having specific numbers of *HBB* mutations detected from 8-24 total sequencing reads.

We further analyzed the indel spectra induced by R02 CRISPR/Cas9 and L4-R4 TALENs respectively. We found that cells injected with TALENs had a broad spectrum of

indels, including peaks of 21-base and 4-base deletions and 6-base insertions **(Figure 19)**. Similarly, the CRISPR/Cas9 system induced a broad spectrum of indels as well, with sharp peaks of 9-base deletions and a 1-base insertion **(Figure 19).** The observation that 9-base deletion is the most frequent indel in cells injected with R02 CRISPR/Cas9 is consistent with that for cells subjected to nucleofection (data not shown). The difference in indel spectra between CRISPR/Cas9 and TALENs is likely due to the specific DNA cleavage induced (blunt ends vs. 4-base overhang) and the corresponding repair mechanism.

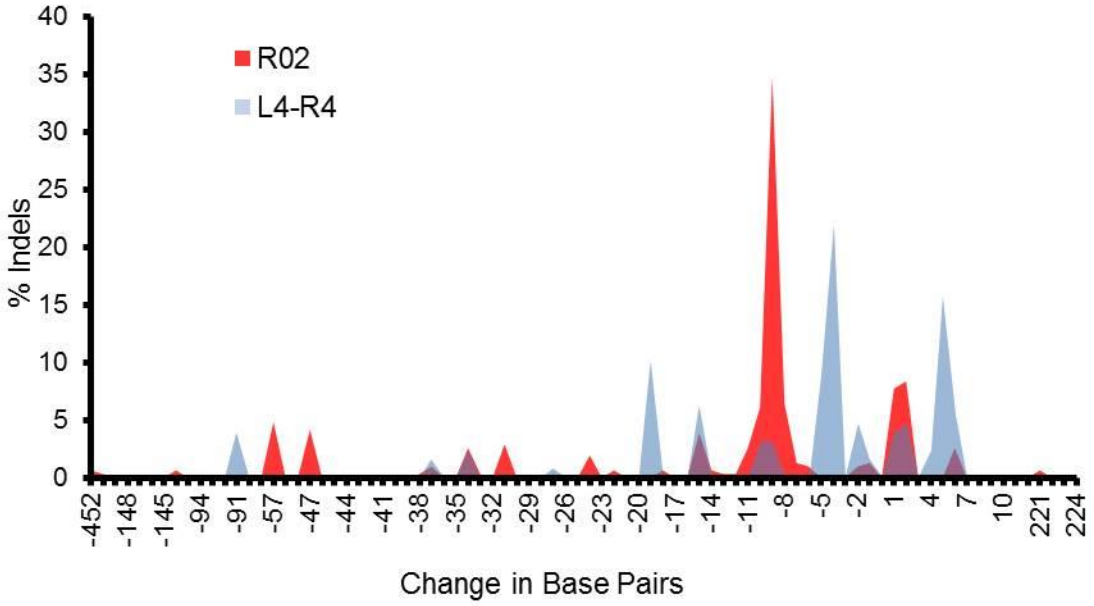


Figure 19: Comparison of indel spectra for cells microinjected with L4-R4 TALENs and R02 CRISPR/Cas9

The indels were analyzed in sequences obtained from single cells injected with nucleases. The change in the number of base pairs resulting from NHEJ repair of DNA cleavage in the *HBB* locus was compiled for each sequence read. The y-axis represents the percentage of indels with specified number of base pair changes.

Quantifying nuclease induced gene modification in microinjected cells

We also investigated HDR-mediated gene modification efficiency for the L4-R4 TALENs and R02 CRISPR/Cas9 system in cells microinjected using glass microcapillaries. The donor template for targeted gene insertion was designed with a GFP expression cassette under the Ubc promoter and flanked by approximately 1 kb arms of homology from the HBB locus[114] **(Figure 20)**. Primers listed in **Supplementary Table 2** were designed to exclusively amplify the integrated GFP sequence at the HBB locus (data not shown).

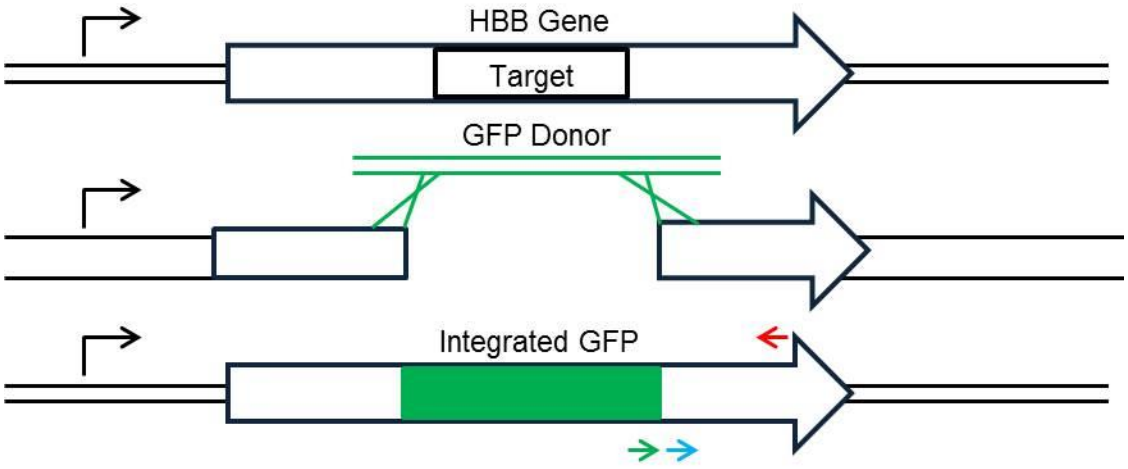


Figure 20: Diagram of GFP reporter system used to detect HDR-mediated gene modification

Nuclease cleavage and resection yields a substrate for HDR which may involve the use of exogenous β-Ubc-GFP donor, flanked by 5' and 3' homologous fragments of the *HBB* sequence (middle and bottom). The dashed lines indicate the HDR process with the donor template. Gene insertion was confirmed by PCR using primers specific for integrated GFP at the target locus, as shown by the green and red arrows (bottom). *HBB* (control) was amplified using primers shown by the blue and red arrows, which bound downstream the 3' homologous region in the HBB locus.

We co-injected the β-Ubc-GFP donor template with L4-R4 TALENs or R02 CRISPR/Cas9 into cells, isolated injected cells using FACS and formed single cell

colonies in a 14 to 16-day culture. Clones were evaluated for GFP expression using fluorescence microscopy. We found that many clones from cells co-injected with *HBB*-targeting nucleases and β-Ubc-GFP donor were positive for GFP fluorescence, but not the case for clones formed with cells injected with donor only **(Figure 21a).** To confirm gene insertion, we extracted the genomic DNA from cells showing GFP fluorescence, and PCR amplified the GFP sequence specifically integrated at the *HBB* site **(Figure 21b)**. We defined clones with HDR-mediated gene modification as having both GFP fluorescence and positive PCR results. To quantify the efficiency of HDR, we determined the number of clones with gene modification, divided by the total number of clones derived from cells injected with both nuclease and donor. This method of quantifying the HDR rate is more rigorous compared to flow cytometry analysis of GFP fluorescence alone[114] because of the additional requirement for PCR detection of GFP integration at the HBB gene locus. Of the 38 clones derived from cells co-injected with R02 CRISPR/Cas9 and donor template, we observed 4 clones (10.5%) with PCR-confirmed gene modification **(Figure 21b)**. In contrast, for clones derived from cells co-injected with L4-R4 TALENs and donor template, only 1.6% had PCR-confirmed GFP gene modification **(Figure 21b)**.

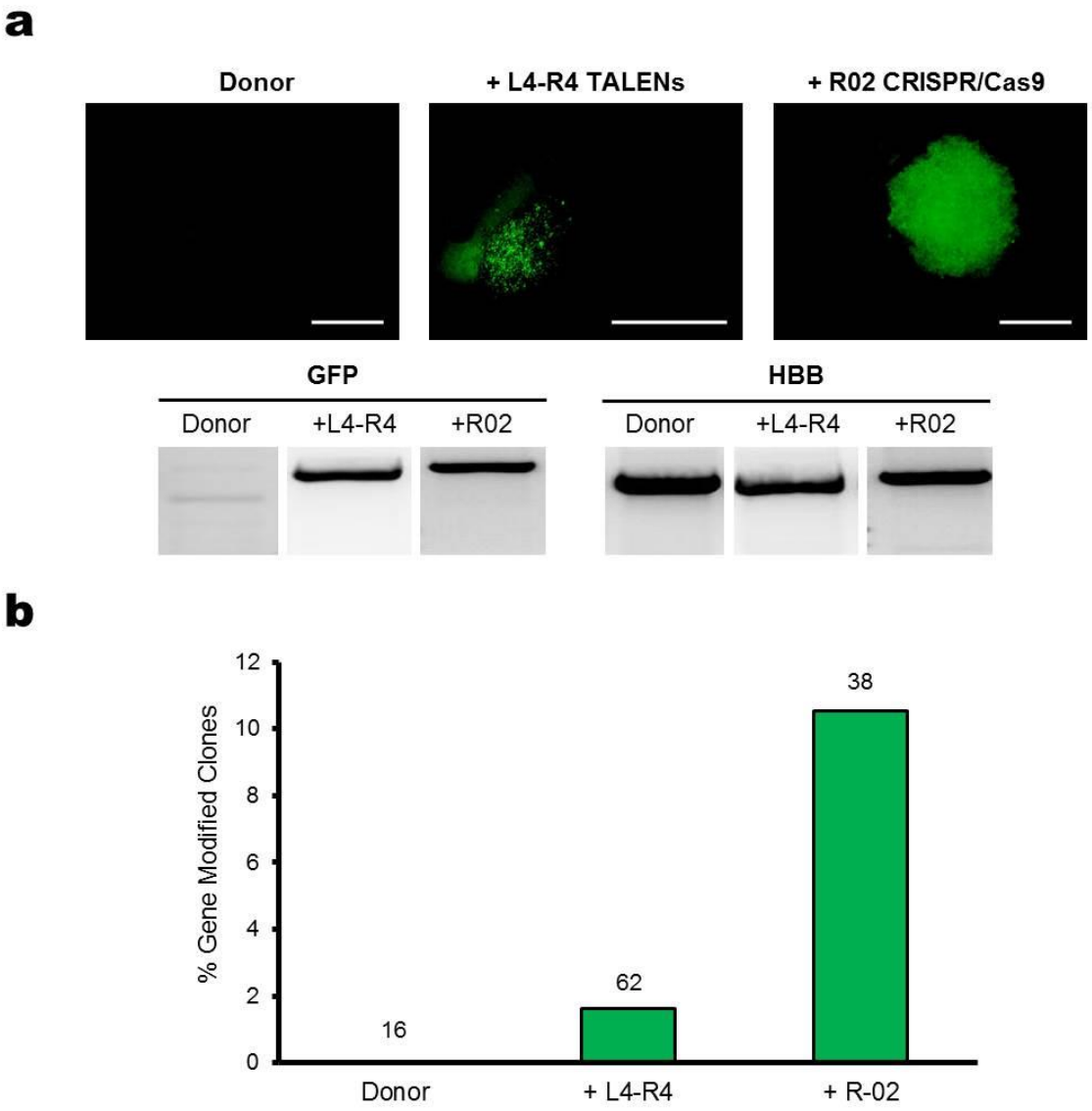


Figure 21: HDR-mediated gene modification in microinjected cells

(a) Fluorescence microscopy images of clones derived from single cells microinjected with β-Ubc donor with or without L4-R4 TALENs or R02 CRISPR/Cas9. Bottom of images are PCR results of integrated GFP or *HBB* (control) for clones. Scale bar corresponds to 500 µm. **(b)** The percentage of clones with HDR-mediated gene modification confirmed by both PCR and fluorescence microscopy. The number of single cell clones analyzed is shown above each bar.

To quantify and compare the frequency of HDR-mediated gene modification using microinjection and nucleofection respectively, cells were nucleofected with the β-Ubc-GFP donor together with *HBB* L4-R4 TALENs or R02 CRISPR/Cas9 and cultured in bulk for up to 21 days. Cells were then analyzed using flow cytometry to obtain the percentage of cells having gene-insertion induced GFP fluorescence, and the results were normalized using that at Day 3. We found that for cells nucleofected with R02 CRISPR/Cas9 plus donor template, there was an 11-fold increase in the normalized amount of GFP-positive cells compared to cells with donor only **(Figure 22a)**. In contrast, there was only a 4-fold increase for cells nucleofected with L4-R4 TALENs plus donor **(Figure 22b)**. The higher amount of GFP-positive cells due to CRISPR/Cas9 induced gene insertion compared to that of TALENs is consistent with the microinjection results **(Figure 21b)**.

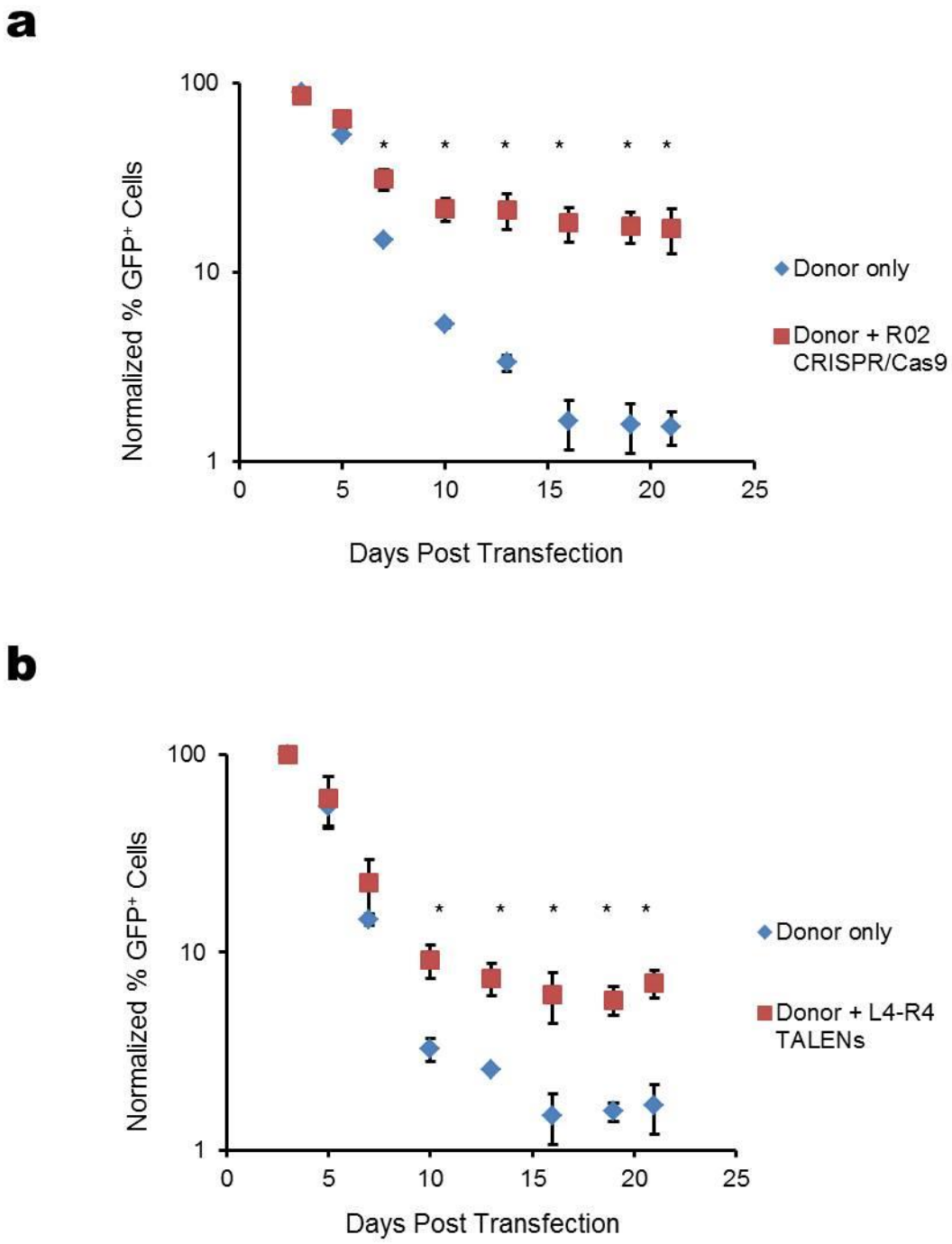


Figure 22: Gene insertion in nucleofected cells

The percentage of GFP-positive cells was quantified using flow cytometry during a 21-day culture after nucleofection with β-Ubc-GFP donor with and without **(a)** R02 CRISPR/Cas9 or **(b)** L4-R4 TALENs. The plots show normalized percentage of GFP-positive cells at specified days post nucleofection. Asterisks indicate significant difference between donor only and donor plus nuclease at specified days. N= 3.

3.3 Discussion

In this work we systematically characterized a microinjection-based method for the direct delivery of genome editing reagents into human K562 cells. Similar to that reported previously[159], we found that adhesion of suspension cells to a surface coated with a sufficient concentration of retronectin facilitates microinjection **(Figure 10)**. We demonstrated the ability of using FACS to generate clones from single injected cells for the analysis of on- and off-target cleavage rates of different nucleases. The results from this study indicate that glass microcapillary-mediated microinjection does not adversely affect the proliferation potential of cells **(Figure 14)** and provides well-controlled delivery of nucleases and donor templates **(Figure 15b)**. Using this microinjection method, we demonstrated high levels of targeted indels **(Figure 18)** and gene modification **(Figure 21b)** in human somatic cells by TALENs and CRISPR/Cas9. This provides a novel approach for performing genome editing with precise control of the amount of nucleases and donors delivered into single cells.

Compared with other delivery methods for genome editing, such as transfection, nucleofection, viral-based delivery and receptor-mediated protein uptake[170], microinjection has several potential advantages, including: (1) precise control of the amount delivered, (2) applicable to a wide variety of cell types (such as primary cells)[159], (3) robust in delivering different forms of genome editing reagents (DNA, RNA, protein)[171], and (4) with minimal waste of reagents[146]. In particular, precise control of the amount of nucleases and donor templates delivered may allow increased gene modification efficiency, and reduced cytotoxicity and off-target effects[147]. Our study is among the first to demonstrate the use of microinjection to deliver genome editing reagents into human somatic cells and achieve a high level of gene modification, including gene insertion. However, although we were able to microinject up to 1000

cells on retronectin coated dishes within 4-6 hours, the microinjection-based approach is low throughput, which is a major limitation. Clearly, there is a need to develop high throughput microinjection systems, especially for applications where processing a large number of cells is required.

3.4 Materials and Methods

HBB-targeting nucleases and donor constructs

The *HBB* NN TALEN (L4-R4) plasmids and β-Ubc-GFP donor template constructs described in[114] were gifts from Dr. Matthew Porteus at Stanford University. The heterodimeric L4- and R4-TALEN target sequences (underlined) separated by a 15-base spacer region: 5'-GCACCTGACTCCTG**T**GGAGAAGTCTGCCGTTACTGCCCT GTGGGG C-3'. The 20-base target sequence (underlined) following the 3-base PAM for the HBB-aiming CRISPR (R02) construct described in[154]: 5'-G**T**GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAC-3'. The nuclease target sites overlap or are in close proximity to the sickle cell mutation (bold) in codon 6 of the *HBB* gene.

Cell culture conditions and transfection

K562 cells (ATCC, Manassas, VA) were grown in Iscove's Modified Dulbecco's Media (IMDM) supplemented with 10% FBS, 2 mM L-glutamine, and 1X penicillin/streptomycin supplement (Invitrogen Life Technologies, Grand Island, NY). For nucleofections, K562 cells were seeded at 1 x 10^6 per well in 6-well dishes. The next day, cells were nucleofected with specified constructs along with 100 ng pmaxGFP (Lonza, Walkersville, MD) using SF cell line 4D-Nucleofection kit (Lonza) according to the manufacturer's protocol in biological triplicates. At 24 hours after nucleofection, the

growth medium was changed in each well. The nucleofection efficiency was determined as the percentage of GFP expressing cells at 3 days after nucleofection using an Accuri C6 flow cytometer (BD Biosciences, Franklin Lakes, NJ). Suspension cultured K562 cells and all derivative clones were grown and maintained in the media conditions listed above in a humidity-controlled incubator with 5% CO_2 at 37°C.

Microinjection on retronectin coated dishes

The coating solution was prepared in Dulbecco's PBS buffer with human recombinant fibronectin fragment CH-296 (Retronectin; TAKARA BIO, Madison, WI) at a final concentration of 100 μg mL^{-1} or specified. The coating solution was aliquoted as 100 uL to coat a circular area of 50.3 mm^3 on an untreated polystyrene 30 mm dish. The coated dishes were incubated overnight at 4°C and washed with 2% BSA. Roughly 20 x 10^4 cells were seeded and attached to each retronectin coated dish by incubating for roughly 2 hours at 37°C. Cells were detached from the dishes by gentle pippetting. The percentage of viable cells detached from the coated surface was calculated using trypan blue staining and automated cell counter (BIO-RAD Laboratories, Hercules, CA). In each microinjection experiment, roughly 1000 cells were injected on 4-5 separate retronectin coated polystyrene dishes.

Sterile glass microcapillaries with 0.5 μm inner tip diameter (Femtotips Narrow; Eppendorf, Hamburg, Germany) were assembled to a capillary holder. The microcapillary position and cell injection was controlled using a programmable InjectMan NI 2 micromanipulator (Eppendorf). Cells were visualized by phase and fluorescence microscopy using a Delta Vision Microscope system equipped with a computer-controlled stage. The pressure applied to release the injection solution from the microcapillary was supplied by a FemtoJet injector (Eppendorf) with built-in air compressor and programmable injection pressure settings to ensure reproducible

injections. The pressure settings applied for cell injections: injection pressure of 60 hPa, injection time of 0.1 seconds, and counter-pressure of 30 hPa. Injections were performed at room temperature. Cells attached to retronectin coated dishes were visualized by phase contrast microscopy. The microcapillary tips were lowered over the cells until both were in the same focal plane, the injection level was defined and programmed into the semiautomatic micromanipulator, which controlled the injection movement at a 45° angle. Injected cells were assessed by fluorescent microscopy to verify successfully injected cells immediately after injection. The percentage of successfully injected cells was determined as the number of fluorescent cells divided by the number of cells injected X 100. The percentage of intact cells was determined as the number of fluorescent cells at 2 hours after injection divided by the number of fluorescent cells immediately after injection X 100. The volume released by the microcapillary was estimated by performing injections using various injection pressures into mineral oil droplets on a microscope slide using the protocol described in[164]. The injection volume was calculated using the equation $V = 4/3\ (\pi r^3)$, where V is the volume and r is the radius of the injected liquid, forming a sphere within the oil droplet. We found the injection volume has an exponential dependence on pressure and is consistent with predictions made in[172].

Preparation of microinjection solution

Injection solutions were prepared in sterile, cold PBS. 10,000 MW dextran-alexa fluor 488 (FITC-dextran; Invitrogen) or 10,000 MW dextran-alexa fluor 594 (TRITC-dextran; Invitrogen) were adjusted to 1 mg mL^{-1}. Unless specified, pmaxGFP was adjusted to 200 ng μL^{-1}, L4-R4-TALEN plasmids were adjusted to 200 ng μl^{-1}, R02 CRISPR/Cas9 plasmid was adjusted to 200 ng μL^{-1}, and β-Ubc-GFP donor template vector was adjusted to 200 or 400 ng μL^{-1}. The pmaxGFP and β-Ubc-GFP donor

template constructs contain GFP isolated from the jellyfish *Aequorea Victoria.* FITC-dextran was co-injected into cells along with L4-R4 TALENs or R02 CRISPR/Cas9 constructs as a marker for successful injection. Injection solutions were centrifuged at 13,000 x g for 20 min at 4°C and the supernatant was directly loaded into microcapillaries for injection.

Isolation of microinjected cells

Cells microinjected with solution supplemented with FITC-dextran were separated by FACS using a FACS Aria II system (BD Biosciences). Cells were gated according to FITC fluorescence intensity levels. For the cell doubling experiments shown in **Figure 14**, cells were stained using To-pro3 (Invitrogen) and further gated for viable, injected cells. The precision mode of sort was set to single cell to ensure high purity. Injected or non-injected cells were deposited 1 cell per well directly into 96-well (Nunc; Thermo Fisher Scientific, Waltham, MA) or 72-well (Terasaki MicroWell; VWR International, Radnor, PA) tissue culture plates containing growth medium. Single injected cells were expanded into clones in a 14-16 day culture.

T7E1 mismatch detection assay

A schematic of the T7E1 detection assay is shown in **Figure 16**. Cleavage activity was quantified in pooled or individual clones derived from single cells microinjected with L4-R4 TALENs or R02 CRISPR/Cas9 expressing plasmids. The off-target activity was measured in the *HBD* and *GRIN3A* loci for L4-R4 TALENs and R02 CRISPR/Cas9 respectively as in[114,154]. The genomic DNA from pelleted cells were processed for PCR amplification as described in[173]. The genomic DNA was harvested using QuickExtract DNA extraction solution (Epicenter Biotechnologies, Madison, WI) and subjected to PCR amplification of on- and off-target loci using primers listed in

Supplementary Table 1. All PCR reactions in 50 μL volume consisting of 1.5 μl genomic DNA were performed using AccuPrime Taq DNA High Fidelity Polymerase kit (Invitrogen) according to the manufacturer's protocol for 40 cycles (94°C for 30 sec, 60°C for 30 sec, 68°C for 45 sec). 200 ng of PCR product supplemented with 1X Accuprime buffer II were processed using cycles of melting and re-annealing (95°C for 10 min, 95-85°C at -2°C/s, 85-25°C at -0.1°C/s). T7E1 (New England Biolabs, Ipswich, MA) was added to a final concentration of 250 units/mL and incubated at 37°C for 60 minutes for digestion of mismatch duplexes. Reactions were resolved on a 2% agarose Tris-acetate-EDTA gel stained with ethidium bromide and observed with a UV imaging station. The intensity of bands corresponding to cleaved and uncleaved PCR product was measured by densitometry analysis using ImageJ software. The percentage of indels was estimated using the equation: 100 X (1-(1-fraction cleaved) $^{1/2}$) as described in[173].

Quantification of HDR-mediated gene modification

For microinjection experiments shown in **Figure 21**, cells were microinjected with 2:1 ratio of L4-R4 TALENs to β-Ubc-GFP donor or 1:2 ratio of R02 CRISPR/Cas9 to β-Ubc-GFP donor. Successfully injected cells were FACS deposited as 1 cell per well into 96-well plates and cultured for 14-16 days in growth medium. GFP fluorescence was observed in single cell clones using fluorescence microscopy. Genomic DNA from individual clones that were positive for GFP fluorescence was harvested and subjected to PCR using *HBB* integrated GFP amplification primers listed in **Supplementary Table 2**. Cells with HDR-mediated gene modification were defined as having both GFP fluorescence and PCR amplified integrated GFP.

In nucleofection experiments shown in **Figure 22**, cells (1 X 10^6) were seeded in 6-well plates and nucleofected with 2 μg R02 CRISPR/Cas9 or L4-R4 TALENs with and

without 10 ug β-Ubc-GFP donor plasmid. pUC18 (Mock) plasmid was added to bring the total DNA amount to 12 ug for each reaction. Cells were incubated in growth medium for 21-days and the percentage of GFP cells was analyzed using an Accuri C6 flow cytometer at specified time points. The percentage of GFP-positive cells was normalized using that at Day 3.

Sanger sequencing for on-target indel rates

Individual clones derived from single cells microinjected with L4-R4 TALENs or R02 CRISPR/Cas9 were PCR amplified using barcoded *HBB* primers listed in **Supplementary Table 3.** PCR amplicons from clones that had positive T7E1 results were separately ligated into a vector and transformed into competent cells using NEB PCR Cloning kit (New England Biolabs) or Zero Blunt TOPO Cloning kit (Invitrogen) according to the manufacturers' protocols. Plasmid DNA from picked *E. coli* colonies were purified and sequenced using Sanger DNA sequencing technology (Operon, Huntsville, AL). Barcodes ensured that sequences were properly matched to individual clones derived by from single microinjected cells.

Statistical analysis

Significance was determined from three or more replicates or samples using Student's t-test or one-way ANOVA. P-values < 0.05 were considered statistically significant. Bars or data points in plots are shown as statistical mean ± standard deviation.

CHAPTER 4: COMPARISON OF CRISPR/CAS9 NUCLEOFECTED AS DNA, MRNA, AND PROTEIN IN K562 CELLS

4.1 Introduction

The recently developed RNA-guided engineered nucleases for robust gene editing is the type II CRISPR/ Cas9 protein from *Streptococcus pyogenes*[134]. Providing protection from invading nucleic acids, CRISPR/Cas9 systems play a role in the adaptive immune systems in bacteria[97]. The Cas9 endonuclease is directed to a target site in the genome by a sgRNA through Watson-Crick base-pairing rules[97]. The Cas9 endonuclease cleaves the complementary 20 nucleotide target site specified on the sgRNA immediately 5' of the NGG PAM sequence [97]. Redirecting the CRISPR/Cas9 to nearly any desired target site in the genome requires modification of the targeting sgRNA sequence while the other components remain fixed, making the process of developing gene-specific CRISPR/Cas9 systems simpler compared to ZFN and TALEN nuclease platforms.

The CRISPR/Cas9 system activates a DSB at its target locus, triggering repair by NHEJ or HDR [174]. DSB repair via the NHEJ pathway results in indels at the break site, generated to obtain alignment in complementary bases for repair, and gene disruption [67]. The HDR pathway involves high fidelity repair of the broken ends using homologous sequences found in sister chromatids, homologous chromosomes or exogenous donor template DNA, and can be exploited for targeted gene modification [67]. Because of their versatility, CRISPR/Cas9 nucleases have been applied for gene editing in human cell lines[175,176], efficient generation of a wide range of transgenic animal models[103-105,107,177], and for engineering plants and crops[108-110]. One unique feature of the system is that the

Cas9 endonuclease can be co-delivered with two or more sgRNAs targeting multiple sites simultaneously, enabling multiplex disruption within the genome[98,107]. The disadvantage of the CRISPR/Cas9 platform is its high frequency of cleavage activity at off-target sites in the genome, potentially resulting in gross chromosomal deletions and other types of untoward chromosomal rearrangements[99-101]. The recent application of ZFNs for targeting mutations causing X-linked-SCID[120] and sickle cell disease[121] in human $CD34^+$ cells shows that gene editing tools offer promising curative approaches. The facile process of making CRISPR/Cas9 nucleases and their robust gene targeting activity makes them highly desirable, providing the impetus to investigate approaches to enhance their safety for clinical applications.

The development of methodologies for delivering gene editing tools is critical for their translation into clinical therapies. IDLV vectors have been shown to provide high levels of nuclear delivery[137], but risks for insertional mutagenesis impedes their clinical application. Typically gene editing tools encoded on plasmid DNA are delivered into human cell lines with potential insertion of plasmid DNA fragments into the target locus or off-target sites in the genome[116], and uncontrolled nuclease expression in cells[100]. Alternatively, delivery of transcribed mRNA encoding nucleases has resulted in high levels of gene editing[120,121]. Although RNA delivery is less toxic in cells compared to plasmid DNA[178], nucleases encoded on mRNA can provide high frequencies of off-target gene modification[121]. Purified protein has been proposed as a strategy to achieve controlled nuclease exposure in cells[138,147]. However, the delivery of purified proteins into cells is challenging due to membrane barriers. Recent studies have shown that nucleofection and cationic lipid reagents can enable efficient delivery of Cas9 RNP complexes into human cell lines, resulting in high frequencies of on-target activity[179,180]. To our knowledge, there have not been previous studies that directly compare CRISPR/Cas9 nucleases delivered as a plasmid DNA, mRNA, and RNP complexes.

In this study, we optimize nucleofection of *HBB* targeting CRISPR/Cas9 nucleases in the human leukemia K562 cell line. We compare the on- and off-target activity of the *HBB*-aiming CRISPR/Cas9 delivered as plasmid DNA, mRNA, and RNPs. This study provides insight into delivery strategies that improve the safety of the CRISPR/Cas9 system.

4.2 Results

Nucleofection of protein into K562 cells

We investigated whether it was feasible to deliver proteins into cells using nucleofection. We used Alexa Fluor 488 conjugated bovine serum albumin (BSA) as a marker for successful protein uptake into cells. K562 cells were analyzed using flow cytometry at 2 hours after nucleofection. The results revealed dose-dependent uptake of BSA into cells with the highest uptake observed (96.8%) when 60 ug of Alex Fluor 488 BSA was nucleofected into cells **(Figure 23)**. These results confirm that purified proteins can be nucleofected into cells with high delivery efficiency[179].

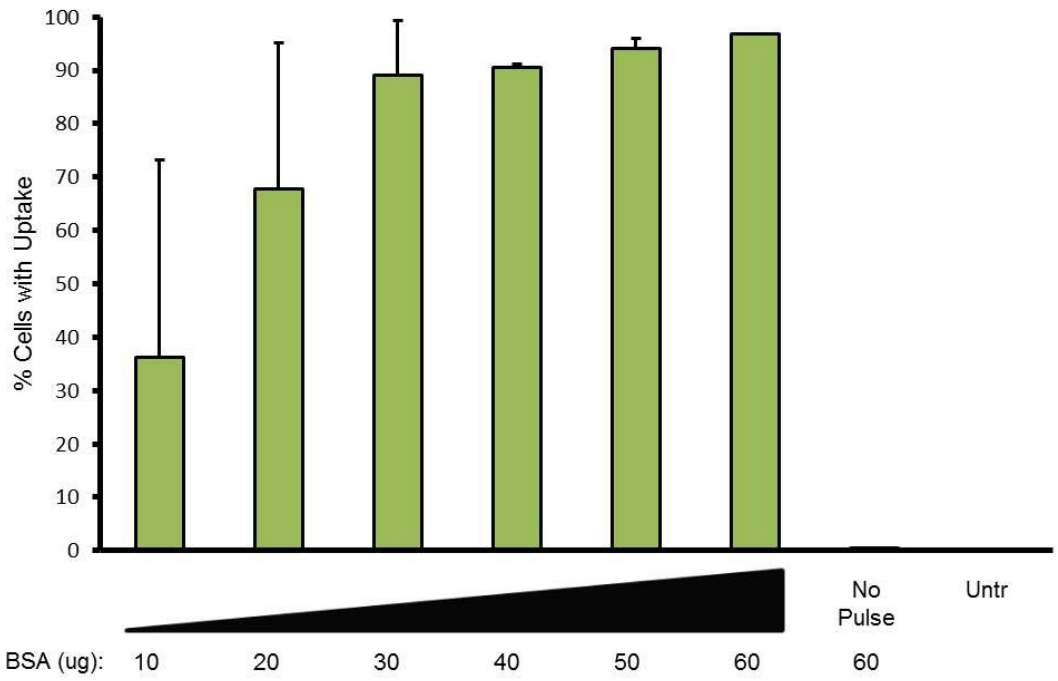


Figure 23: Nucleofection-mediated uptake of BSA into K562 cells

Plot of percentage of uptake in cells nucleofected with specified amounts of Alexa Fluor 488 conjugated BSA. Cells were analyzed using flow cytometry at 2 hours after nucleofection. Control cells were incubated with 60 ug Alexa Fluor 488 conjugated BSA without nucleofection or untreated. Bars represent mean percent uptake in cells ± standard deviation. N = 4.

Enhancing specificity for CRISPR/Cas9 via RNP and mRNA delivery

We optimized the nucleofection of *HBB*-aiming R02 CRISPR/Cas9 into K562 cells. Cells were harvested at 3 days after nucleofection and the genomic DNA was subjected to mutation detection assays to quantify indels. On-target indels were measured using the T7E1 assay and the off-target activity was measured in *GRIN3A*[99] using TIDE[136]. We observed dose-dependent on- and off-target cleavage activity in K562 cells nucleofected with plasmid DNA encoding for the R02 CRISPR/Cas9 system **(Figure 24a)**. The highest on-target activity measured was 40% with a corresponding

9.5% off-target activity when 1 ug of R02 plasmid DNA was nucleofected into cells **(Figure 24a)**. We optimized the amount of synthetic R02 sgRNA and Cas9 mRNA nucleofected into cells. The Cas9 mRNA encoded for 2 NLS sequences, similarly to the R02 CRISPR/Cas9 plasmid DNA. Nucleofecting different amounts of sgRNA with a fixed amount of Cas9 mRNA into cells revealed that the on-target indels depended on the sgRNA dose **(Figure 24b)**. We observed the highest on-target indels (58.3%) when 10 ug sgRNA along with 2 ug Cas9 mRNA was nucleofected into cells. Interestingly, the dose of sgRNA along with the Cas9 mRNA did not influence the off-target activity **(Figure 24b)**. Recombinant Cas9 endonuclease proteins designed with a His-tag for purification, FLAG-tag, and two NLS sequences, one on both ends, were purified from *Escherichia coli*. The Cas9 proteins were incubated with R02 sgRNA for 10 minutes at room temperature to stimulate RNP complex formation prior to nucleofection. When the sgRNA-Cas9 molar ratio was maintained at 1.2, we observed negligible levels of on-target indels (<4%) even for high amounts of Cas9 protein **(Figure 24c)**. With a sgRNA-Cas9 molar ratio of 10.3, we observed high levels of on-target indels (33%) with RNP delivery **(Figure 24c)**. There was undetectable off-target indels induced by the CRISPR/Cas9 RNP complexes in cells (data not shown). These results show that robust indels can be generated by CRISPR/Cas9 delivered as mRNA or RNP complexes that can be controlled by adjusting the dose.

We next investigated the effects of CRISPR/Cas9 delivery as plasmid DNA, mRNA, and RNP complexes on the targeting specificity. We quantified the off- to on-target ratio for each delivery strategy providing the highest on-target indels. Although the R02 CRISPR/Cas9 mRNA delivery provided higher on-target indels compared to the plasmid DNA, it had a lower ratio of off- to on-target indels, which indicates higher specificity **(Figure 24d)**. The lowest off- to on-target ratio was observed for the RNPs, whereas the plasmid DNA had the highest ratio **(Figure 24d)**. Interestingly, the RNP

provided significantly higher specificity relative to the mRNA **(Figure 24d)**. These results indicate that CRISPR/Cas9 nucleofected as RNPs and mRNA provides higher specificity compared to plasmid DNA.

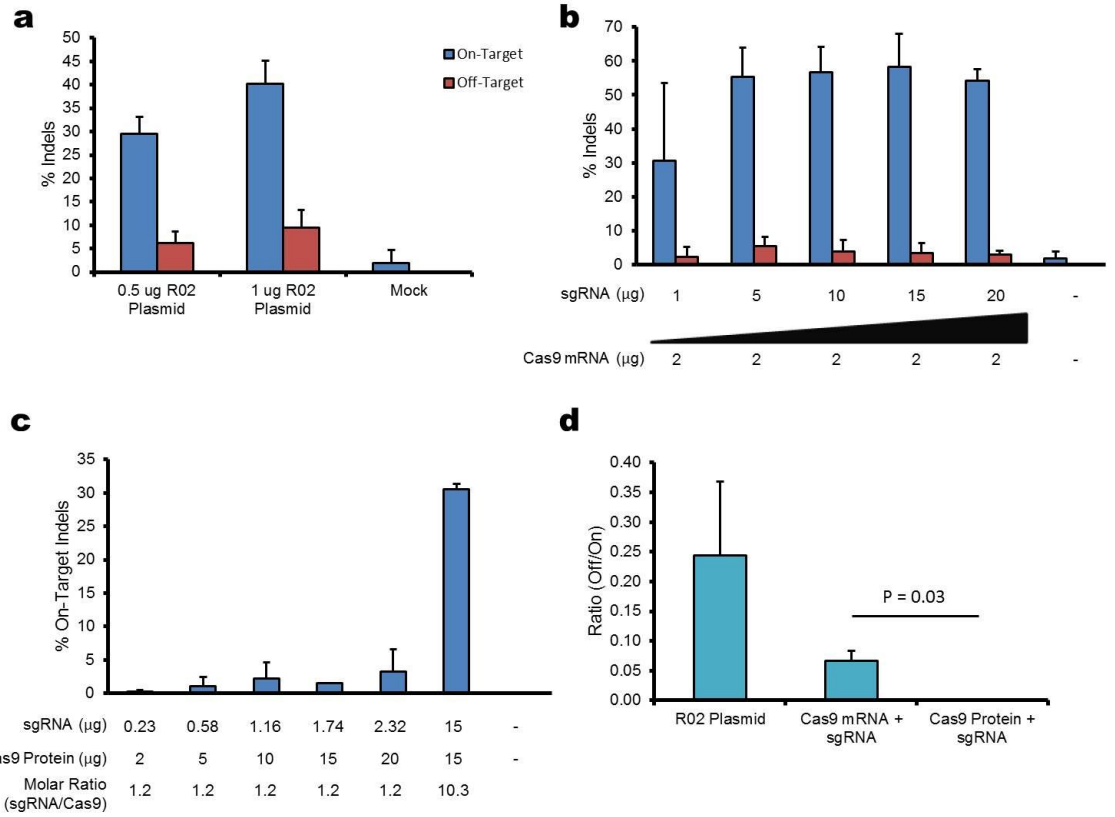


Figure 24: Indels by R02 CRISPR/Cas9 nucleofected as plasmid DNA, mRNA, and RNPs into K562 cells

(a) Plot of on- and off-target indels induced by the *HBB*-aiming R02 CRISPR/Cas9 delivered as a plasmid DNA into cells. **(b)** Indels induced by Cas9 mRNA with different amounts of synthetic R02 sgRNA specified below the plot. **(c)** On-target indels induced by different amounts of Cas9 purified protein and sgRNA complexes. Specified below the plot are the amounts of sgRNA and Cas9 protein, and sgRNA-Cas9 molar ratios. Indels induced by the control was shown in each plot. **(d)** Plot showing the ratio of off- to on-target indels induced for plasmid DNA, mRNA, and RNP complex conditions that provided the highest on-target indels. Bars represent the average of 2 biological replicates ± the standard deviation. The p-value for is shown to indicate statistical significance.

HDR in K562 cells nucleofected with CRISPR/Cas9 and donor vector

Next, we investigated the frequency of HDR-mediated gene medication in K562 cells nucleofected with the CRISPR/Cas9 plasmid DNA, mRNA, and RNPs along with donor plasmid DNA. The donor vector contained an EcoRI restriction site, and right and left homology arms to the *HBB* gene, which included silent point mutations to prevent binding and cleavage by the R02 sgRNA. The insertion of the EcoRI restriction site via the HDR pathway **(Figure 25a)** was quantified using a restriction fragment length polymorphism (RFLP) assay (**Figure 26)** at 3 days after nucleofection. The amount of donor vector in the initial nucleofection solution was the same for each sample to ensure that any potential difference in HDR was due to the delivery strategy. We nucleofected the same amount of R02 CRISPR/Cas9 plasmid DNA, mRNA, and RNP used in **Figure 24d**. We observed similar rates of HDR (roughly 20%) in cells nucleofected with CRISPR/Cas9 plasmid DNA, mRNA, and RNPs **(Figure 25b)**. This result indicates that CRISPR/Cas9 delivered via RNPs and mRNA provides similar levels of HDR as plasmid DNA.

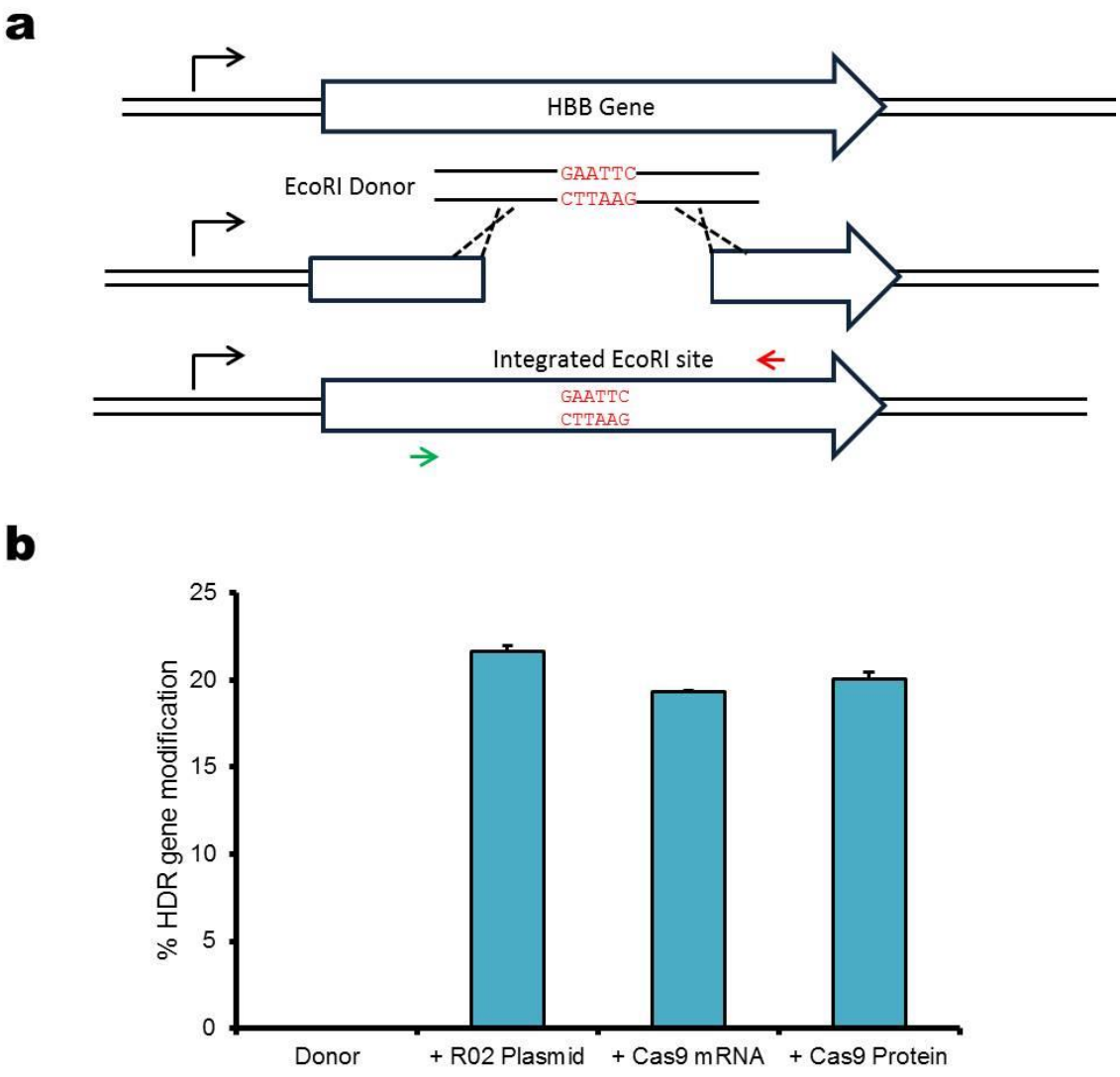


Figure 25: Gene modification by R02 CRISPR/Cas9 system and donor vector quantified using RFLP assay

(a) Schematic of the RFLP reporter system used to detect HDR-mediated gene modification. Nuclease cleavage and resection yields a substrate for HDR which may involve the use of exogenous HBB-EcoRI donor, flanked by 5' and 3' homologous fragments of the *HBB* gene containing silent point mutation to prevent binding and cleavage by the R02 sgRNA. The dashed lines indicate the HDR process with the donor vector. Gene modification was confirmed by PCR using primers specific for the target locus, as shown by the green and red arrows, and the addition of EcoRI restriction enzyme. **(b)** Plot of the percentage of alleles repaired via the HDR pathway in cells nucleofected with R02 CRISPR/Cas9 plasmid DNA, mRNA, and RNPs along with HBB-EcoRI donor vector. Bars represent the average of 2 biological replicates ± the standard deviation.

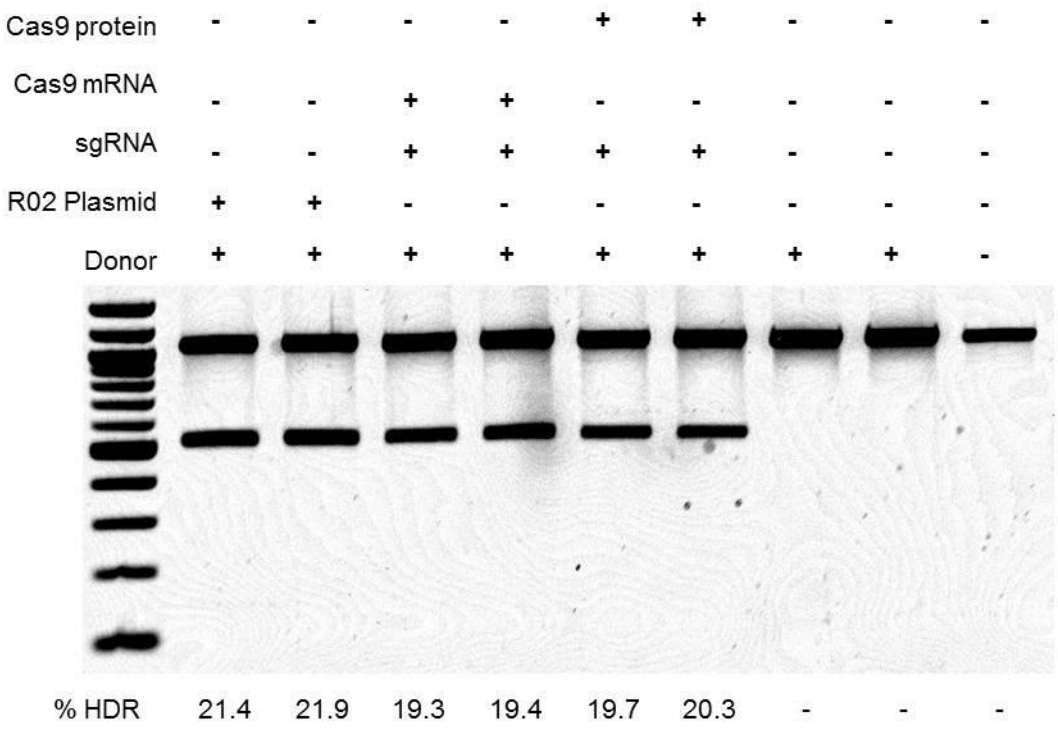


Figure 26: HDR frequency by R02 CRISPR/Cas9 and HBB-EcoRI donor vector measured using the RFLP assay in K562 cells

PCR products from genomic DNA extracted from cells nucleofected with R02 CRISPR/Cas9 plasmid DNA, mRNA, or RNP complexes along with HBB-EcoRI donor vector. Control cells were nucleofected with donor only or untreated. Purified PCR amplicons (200 ng) were treated with EcoRI restriction enzyme before loading on a gel. The bottom, lower molecular weight band represents the cleaved band. Shown below is the percentage of HDR quantified using densitometry analysis. Biological duplicates are shown for each sample.

4.3 Discussion

A major motivation for this study is the urgent need for delivery methodologies to enhance the safety of gene editing tools. There have been several recent studies focusing on delivering nucleases as purified proteins[138,170,181] as a strategy to enhance the nuclease specificity, which is particularly a concern for CRISPR/Cas9 systems[99]. In this study, we compare the delivery of *HBB*-targeting CRISPR/Cas9 nucleases delivered

as plasmid DNA, mRNA, and RNPs **(Figure 24)**. In our comparison studies, we used Cas9 endonucleases that have a dual NLS sequence. For the first time, we show that CRISPR/Cas9 RNPs can be nucleofected with high efficiency and provide higher specificity compared to mRNA and plasmid DNA **(Figure 25d)**. The delivery of CRISRP/Cas9 via mRNA nucleofection provides higher activity and specificity compared to plasmid DNA, although the difference observed is not significant. For the RNP delivery, we show that a critical factor for optimizing cleavage activity is the sgRNA-Cas9 protein molar ratio. A sgRNA-Cas9 protein molar ratio of 10.3 provided high levels of on-target indels **(Figure 24c)**. We also found that the delivery of CRISPR/Cas9 via RNP and mRNA provided robust HDR-mediated gene modification that was similar to plasmid DNA **(Figure 25b)**. It will be interesting to further investigate whether the CRISPR/Cas9 activity can be further optimized using smaller amounts of purified Cas9 protein to improve the efficiency of the approach and eliminate the waste of reagents.

In conclusion, we show that *HBB* CRISPR/Cas9 nucleofected as RNPs and mRNA provide higher specificity compared to plasmid DNA and without altering the levels of HDR-mediated gene modification in human cells. The safety concerns associated with CRISPR/Cas9 can be addressed by mRNA and especially RNP delivery, which have the potential to eliminate off-target cleavage activity and broaden the spectrum of its application.

4.4 Materials and Methods

CRISPR/Cas9 and donor constructs

The *HBB*-aiming CRISPR (R02) construct described in [99] has a 20-base target sequence (underlined): 5'-G**T**GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAC-3'.

The target site is in close proximity to the sickle cell mutation (bold) in codon 6 of the *HBB* gene. The mRNA encoding the R02 sgRNA and Cas9 were ordered from TriLink Biotechnologies (San Diego, CA). The Cas9 plasmid DNA, mRNA, and purified protein contained a dual NLS to increase importation into the cell nucleus. Purified Cas9 proteins were generously supplied by Dr. Charles Gersbach at Duke University. Specified amounts of Cas9 protein and sgRNA was mixed and incubated at room temperature for 10 minutes to induce RNP complex formation. The HBB-EcoRI donor vector used to quantify HDR-mediated gene targeting was designed with right and left homology arms from the *HBB* locus isolated from K562 cells that contained an EcoRI restriction site cloned into a puc18 plasmid. The construct sequence of the HBB EcoRI donor is found in the supplementary information.

K562 cell culture and nucleofection

K562 cells (ATCC, Manassas, VA) were grown in RPMI 1640 (Hyclone, Logan, UT) supplemented with 10% fetal bovine serum and 2 mM L-glutamine (Invitrogen Life Technologies, Grand Island, NY). K562 cells (2×10^5) were nucleofected with a 4D nucleofector (Lonza, Walkersville, MD) using Amaxa SF Cell Line kit (V4XC-2032) and program FF-120 according to the manufacturer's protocol. In the BSA uptake experiments, K562 cells were nucleofected with specified amounts of BSA Alexa Fluor 488 conjugates (Life Technologies, Carlsbad, CA) and analyzed using an Accuri C6 flow cytometer (BD Biosciences, Franklin Lakes, NJ) at 2 hours after nucleofection. Suspension cultured K562 cells were grown and maintained in the media conditions listed above in a humidity-controlled incubator with 5% CO_2 at 37°C.

Measuring indels using TIDE and T7E1 assay

Cleavage activity was quantified in pooled K562 cells nucleofected with R02 CRISPR/Cas9 plasmid DNA, mRNA, or RNP complexes. The off-target activity was measured in the *GRIN3A* locus [99]. The genomic DNA was harvested using QuickExtract DNA extraction solution (Epicenter Biotechnologies, Madison, WI) and subjected to PCR amplification of on- and off-target loci using primers listed in **Supplementary Table 5**. All PCR reactions in 50 μL volume consisting of 1.5 μl genomic DNA were performed using AccuPrime Taq DNA High Fidelity Polymerase kit (Invitrogen) according to the manufacturer's protocol. For the T7E1 analysis of off-target indels, 200 ng of purified *GRIN3A* PCR product supplemented with 1X Accuprime buffer II were processed using cycles of melting and re-annealing (95°C for 10 min, 95-85°C at -2°C/s, 85-25°C at -0.1°C/s). T7EI (NEB, Ipswich, MA) was added to a final concentration of 250 units/mL and incubated at 37°C for 60 minutes for digestion. Reactions were resolved on a 2% agarose Tris-EDTA gel stained with ethidium bromide. The percentage of off-target indels was estimated by densitometry analysis using ImageJ software and the equation: 100 X (1-(1-fraction cleaved) $^{1/2}$) as described in [173]. For quantifying the on-target indels using TIDE [136], purified *HBB* PCR product were Sanger-sequenced using the forward and reverse primers separately, and subsequently analyzed using the online TIDE software (accessible at http://tide.nki.nl).

RFLP assay for quantifying HDR-mediated gene modification

K562 cells were nucleofected with HBB-EcoRI donor vector with and without R02 CRISPR/Cas9 delivered as plasmid DNA, mRNA or RNPs. The genomic DNA was harvested from the cell pellet using QuickExtract DNA extraction solution and subjected to PCR amplification of the *HBB* gene using RFLP primers listed in **Supplementary**

Table 6. 200 ng of the purified PCR product was supplemented with 1X CutSmart and 10 units of EcoRI enzyme (NEB) and was incubated at 37C for 1 hour for digestion. The samples were resolved on a 2% TRIS-EDTA gel stained with ethidium bromide. The cleaved bands corresponding to alleles containing the EcoRI site were quantified by densitometry analysis using ImageJ software.

Statistical analysis

Significance was determined using the Student's t-test. P-values < 0.05 were considered statistically significant. Bars in plots are shown as statistical mean ± standard deviation.

CHAPTER 5: NUCLEOFECTION PROTOCOL OPTIMIZATION AND GENOME EDITING IN HUMAN CD34$^+$ CELLS

5.1 Introduction

HSCs are critical substrates for treating SCD using precision gene engineering tools. Recently there have been several studies that demonstrate the application of ZFNs[120,121,151], TALENs[116], and CRISPR/Cas9 systems[182] for gene modification in human CD34$^+$ cells. The most advanced showing gene correction of SCD is the one by Hoban et al. that demonstrated successful gene correction of the SCD mutation in 18.4% of alleles in CD34$^+$ cells isolated from the bone marrow (BM) aspirates of a SCD patient[121]. Functional correction of the SCD mutation was confirmed by analysis of the globin tetramers, which showed 5% induction of HbA in samples treated with *HBB*-targeting ZFNs and donor template[121]. The study also showed high levels of off-target activity by the ZFNs in the homologous *HBD* gene[121]. Although these results are promising, there are concerns about the safety of the gene editing strategy and feasibility as a clinical therapy for SCD. It is of paramount interest to enhance the safety of gene editing tools without compromising the frequency of HDR-mediated gene correction. One strategy that takes advantage of the robust cleavage activity offered by the CRISPR/Cas9 system while limiting its off-target activity is the D10A mutant Cas9 nickases (Cas9n) that targets opposite strands of the target site by a pair of offset sgRNAs[133]. This double-nicking strategy was shown to provide up to >100-fold greater specificity and comparable HDR compared to the wild type Cas9 nuclease[133]. This study will explore the application of Cas9n for gene editing in CD34$^+$ cells. To our knowledge, there are no previous studies that compare RNA-guided nucleases and nickases in CD34$^+$ cells.

Nucleofection is an attractive approach for delivering gene editing tools because it provides the versatility for delivering nearly any type of molecule, including DNA, mRNA, and proteins[120,179,183]. Although nucleofection has been shown to provide high transfection efficiency (>80%) with a mean survival of roughly 50% in $CD34^+$ cells[184], there are many factors, including the number of cells used in the nucleofection reaction and cell culturing protocol, that may influence the cell viability and gene expression after nucleofection[184]. The absence of standardized cell culturing protocols makes the application of the nucleofection technique particularly challenging in HSCs. Studies that apply nucleofection for delivering gene editing tools into HSCs, typically deliver the nucleases as plasmids[151] or mRNA[121]. To our knowledge, there have not been any studies in the literature that directly compares the delivery of nucleases as plasmid DNA, mRNA, and purified proteins in HSCs.

In this study, we will explore the application of nucleofection for delivering *HBB*-aiming CRISPR/Cas9 nucleases and nickases as mRNA, DNA, and protein into $CD34^+$ cells. We will compare BM and umbilical cord blood (CB) sources of $CD34^+$ cells for cell viability, transfection efficiency, and nuclease cleavage activity. Finally, we will address the feasibility of applying drug selection to enrich for gene corrected $CD34^+$ cells nucleofected with the CRISPR/Cas9 system and donor template.

5.2 Results

Optimization of *in vitro* cell culturing and nucleofection protocols for human $CD34^+$ cells

We applied the Amaxa 4D nucleofector system (Lonza) and P3 Primary Cell nucleofection kits to transfect cryopreserved $CD34^+$ cells from healthy donors. The advantage of using the 4D nucleofector system is its optimized pre-programmed

electrical pulses and buffers for specific cell types and it can be used for nucleofecting small numbers of cells (as few as 5 x 10^4) per reaction in minimal volumes of nucleofection buffer. The 4D nucleofector system has been applied in several studies in the literature to successfully deliver mRNA and plasmid DNA into freshly isolated human HSCs[120,151]. There are many protocols for culturing HSCs available in the literature that involve suspension cultures supplemented with various combinations of early acting cytokines, such as stem cell factor (SCF), Flt-3 ligand (Flt3-L), and interleukin-3 (IL-3), as well as late-activing cytokines, including thrombopoietin (TPO), interleukin-6 (IL-6), and erythropoietin (EPO)[185,186]. The manufacturer's recommendation involved nucleofecting the $CD34^+$ cells on the same day after thawing. We used StemSpan Serum Free Expansion Media (Stem Cell Technologies) as the base media for thawing and culturing $CD34^+$ cells. Early and lineage specific-late-acting human cytokines (SCF, IL-3, IL-6, and EPO) were supplemented to the base media for cell proliferation and differentiation to erythroid progenitors[186]. When following the manufacturer's instructions to nucleofect cells with R02 CRISPR/Cas9 and pmaxGFP plasmid DNA immediately after thawing and using 2 x 10^5 $CD34^+$ cells per reaction, we obtained drastically lower levels of cell viability and nucleofection efficiency in BM and CB $CD34^+$ cells than reported in the literature[184]. At day 2 after nucleofection and culturing in erythroid media, the cell viability was 3.5% and 6% in BM and CB $CD34^+$ cells respectively **(Figure 27a)**. Given the high cell viability in untransfected cells, it is likely that the stress induced by the combination of thawing and nucleofection was the cause for the cell death. With prolonged culturing, we observed incremental increase in the cell viability for both BM and CB $CD34^+$ cells with each successive day of culturing until the cell viability approached that of the untransfected cells, particularly for CB $CD34^+$ cells **(Figure 27a)**. The increase in cell viability is likely due to proliferation of the few surviving cells after nucleofection, which gradually outnumbered the dead cells. The nucleofection efficiency

was estimated by the percentage of cells that were positive for GFP expression. The GFP expression was roughly 50% in the surviving cell population for both CB and BM $CD34^+$ cells, which gradually decreased during the 10-day culture after nucleofection as the successfully transfected cells lost GFP plasmid DNA **(Figure 27b).** In the literature, studies that show high levels of cell survival typically culture the $CD34^+$ cells for 5-48 hours prior to infection using viral vectors or nucleofection[120,151]. We hypothesized that culturing the cells in cell culture media before nucleofection may enable the cells to recover from the thawing process and enhance cell survival after nucleofection. In our modified *in vitro* cell culture and nucleofection protocol, we used expansion media (supplemented with Flt3-L, SCF, TPO, and IL-3)[185], cultured the cells for 48 hours prior to nucleofection, and used 8 x 10^4 cells per reaction. With the protocol changes, we observed a dramatic improvement in the cell viability after nucleofection **(Figure 28a)**. At 24 hours after nucleofection, we observed >50% live cells in the BM and CB $CD34^+$ cells nucleofected with R02 CRISPR/Cas9 and pmaxGFP plasmid DNA **(Figure 28a)**. Similar cell viability was observed in $CD34^+$ cells nucleofected in the absence of any substrate DNA **(Figure 28a)**, which is an indication that nucleofection itself is responsible for cell toxicity whereas the addition of substrate DNA had negligible effects on the cell viability[151]. We applied this improved protocol for nucleofecting $CD34^+$ cells in further experiments.

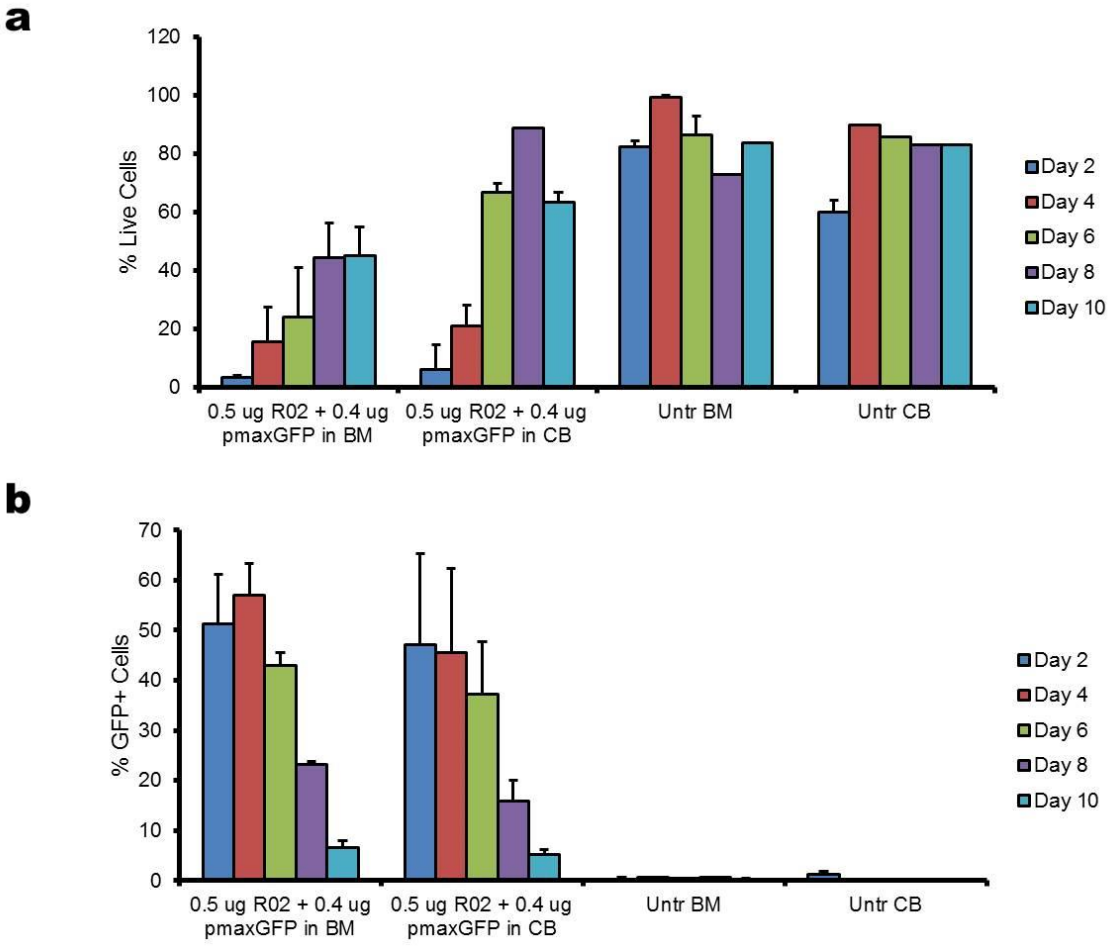


Figure 27: CD34$^+$ cells nucleofected and cultured in erythroid media

BM and CB CD34$^+$ cells were nucleofected (2x10^5 cells per reaction) with specified amounts of R02 CRISPR/Cas9 and pmaxGFP encoding plasmid DNA in duplicates. **(a)** The percentage of live and **(b)** GFP$^+$ cells were monitored over 10 days. Untransfected BM and CB CD34$^+$ cells are shown as controls. Bars represent the mean percentage of live cells or GFP$^+$ cells ± standard deviation.

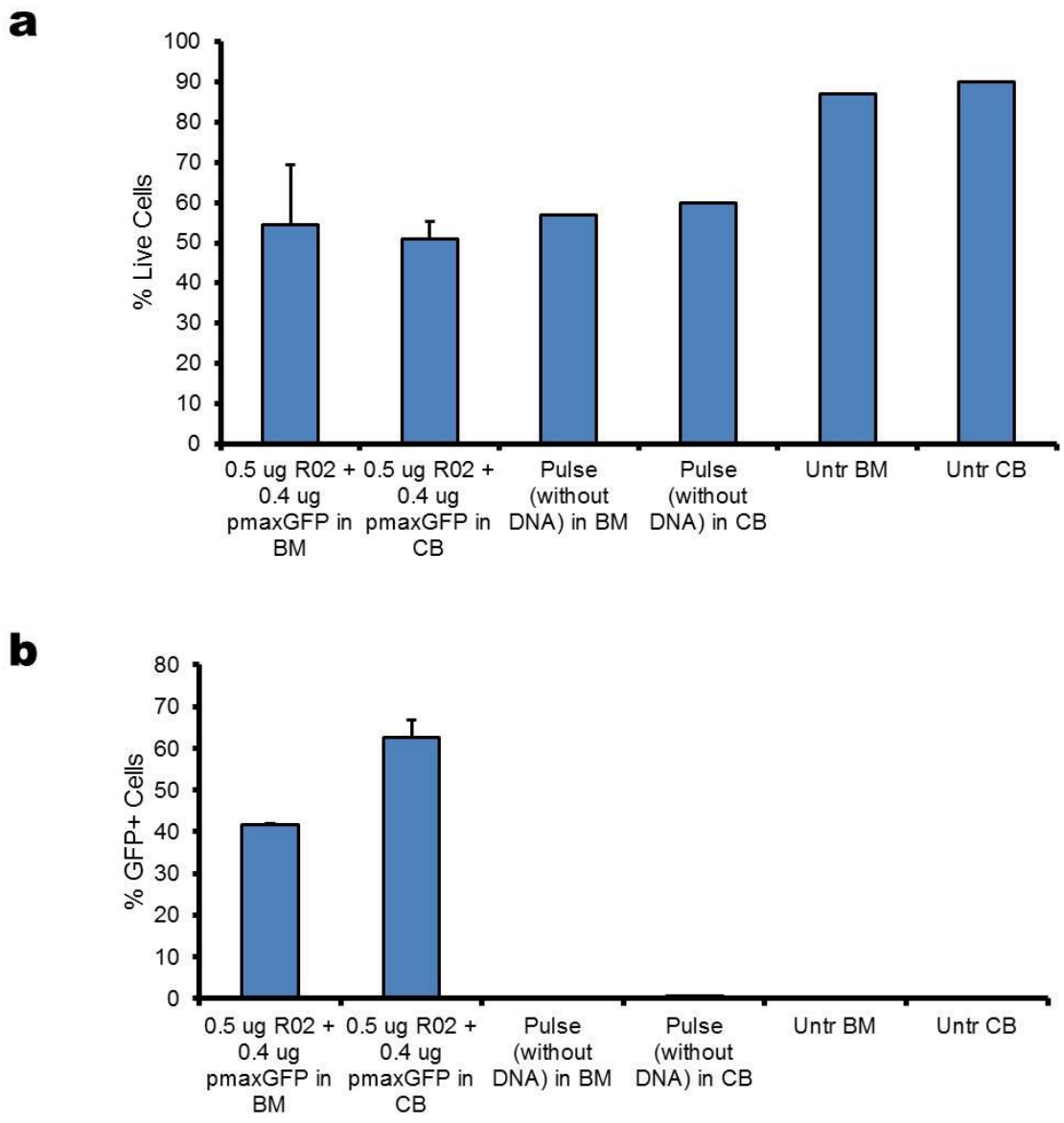


Figure 28: CD34$^+$ cells nucleofected at 48 hours after culture in expansion media

BM and CB CD34$^+$ cells (8x10^4 cells per reaction) were nucleofected with specified amounts of R02 CRISPR/Cas9 and pmaxGFP encoding plasmid DNA after culture for 48 hours in expansion media. Samples were nucleofected in duplicates. **(a)** The percentage of live cells and **(b)** GFP$^+$ cells was determined at 24 hours after nucleofection. Cells that were untransfected or nucleofected without DNA are shown as controls. Bars represent the mean percentage of live cells or GFP+ cells for each sample ± standard deviation.

Comparison of *HBB* CRISPR/Cas9 cleavage activity in CD34$^+$ cells

In the literature, studies have demonstrated high levels of cleavage activity and gene modification in human cells nucleofected with nucleases delivered as DNA[151], mRNA[121], and purified proteins[179]. In Chapter 4, we directly compared CRISPR/Cas9 nucleofected as plasmid DNA, mRNA, and RNPs and showed that RNP and mRNA provide enhanced specificity compared to plasmid DNA in K562 cells **(Figure 24)**. There has not been any comparable study in human CD34$^+$ cells. Given that CD34$^+$ cells are relatively quiescent and progress through the cell cycle infrequently, it is very likely that they will express low levels of nucleases, particularly when delivered as plasmid DNA. The advantage of delivering nucleases as an mRNA or purified protein in CD34$^+$ cells is that active nucleases are present in the cells more rapidly compared to plasmid DNA and have a shorter half-life[179]. We hypothesized that the delivery of the CRISPR/Cas9 system as an mRNA or purified protein will provide higher levels of activity compared to plasmid DNA. We nucleofected BM and CB CD34$^+$ cells with R02 CRISPR/Cas9 plasmid DNA, mRNA, and RNPs and measured the cell viability and the nucleofection efficiency. We formed an RNP complex prior to nucleofection by incubating the Cas9 protein with sgRNA for 10 minutes at room temperature. The nucleofection efficiency was estimated by the percentage of cells that were positive for GFP expression. At 24 hours after nucleofection, we observed similar levels of cell viability (51-71.5%) in the nucleofected cells **(Figure 29a)**. The cell viability for the RNP was slightly higher relative to the plasmid DNA and mRNA delivery, although the difference was not significant **(Figure 29a)**. We observed higher frequencies of GFP expression in cells nucleofected with eGFP mRNA compared to pmaxGFP plasmid DNA, an indicator that mRNA provides faster and more robust gene expression in CD34$^+$ cells **(Figure 29b)**. Interesting, we observed significantly higher nucleofection efficiency in the CB CD34$^+$

cells compared to the BM CD34$^+$ cells **(Figure 29b).** Analysis of the on-target indels using TIDE revealed that the delivery of the R02 CRISPR/Cas9 system as plasmid DNA, mRNA, and RNP provided variable levels of cleavage activities **(Figure 30a).** Delivery of the Cas9 mRNA and sgRNA provided 14% and 16% on-target cleavage activity in BM and CB CD34$^+$ cells respectively **(Figure 30a)**, where the difference in indels was not significant. The off-target activity was undetectable **(Figure 30b)**. On the contrary, we did not observe any activity in neither BM nor CB CD34$^+$ cells nucleofected with the R02 CRISPR/Cas9 plasmid DNA **(Figure 30a).** We observed activity from the RNP complex (17%) in BM, but not in CB CD34$^+$ cells **(Figure 30a)**. Given that the nucleofection efficiency was higher in CB CD34$^+$ cells **(Figure 29b)**, it is likely that the nucleofection efficiency is not the reason for the difference in RNP activity in between BM and CB CD34$^+$ cells. Because the activity in BM CD34$^+$ cells had considerable variability, it is likely that the amount of RNP complex in the nucleofection reaction was not appropriate. The sgRNA-Cas9 molar ratio for the RNP complex nucleofected into CD34$^+$ cells was 14; higher than the ratio used in K562 cells **(Figure 24)**. It is likely that further optimization of RNP complexes in CD34$^+$ cells is necessary to obtain consistent levels of activity.

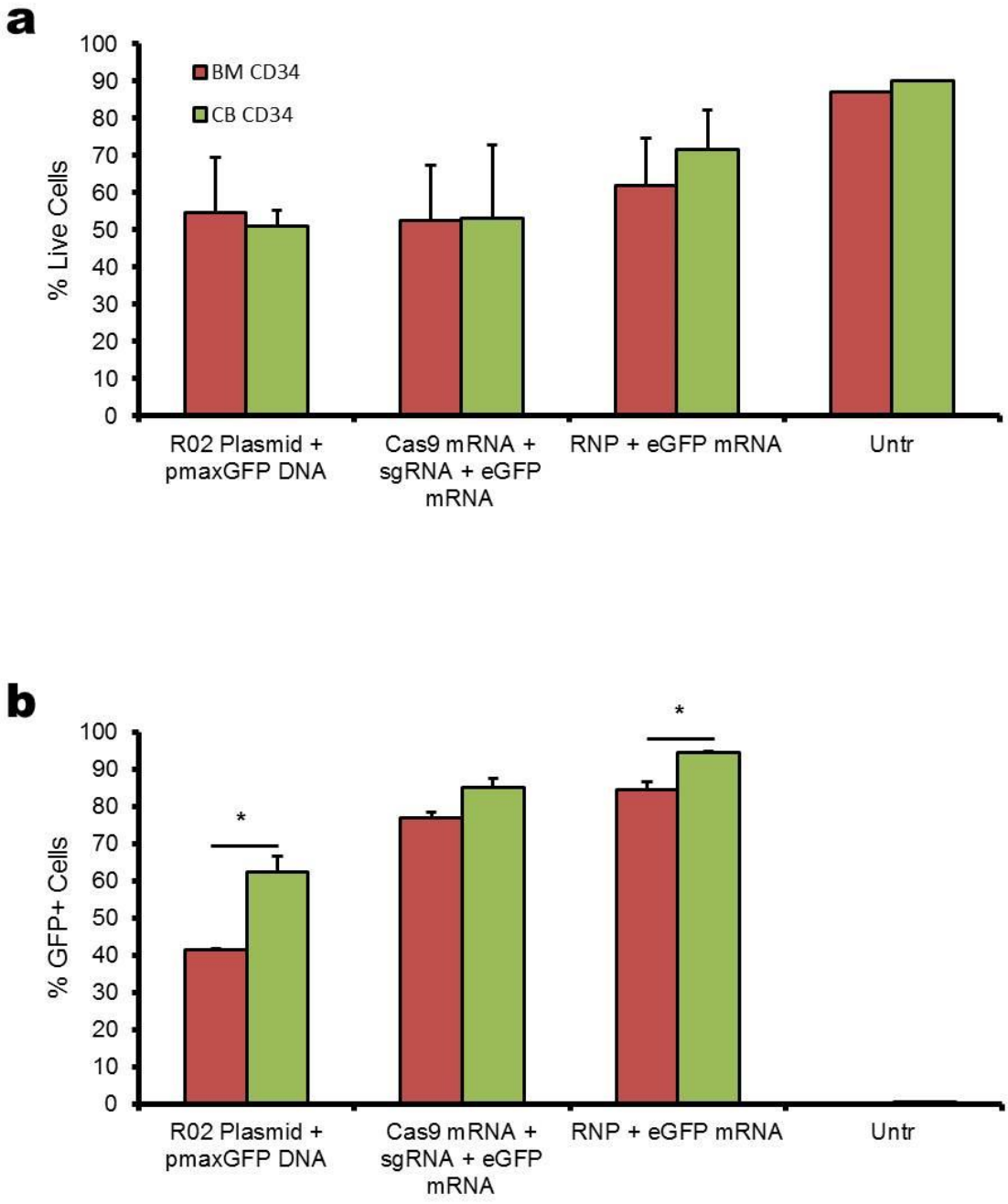


Figure 29: CD34$^+$ cells nucleofected with R02 CRISPR/Cas9 nuclease delivered as plasmid DNA, mRNA, and RNP

BM and CB CD34$^+$ cells were nucleofected with 0.5 μg of R02 plasmid DNA, 1 μg of Cas9 mRNA and 10 μg of R02 sgRNA, or 7.5 μg Cas9 protein and 10 μg R02 CRISPR sgRNA (RNP). **(a)** The percentage of live cells and **(b)** GFP$^+$ cells at 24 hours after nucleofection. Bars represent the mean percentage of live cells or

GFP$^+$ cells ± standard deviation. Asterisks represent statistical significance between the BM and CB data, where $p<0.05$.

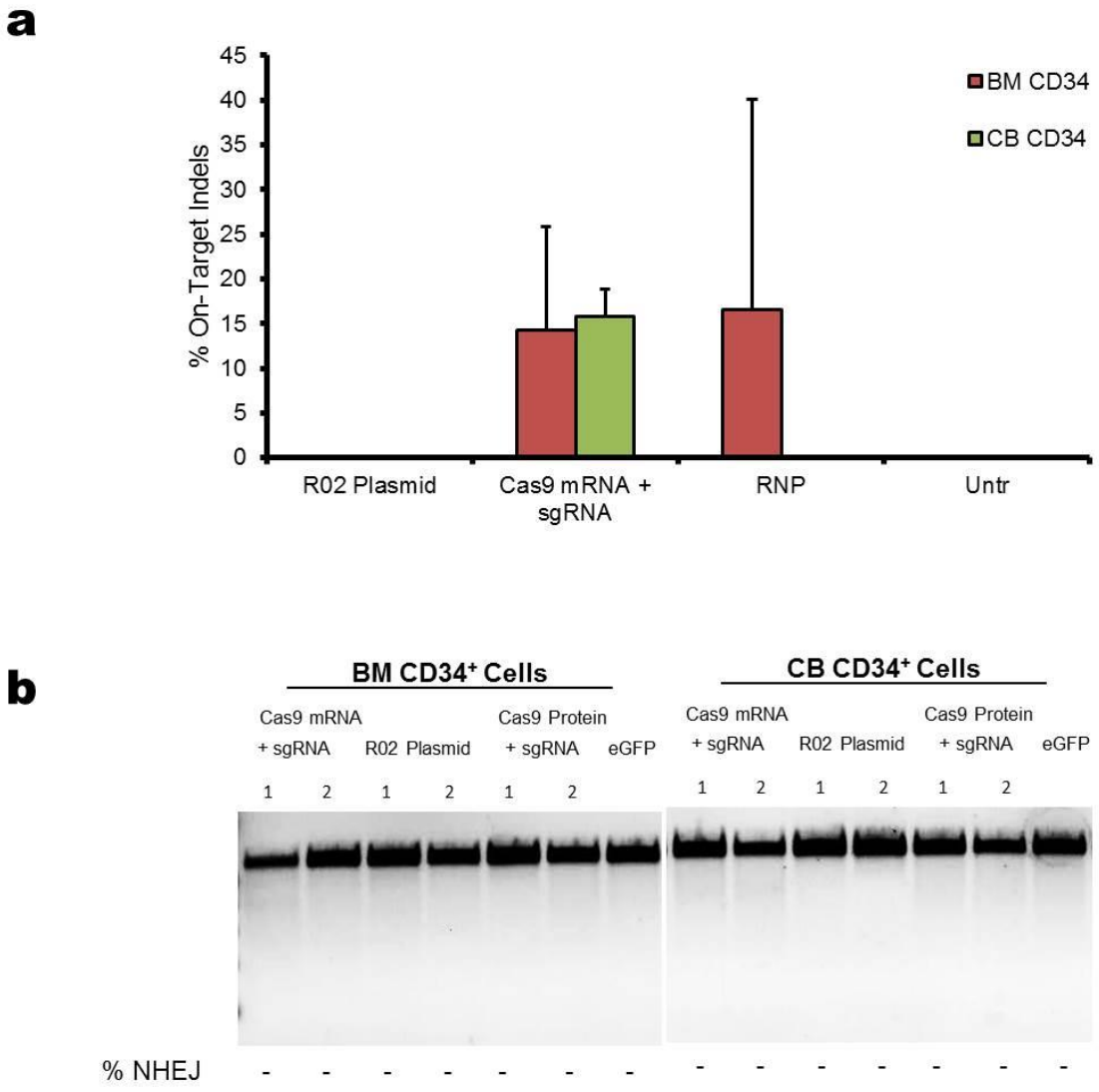


Figure 30: Cleavage activity by R02 CRISPR/Cas9 system in BM and CB CD34+ cells delivered as plasmid DNA, mRNA, and RNP

BM and CB CD34$^+$ cells were nucleofected with 0.5 μg of R02 plasmid DNA, 1 μg of Cas9 mRNA and 10 μg of R02 sgRNA, or 7.5 μg Cas9 protein and 10 μg R02 CRISPR sgRNA (RNP). **(a)** On-target indels measured using TIDE and **(b)** off-target indels measured using the T7E1 mutation detection assay at 3 days after nucleofection.

Indels formed by double Cas9n nucleofected into K562 and CD34$^+$ cells

We next investigated the cleavage activity by an *HBB*-aiming CRISPR/Cas9n system. The double Cas9n have been shown to provide higher specificity compared to wild type Cas9 without compromising the gene-targeting activity[133]. The Cas9n was directed to off-set sites in the *HBB* gene by a pair of sgRNA (R01 and R02) shown in **Figure 31**. K562 cells nucleofected with plasmid DNA encoding for the R01/R02 Cas9n pair, had 25% on-target activity **(Figure 32c)**, which was slightly lower the R02 CRISPR/Cas9 **(Figure 24)**. We confirmed that the K562 cells nucleofected with the R01/R02 nickase plasmid DNA had reasonable cell viability and nucleofection efficiency **(Figure 32).** However, when we delivered the R01/R02 Cas9n in BM and CB CD34$^+$ cells as mRNA and plasmid DNA we did not observe any on-target indels (data not shown). We confirmed that there was reasonable cell viability and nucleofection efficiency in BM and CB CD34$^+$ cells nucleofected with plasmid DNA and mRNA encoding the R01/R02 Cas9n pair **(Figure 33)**. It is important to note that the amount of each R01/R02 sgRNA nucleofected into CD34$^+$ cells along with the Cas9n mRNA was half the amount of sgRNA that we nucleofected for the Cas9 nuclease. We hypothesize that increasing the sgRNA for the Cas9n pair may help to improve the activity in CD34$^+$ cells.

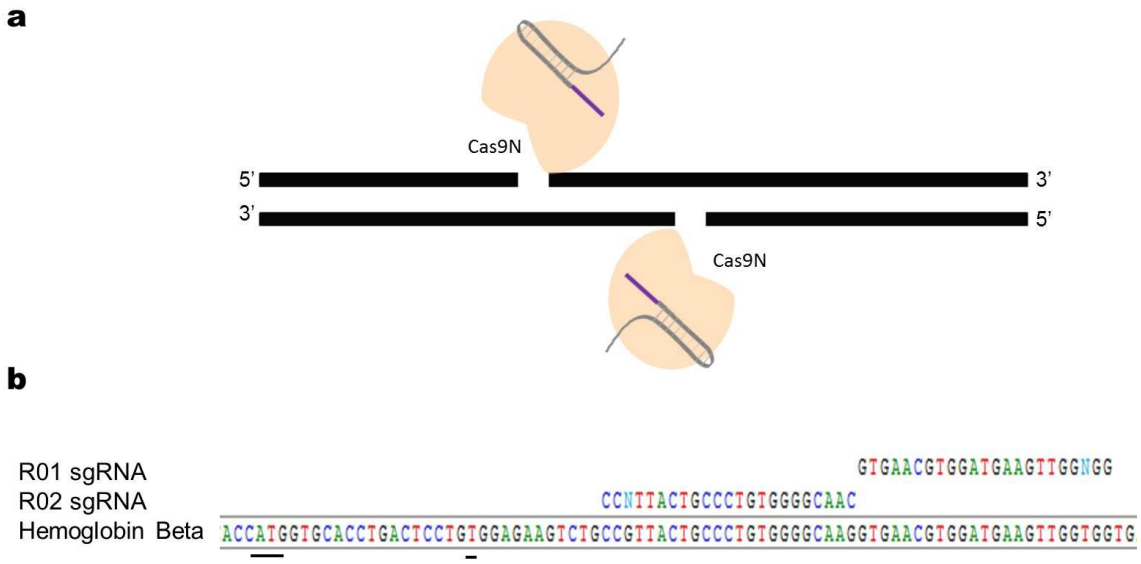


Figure 31: CRISPR/Cas9 nickase system targeting the *HBB* gene

(a) Schematic of a pair of CRISPR/Cas9 nickases activating a staggered DSB on opposite strands of the DNA. Each Cas9 nickase is guided by two different sgRNAs. **(b)** Targeting sequences for the R01 and R02 CRISPR sgRNA aligned to the *HBB* locus. The CRISPR sgRNAs are shown complementary to the reverse stand and is listed to the right and left of the PAM sequence for the R02 and R01 CRISPR sequences respectively. The ATG start codon and the sickle cell mutation are underlined. The A, T, C, and G nucleotides are shown in green, red, blue, and black respectively for clarity.

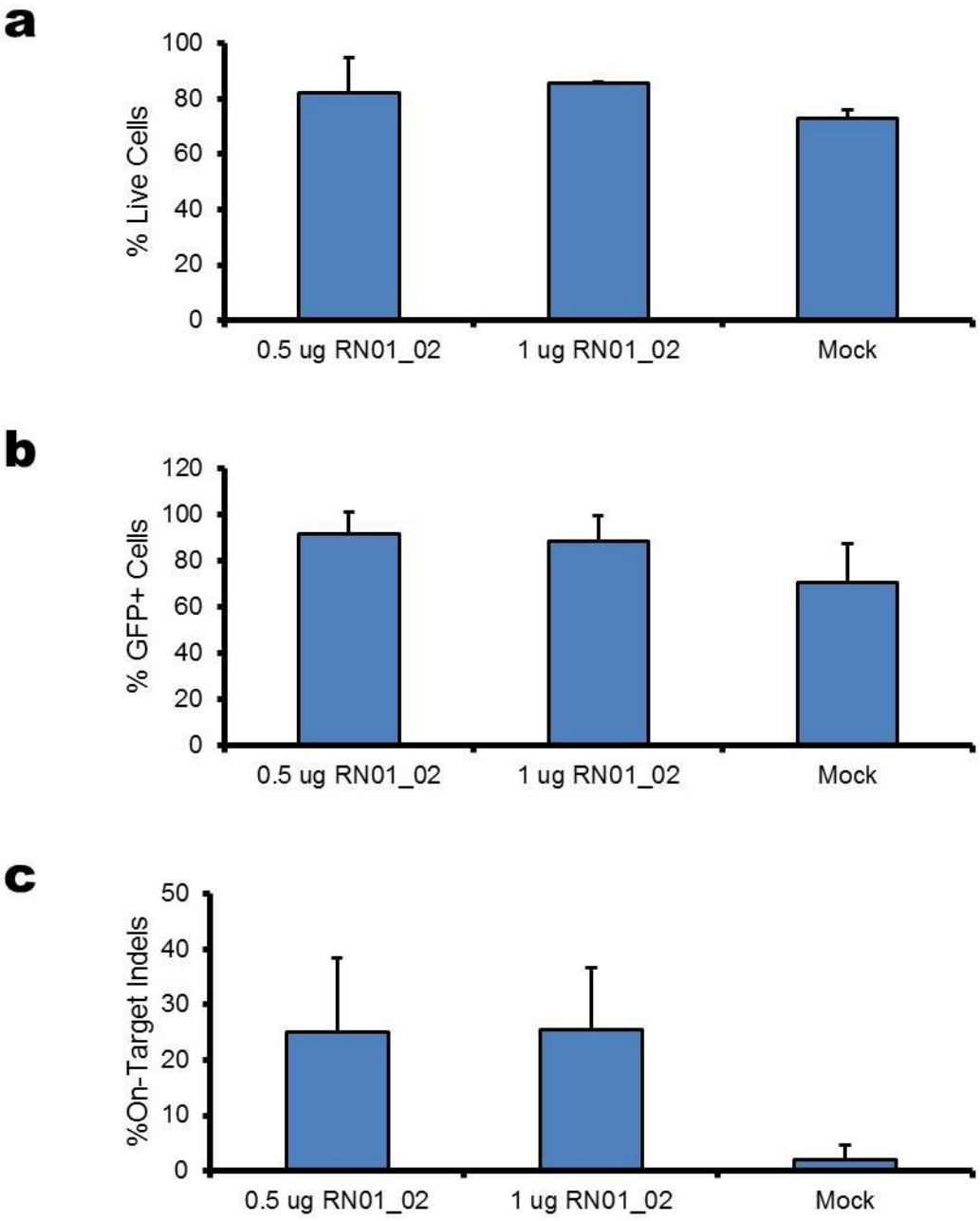


Figure 32: K562 cells nucleofected with R01/R02 CRISPR/Cas9 nickase system

(a) Percentage of live cells, **(b)** GFP+ cells, and **(c)** on-target indels in cells nucleofected with 0.5 ug and 1 ug RN01_02 CRISPR/Cas9 nickase + 400 ng pmaxGFP plasmid DNA. The percentage of live and GFP+ cells were measured at 24 hours after nucleofection. The on-target indels was assessed using TIDE analysis at 3 days after nucleofection. Bars represent the statistical mean ± standard deviation.

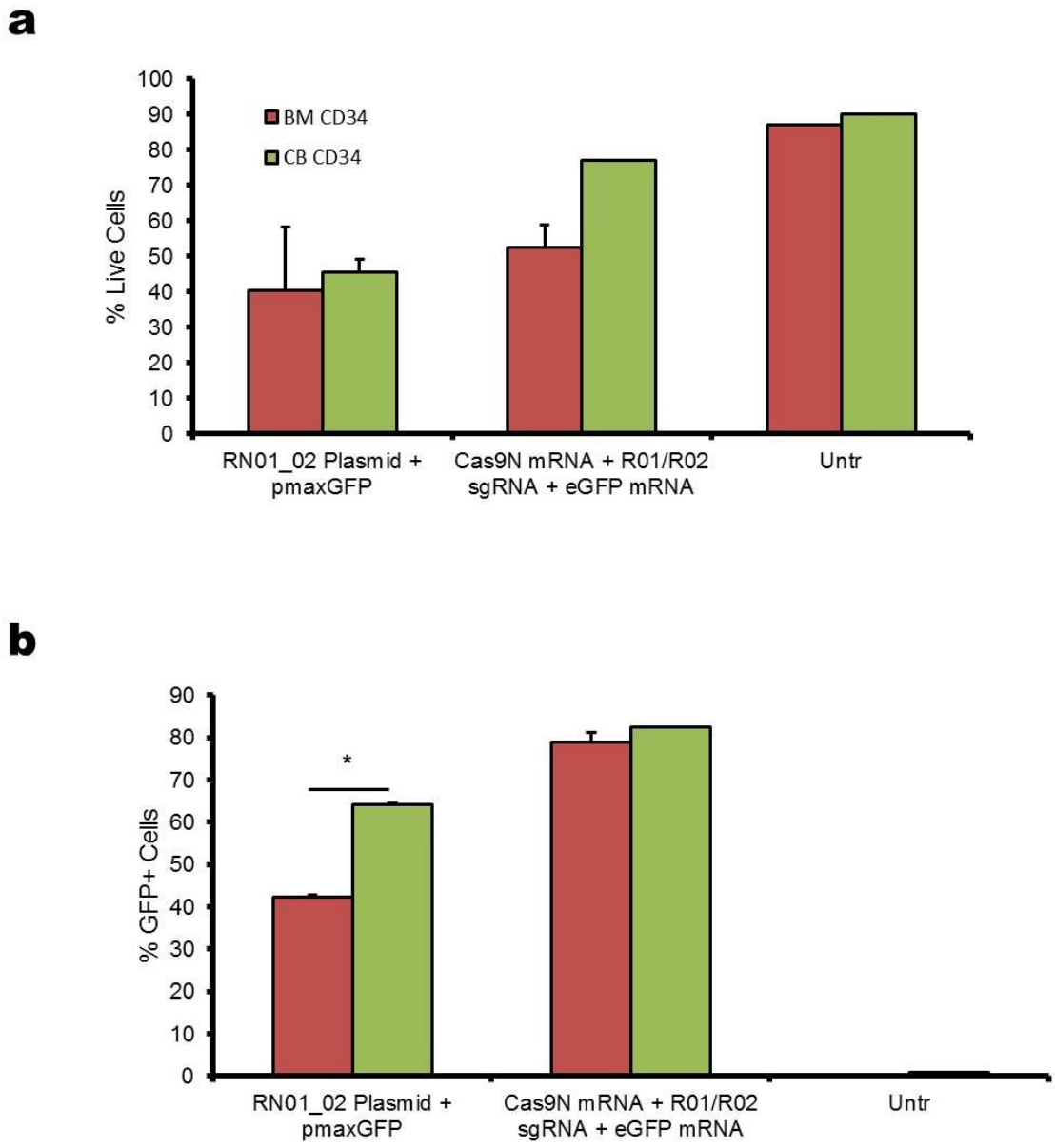


Figure 33: CD34$^+$ cells nucleofected with R01/R02 CRISPR/Cas9 nickase plasmid DNA and mRNA

BM and CB CD34$^+$ cells were nucleofected with 0.5 μg R01/R02 CRISPR/Cas9 nickase plasmid DNA or 1 μg Cas9 nickase mRNA and 5 μg each R01 and R02 sgRNAs. 0.4 μg pmaxGFP plasmid DNA or 0.5 μg of egFP mRNA were added to estimate the nucleofection efficiency. **(a)** The percentage of live cells and **(b)** GFP$^+$ cells at 24 hours after nucleofection. Bars represent the mean percentage of live cells or GFP$^+$ cells ± standard deviation. Asterisks represent statistical significance between the BM and CB CD34$^+$ cell data, where $p<0.05$.

Drug selection of gene modified K562 cells nucleofected with sickle β-globin donor and CRISPR/Cas9-derived nucleases and nickases

Next we investigated HDR-mediated gene modification in cells nucleofected with the CRISPR/Cas9 nucleases and nickases along with donor template. The donor applied in the study contained 5' and 3' homologous sequence from the *HBB* gene, and the sickle cell *HBB* cDNA, which can be used for converting wild type to sickle *HBB* **(Figure 34a)**. The βS-donor contained the P140K mutant O^6-methylguanine methyltransferase (P140K-MGMT) drug selection cassette and mCitrine gene for enrichment and quantification of gene modified cells **(Figure 34a)**. The P140K-MGMT confers resistance to the combined treatment of O^6 benzylguanine (BG) and chemotherapeutic agents, such as 1,3 bis(2-choloroethyl)-1-nitrosourea (BCNU)[187]. The study by Zielske et al showed high levels of *in vivo* gene selection and reconstitution of human $CD34^+$ cells transduced using lentiviral vectors containing the P140K-MGMT gene using nonmyeloablated NOD/SCID mice[188]. The Zielske study showed that engraftment of human HSCs in a nonmyelablated NOD/SCID host required drug treatment. Only after the second round of BG/BCNU treatment was there engraftment by the human cells in the mice, which consisted of 20% of the bone marrow cells[188]. Substantial *in vivo* enrichment of gene modified and drug resistant HSCs has also been observed in nonmyeloablated β-thalassemia mice with a 66% success rate[189]. We first tested the drug selection enrichment approach using the βS donor while comparing the R01/R02 Cas9n and the R02 CRISPR/Cas9 system for HDR-mediated gene modification in K562 cells. The procedure for the drug selection experiments in K562 cells is shown in **Figure 34b**. When we nucleofected K562 cells with the βS donor and R01/R02 Cas9n, we observed low levels of mCitrine positive cells (1.9%) without drug selection using BG/BCNU **(Figure 35a).** We observed higher levels of $mCitrine^+$ cells

with the R02 CRISPR/Cas9 (6.4%) prior to BG/BCNU treatment **(Figure 35b)**. With each round of drug selection we observed successive enrichments of gene modified cells for both R01/R02 Cas9n and R02 CRISPR/Cas9 **(Figure 35)**. By the third round of drug selection, we observed a 16-fold enrichment of mCitrine positive cells with the R01/R02 Cas9n and 4-fold enrichment with the R02 CRISPR/Cas9 **(Figure 35).** On the contrary, we observed undetectable levels of mCitrine^{+} cells in controls nucleofected with βS donor only **(Figure 35c)**, which is evidence that the drug selection was specific for cells positive for HDR-mediated gene modification. The higher fold increase in gene modified cells for the R01/R02 Cas9n pair relative to the R02 CRISPR/Cas9 is likely due to differences in drug selection efficiency during the experiments. The higher level of mCitrine positive cells for the samples nucleofected with the βS donor and R02 CRISPR/Cas9 relative to the R01/R02 Cas9n prior to drug treatment is an indication that higher HDR is obtained with the CRISPR/Cas9 system[133].

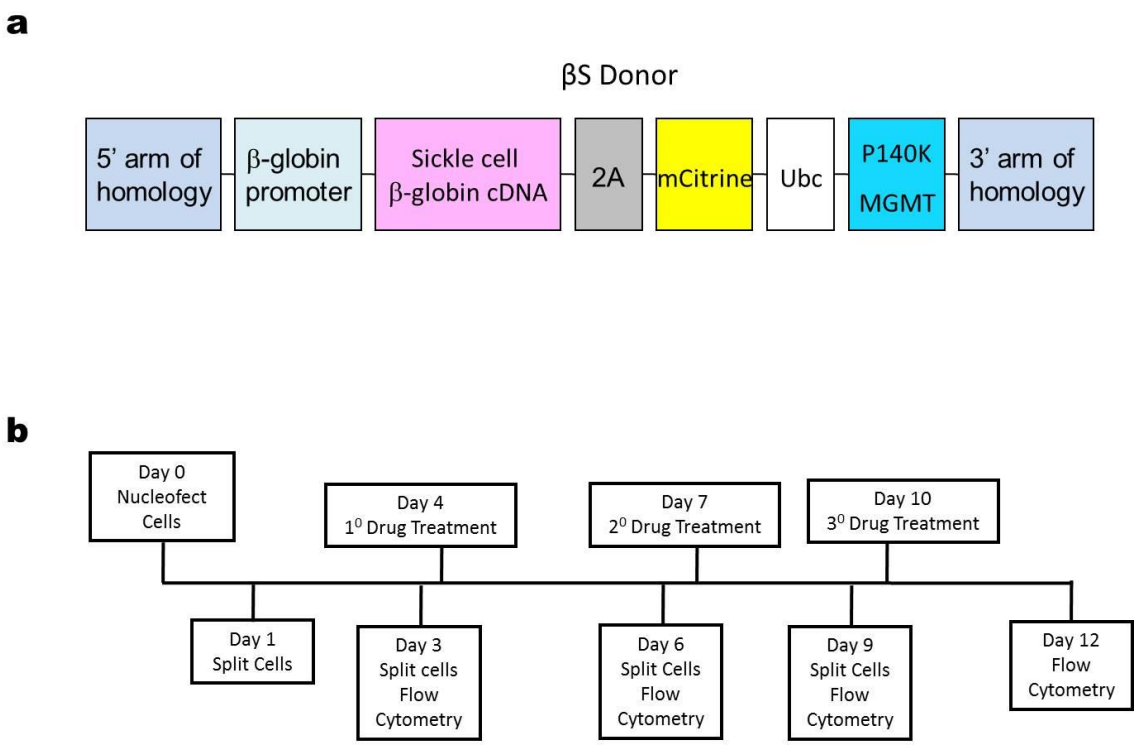


Figure 34: Drug selection for cells positive for HDR-mediated gene modification

(a) Schematic of the βS donor plasmid construct containing the sickle cell β-globin cDNA, mCitrine, and P140K MGMT drug selection cassette. When repair of a DSB is mediated via the HDR pathway, the entire construct will be inserted into the *HBB* gene. BG/BCNU drug selection was used to enrich for cells that were positive for mCitrine fluorescence, which was a marker for HDR. **(b)** Overview of the drug selection experiments in K562 cells. Cells were nucleofected with 2 μg R02 CRISPR/Cas9 nuclease or 1 μg each of R01 and R02 CRISPR/Cas9 nickase pair encoding plasmid DAN (1x10^6 cells per reaction) along with 8 μg of βS donor and were subsequently subjected to 3 rounds of drug selection. Flow cytometry was used to quantify the percentage of mCitrine+ positive cells.

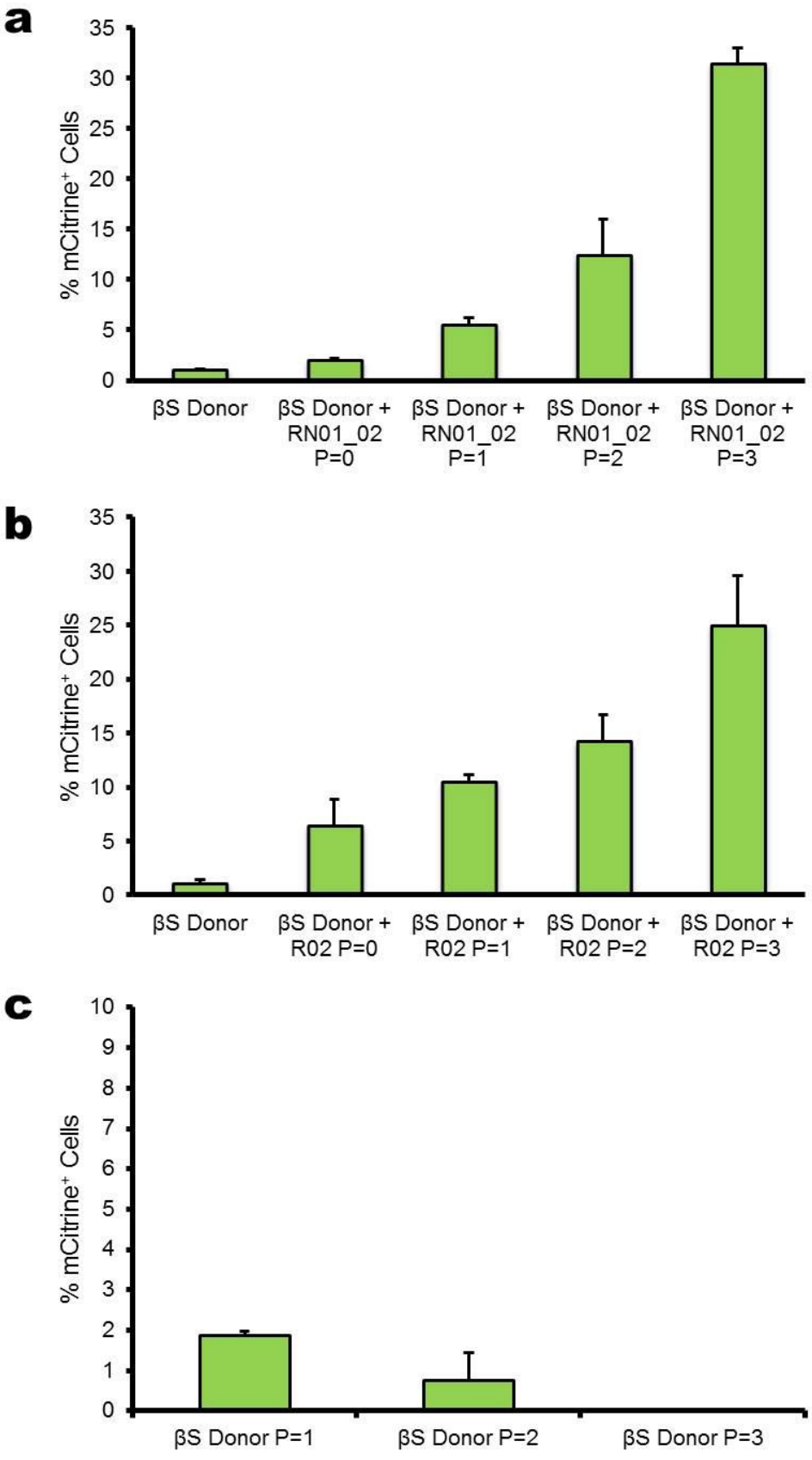

a
% mCitrine+ Cells
35
30
25
20
15
10
5
0
βS Donor
βS Donor + RN01_02 P=0
βS Donor + RN01_02 P=1
βS Donor + RN01_02 P=2
βS Donor + RN01_02 P=3
b
% mCitrine+ Cells
35
30
25
20
15
10
5
0
βS Donor
βS Donor + R02 P=0
βS Donor + R02 P=1
βS Donor + R02 P=2
βS Donor + R02 P=3
c
% mCitrine+ Cells
10
9
8
7
6
5
4
3
2
1
0
βS Donor P=1
βS Donor P=2
βS Donor P=3

Figure 35: Drug selection of gene modified cells nucleofected with the *HBB* CRISPR/Cas9 nickase and nucleases in K562 cells

K562 cells were nucleofected with the βS donor plasmid along with **(a)** R01 and R01 CRISPR/Cas9 nickase plasmid DNA or **(b)** R02 CRISPR/Cas9 nuclease plasmid DNA. **(c)** Cells nucleofected with βS donor plasmid only is shown as a control. Samples nucleofected with donor and the nickase or nuclease were subjected to 3 pulses of drug selections (P = 1, 2, and 3) using BG/BCNU. Bars represent the mean percentage of mCitrine positive cells ± standard deviation, where n = 3.

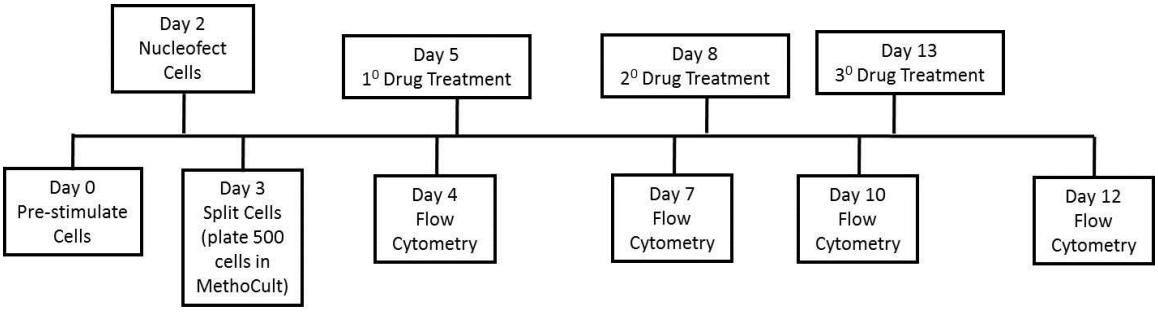


Figure 36: Overview of the drug selection experiments in CB CD34$^+$ cells

Cells (8x10^4 cells per reaction) were nucleofected with 1 μg Cas9 mRNA and 10 μg R02 sgRNA or 7.5 μg Cas9 protein and 10 μg R02 sgRNA along with 0.5 μg βS donor and were subsequently subjected to 3 rounds of drug selection. On day 3 after nucleofection, 500 cells were plated in MethoCult media. Flow cytometry was used to quantify the percentage of mCitrine+ positive cells.

HDR-mediated gene modification in CD34$^+$ cells with drug treatment

We next investigated the application of drug selection to enrich for CD34$^+$ cells gene modified using the R02 CRISPR/Cas9 system and βS donor. Because the HDR repair of nuclease induced DSBs preferentially occurs during the S-phase and is less likely to occur in quiescent CD34$^+$ cells[190], drug selection may increase the probability of detecting HDR in CD34$^+$ cells. CB CD34$^+$ cells were nucleofected with βS donor along with R02 CRISPR/Cas9 mRNA or RNP, were cultured in erythroid media, and treated

with 3 rounds of drug treatment using BG/BCNU **(Figure 36)**. At 24 hours after nucleofection, $CD34^+$ cells were plated in H4434 MethoCult (Stem Cell Technologies) media (500 cells per plate) and cultured for 14 days. The erythroid and myeloid CFC units for each sample was quantified **(Figure 37a)**. We observed variable amounts of CFC units, but the difference was not significant for each condition **(Figure 37a)**, an indicator that the delivery of donor template with and without R02 CRISPR/Cas9 as mRNA or protein did negligibly affect the hematopoietic potential of $CD34^+$ cells *in vitro*. Furthermore, neither DNA nor mRNA substrate delivered using nucleofection had an adverse effect on the HSC functionality *in vitro* and the number of CFC units plated was consistent with[120]. We only observed detectable, but low levels of mCitrine positive cells after the second round of BG/BCNU drug selection **(Figure 37b)**. The delivery of donor template along with R02 CRISPR/Cas9 mRNA or RNP did not lead to the enrichment of mCitrine positive CD34+ cells **(Figure 37b)** as observed with the K562 cells **(Figure 35)**, an indicator that we could not detect HDR mediated gene modification in the $CD34^+$ cells. The absence gene modification was not surprising given that HDR occurs rarely in CD34+ cells. Furthermore, the βS donor template does not have a built in promoter to enable robust mCitrine gene expression and is, instead, driven by the endogenous ^{HBB} locus, which has low levels of expression in $CD34^+$ cells.

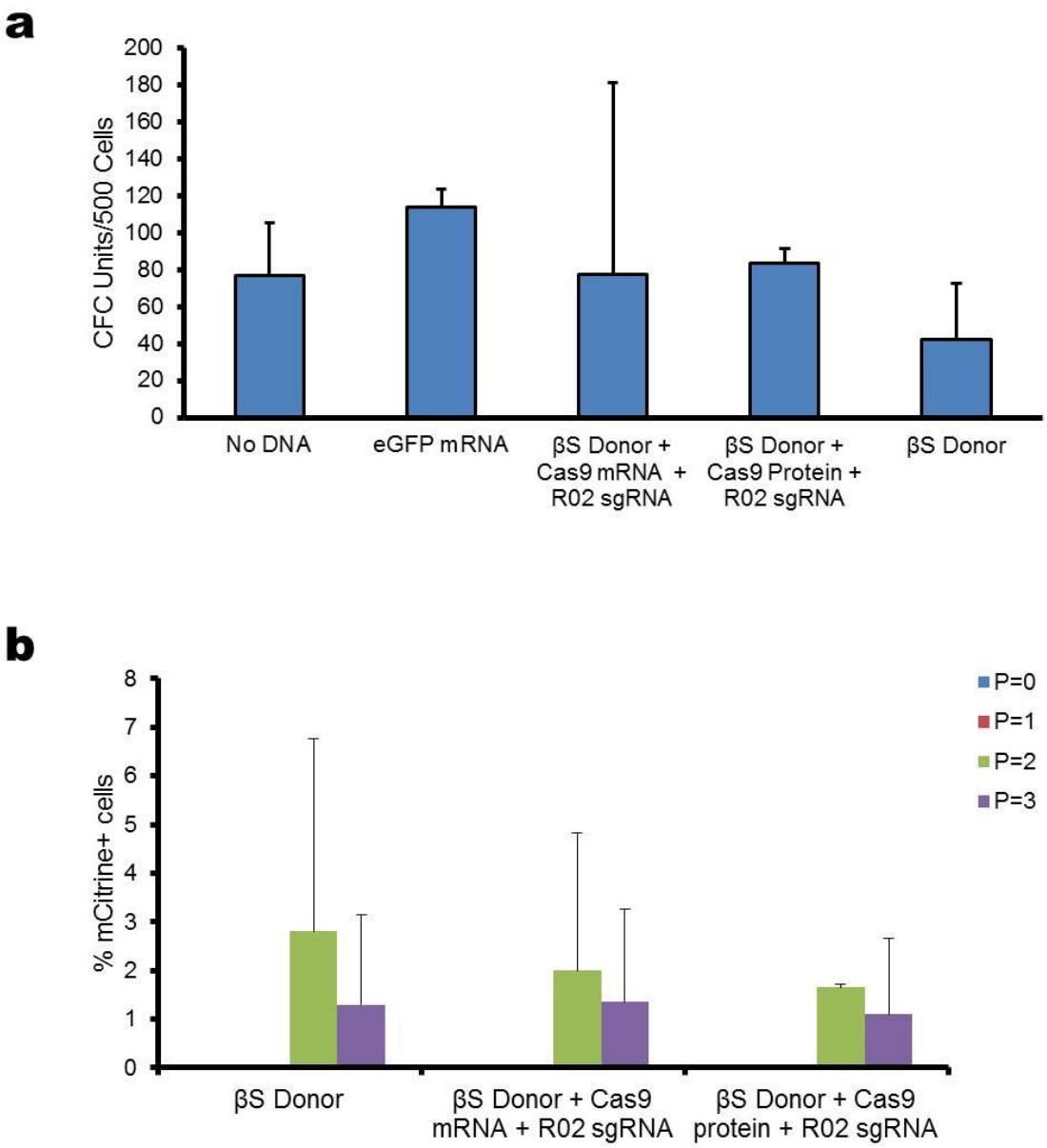


Figure 37: CB CD34^{+} cells nucleofected with βS donor along with R02 CRISPR/Cas9 mRNA or RNP

(a) The number of CFC units measured from 500 cells plated at 24 hours after nucleofection. **(b)** The percentage of mcitrine positive cells after 3 pulses of drug selection. Bars represent the mean CFC units or percentage of mCitrine positive cells ± standard deviation, n = 2.

5.3 Discussion

CRISPR/Cas9 nucleases are promising tools for therapeutic gene correction because of their high cleavage activity and their simpler design compared to other nuclease platforms. However, the major disadvantage of CRISPR/Cas9 systems is that they are associated with high off-target activity[99-101]. Cas9n provide an alternative that can reduce the off-target effects, while enabling high cleavage and gene targeting activity provided through a sgRNA pair[133]. One challenge for applying these systems in $CD34^+$ cells is obtaining high levels of nuclease or nickase expression and sufficient amounts of donor for HDR-mediated gene modification. In this study, we addressed the challenge of delivery by optimizing the *in vitro* culture and nucleofection of $CD34^+$ cells. We compared the delivery of the *HBB*-aiming R02 CRISPR/Cas9 and R01/R02 Cas9n pair for cleavage activity in $CD34^+$ cells derived from BM and CB, and investigated the application of P140K-MGMT and drug selection to enrich for gene modified $CD34^+$ cells.

Our results indicate that culturing $CD34^+$ cells in expansion media for 48 hours prior to nucleofection was critical for obtaining high levels of cell viability and nucleofection efficiency **(Figure 28)**. We found that the CB $CD34^+$ cells consistently showed higher gene expression compared to the BM derived $CD34^+$ cells, which may be explained by differences in the activities of transcriptional and translational machinery. $CD34^+$ cells nucleofected with GFP plasmid DNA showed lower gene expression compared to mRNA **(Figure 33)** consistent with lower transcriptional activity in these cells. We found that delivering the R02 CRISPR/Cas9 mRNA or RNP (for BM $CD34^+$ cells) provided high levels of indels, whereas there was undetectable activity when nucleases were delivered as plasmid DNA **(Figures 30)**. The considerable variability observed in BM $CD34^+$ cells nucleofected with Cas9 RNP indicates that the amount of sgRNA and Cas9 protein delivered into the cells likely was inappropriate. We also

observed undetectable levels of indels by the R01/R02 Cas9n pair delivered as mRNA and RNP, likely because of the doses used in the nucleofection reactions. Future studies should focus on further optimizing the amount of the R02 Cas9 RNP and the R01/R02 nickases to obtain high levels of cleavage activity in $CD34^+$ cells. These results suggest that CRISPR/Cas9 systems encoded on plasmid DNA is not effective for obtaining high levels of nuclease activity, but, rather, future studies should apply RNP or mRNA as delivery approaches in $CD34^+$ cells.

Drug selection of gene modified cells using the P140K-MGMT cassette has many clinical applications, including gene correction[187]. We showed that the delivery of donor template encoding P140K-MGMT along with the R02 CRISPR/Cas9 and R01/02 double Cas9n can be used to enrich for gene modified K562 cells **(Figure 35)**. Further optimization of the drug selection protocol may help to enhance the enrichment observed in this study. We did not observe any gene modification in $CD34^+$ cells, which is likely because of the donor template and assay design. The gene modified cells was estimated by the expression of the fluorescent marker mCitrine, driven by the endogenous *HBB* promoter, which has low activity in K562 cells or quiescent $CD34^+$ cells. Future studies should focus on developing a donor template that will provide high levels of HDR in $CD34^+$ cells. Single stranded oligonucleotides (ssOGN) may represent a more efficient design for a donor template to use in $CD34^+$ cells. In the study by Hoban et al, the delivery of various different ssOGN designs provided high levels of gene modification in $CD34^+$ cells[121].

5.4 Materials and Methods

HBB CRISPR/Cas9 nuclease and nickases and βS donor constructs

The R01 and R02 Cas9n were prepared by cloning the 20-base guide sequences into a pX335-U6-SpCas9n (D10A) plasmid using the BbsI restriction site. The 20-base target sequence (underlined) following the 3-base PAM for the R02 construct: 5'-G**T**GGAGAAGTCTGCCGTTACTGCCCTGTGGGGCAAC-3'. Following the R01 target is the complementary sequence for the R01 CRISPR (underlined) following the 3-base PAM: CCNCCAACTTCATCCACGTTCAG. The nuclease target sites are in close proximity to the sickle cell mutation (bold) in codon 6 of the *HBB* gene. Synthetic mRNA for Cas9n, Cas9, R01, R02 sgRNA, and eGFP was purchased from Tri-Link BioTechnologies. The βS donor was a gift from Dr. Matthew Porteus at Stanford University. Cas9 purified protein was provided by Dr. Charles Gersbach at Duke University. Cas9 RNP complexes were formed by incubating the sgRNA and purified Cas9 protein for 10 minutes at room temperature.

***In vitro* culture and CFC assay in CD34$^+$ cells**

Cryopreserved CD34$^+$ cells were purchased from Stem Cell Technologies. The erythroid differentiation media consisted of StemSpan SFEM medium supplemented with human cytokines: EPO 3 units mL^{-1}, SCF 25 ng mL^{-1}, IL-3 10 ng mL^{-1}, and IL-6 ng mL^{-1}; all purchased from R&D systems. The expansion medium consisted of StemSpan SFEM medium supplemented with human cytokines: SCF 100 ng mL^{-1}, IL-3 20 ng mL^{-1}, TPO 100 ng mL-1, and Flt3-L 100 ng mL^{-1}; all purchased from Peprotech. CD34$^+$ cells shown in Figures 27 and 28 were nucleofected (2 x 10^5 cells per reaction) immediately after thawing. In proceeding experiments, 2.5 x 10^5 cells per mL were stimulated in expansion media for 48 hours and then nucleofected (8 x 10^4 cells per reaction) with specified amounts of plasmid DNA, mRNA, and proteins using the P3 Primary Cell 4D-Nucleofector X kit (Lonza) and program EO-100. Cells cultured *in vitro* were maintained at a density of 1-5 x 10^5 cells mL^{-1} for cultures proceeding longer than 3 days. The

transfection efficiency was estimated by the percentage of GFP positive cells using an Accuri C6 flow cytometer (BD Biosciences, Franklin Lakes, NJ). The cell viability was estimated using trypan blue exclusion dye and an automated cell counter (BioRad, Hercules, CA). For the CFC assay, 500 cells per plate were seeded at 24 hours after nucleofection in H4434 MethoCult (Stem Cell Technologies), a methylcellulose-based medium with human recombinant cytokines to support the growth of progenitor colonies of the erythroid and myeloid lineages. At 14 days after plating, the total CFC units were enumerated according to the manufacturer's instructions.

***In vitro* culture and nucleofection into K562 cells**

K562 cells (ATCC, Manassas, VA) were grown in RPMI 1640 (Hyclone, Logan, UT) supplemented with 10% fetal bovine serum and 2 mM L-glutamine (Invitrogen Life Technologies, Grand Island, NY). K562 cells were nucleofected using SF cell line 4D-Nucleofection kit (Lonza) according to the manufacturer's protocol. The transfection efficiency was estimated by the percentage of GFP positive cells using an Accuri C6 flow cytometer. The cell viability was estimated using trypan blue exclusion dye and an automated cell counter.

Analysis of on- and off-target indels

Cleavage activity was quantified in pooled nucleofected K562 and $CD34^+$ cells. The off-target activity for R02 CRISPR/Cas9 was measured in the *GRIN3A* locus [99]. The genomic DNA was harvested using QuickExtract DNA extraction solution (Epicenter Biotechnologies, Madison, WI) and subjected to PCR amplification of on- and off-target loci using primers listed in Chapter 4 **Supplementary Table 5**. All PCR reactions in 50 μL volume consisting of 1.5 μl genomic DNA were performed using AccuPrime Taq DNA High Fidelity Polymerase kit (Invitrogen) according to the manufacturer's protocol. For

the T7E1 analysis of off-target indels, 200 ng of purified GRIN3A PCR product supplemented with 1X Accuprime buffer II were processed using cycles of melting and re-annealing (95°C for 10 min, 95-85°C at -2°C/s, 85-25°C at -0.1°C/s). T7EI (NEB, Ipswich, MA) was added to a final concentration of 250 units/mL and incubated at 37°C for 60 minutes for digestion. Reactions were resolved on a 2% agarose Tris-EDTA gel stained with ethidium bromide. The percentage of off-target indels was estimated by densitometry analysis using ImageJ software and the equation: 100 X (1-(1-fraction cleaved) $^{1/2}$) as described in[173]. For quantifying the on-target indels using TIDE [136], purified HBB PCR product were Sanger-sequenced using the forward and reverse primers separately, and subsequently analyzed using the online TIDE software (accessible at http://tide.nki.nl).

Drug selection of gene modified K562 and cord blood CD34$^+$ cells

For experiments in K562 cells, cells were subjected to drug treatments on Day 4, 7, and 10 after nucleofection **(Figure 34b)**. K562 cells at a density of 1 x 10^5 cells were incubated for 1 hour with 50 uM O^6-BG (Sigma), and followed a by second 1 hour incubation with 40 uM BCNU (Sigma) at 37°C. Treated cells were washed with in pre-warmed PBS buffer two times and resuspended in fresh culture media. A similar procedure was used for CD34$^+$ cells with the drug treatments completed on different days **(Figure 36)**. The percentage of cells with mCitrine expression, a marker for gene modification, was estimated using an Accuri Flow cytometer.

CHAPTER 6: CONCLUSIONS AND FUTURE PERSPECTIVES

Precision gene editing tools are monumental scientific discoveries having the potential to usher a new era of curative therapeutics involving gene correction for hemoglobinopathies. Compared to gene therapy, gene correction through gene editing reagents is potentially safer and more effective because it involves the replacement of the aberrant HBB with the wild type or anti-sickling variant, eliminating the risks associated with random gene insertion using viral vectors[43]. The gene therapy trials for β-thalassemia and SCD is still in progress, but reports so far have shown correction in some elements of the disease phenotype, including the RBC counts and hemoglobin levels, obviating the need for RBC transfusion[50,51]. The observation of clonal expansion in cells due to semi-random lentiviral-induced gene disruption shows that risks for gene therapy cannot be overlooked. Recently *HBB* ZFNs and donor template was used to correct the SCD mutation in patient derived BM $CD34^+$ cells, providing evidence of the therapeutic potential of the gene correction approach[121].

In chapter 1, we reviewed DNA binding protein and RNA-guided nucleases and discussed the pros and cons for each nuclease platform. The shared challenge for gene editing tools is designing the nuclease and donor template to achieve high levels of HDR-mediated gene modification and specificity, particularly within clinically relevant cells. Because of the difficulties of designing a nuclease that does not have any off-target effects, it is important that the off-target sites for a specific nuclease design are thoroughly assessed to ensure that they will not have any adverse effects in a patient. The CRISPR/Cas9 system has the advantage of a simpler design, but it is associated with substantial off-target activities[99-101]. In Chapter 3, we showed that the R02 CRISPR/Cas9 nuclease has higher activity compared to TALENs, therefore, RNA-

guided nucleases should still be investigated as a potential tool for genome editing of the SCD mutation. Future studies should investigate whether CRISPR/Cas9 systems can induce higher frequencies of gene targeting compared to the ZFN pair used in [121]. However, strategies will be needed to curb the off-target effects of the CRISPR/Cas9 nuclease.

In Chapter 3, we developed a microinjection method for K562 cells using glass microcapillaries and retronectin immobilization of cells on a surface. This work is significant because it involves adapting microinjection, traditionally used for adherent cells or, more frequently, oocytes, and using it to deliver gene editing tools into human somatic cells. For the first time, we showed that microinjection can be used to deliver gene editing tools and obtain high levels of gene editing (both NHEJ and HDR) in a human somatic cell. We showed that the advantage of microinjection is the precise control of the amount of the delivered gene editing tools in cells, and uniformity in delivery between single cells; enabling amplification of the effects from the delivered material in the cells. Evidence of the control provided with injection is the high frequencies of K562 cells observed that contained tri-allelic indels after injection with R02 CRISPR/Cas9 and L4-R4 TALENs without drug selection, which was considerably higher than the frequencies observed in ZFNs nucleofected into mammalian cell lines [191]. Future work should focus on approaches that can be used to further optimize the levels of gene targeting in cells microinjection with nuclease and donor template. In our experiments we only delivered nucleases and donor as a plasmid DNA. It will be interesting to investigate gene targeting by nucleases microinjected as mRNA or proteins as well as compare ssOGN and plasmid DNA donor, and different ratios of donor to nucleases. Microinjection would be ideal for these experiments since the delivery efficiency would be the same for any delivered substance, including all of the components in the injection solution, as long as the injection parameters are fixed.

In chapter 4, we used nucleofection to show that high levels of gene targeting is feasible for the R02 CRISPR/Cas9 system delivered as an mRNA and RNP complex. The RNP and mRNA nucleases provided higher specificity compared to the plasmid DNA [179]. Interestingly, the RNP delivery provided even more specificity compared to mRNA, while the frequency of HDR-mediated gene modification was roughly the same for each delivery approach. It will be interesting to investigate whether gene targeting by the RNP delivery strategy can be enhanced by improving the complex formation between the Cas9 and sgRNA, further optimizing the amounts used for delivery, or by using better quality Cas9 proteins.

A second challenge with therapeutic gene correction is obtaining high levels of nuclease expression and amount of donor template in quiescent $CD34^+$ cells. In chapter 5, we developed *in vitro* culture and nucleofection protocols and showed that they provided high levels of nucleofection efficiency and reasonable cell viability. Although we observed high levels of gene expression from both pmaxGFP plasmid DNA and eGFP mRNA, we only observed consistent nuclease activity in $CD34^+$ cells nucleofected with R02 CRISPR/Cas9 mRNA and not plasmid DNA. This difference suggests that expression of active nucleases is more efficient with mRNA delivery. The delivery of the CRISPR/Cas9 system as RNPs provided promising results, but further work is needed to optimize its activity in $CD34^+$ cells. Lastly, we explored drug selection using donor containing the P140K-MGMT cassette as a strategy to enrich for nuclease induced gene modified $CD34^+$ cells *in vitro*. The P140K-MGMT drug selection marker is attractive since it has already been shown to be safe and is currently used in clinical trials for enhancing chemotherapy[192]. We showed that the drug selection strategy was feasible in K562 cells. However, an optimized procedure and donor template would be necessary to show proof-of-principle of the drug selection approach in human $CD34^+$ cells.

APPENDIX: SUPPLEMENTARY INFORMATION

Chapter 3 Supplementary Information

Table 1: Sequences of primers used to amplify the endogenous genes for T7E1 mutation detection assays

Gene	Sequence
HBB-F	AGGCACCGAGCACTTTCTTGCC
HBB-R	ACCCTGTGGAGCCACACCCTA
HBD-F	GAGGTTGTCCAGGTGAGCCAGGCCATCAC
HBD-R	CTGCTGAAAGAGATGCGGTGGGGAGATATGTA
GRIN3A-F	GTTTCTAAGAGCGGTGGCTCTCA
GRIN3A-R	CTGCCCCATCTATGCTTGGGA

Table 2: Sequences of primers used for PCR confirmation of HR-mediated GFP integration in the *HBB* locus

Gene	Sequence
GFP Integration-F	CGACAACCACTACCTGAGCA
GFP Integration-R	AGCAGAATGGTAGCTGGATTG
HBB Control-F	TGGTGGTGAGGCCCTGGGCAGGTTG
HBB Control-R	TAAAAGCAGAATGGTAGCTGGATT

Table 3: Sequences of primers used for amplifying *HBB* locus in single cell clones for T7E1 assay and Sanger sequencing.

Primer	Sequence
Beta 4F-Tag1	atcgAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag2	cagaAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag3	gctaAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag4	tgacAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag5	acgtAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag6	catgAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag7	gtgaAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag8	tagcAGGCACCGAGCACTTTCTTGCC
Beta 4F-Tag9	agtcAGGCACCGAGCACTTTCTTGCC
Beta 4R-Tag1	atcgACCCTGTGGAGCCACACCCTA
Beta 4R-Tag2	cagaACCCTGTGGAGCCACACCCTA
Beta 4R-Tag3	gctaACCCTGTGGAGCCACACCCTA
Beta 4R-Tag4	tgacACCCTGTGGAGCCACACCCTA
Beta 4R-Tag5	acgtACCCTGTGGAGCCACACCCTA
Beta 4R-Tag6	catgACCCTGTGGAGCCACACCCTA
Beta 4R-Tag7	gtgaACCCTGTGGAGCCACACCCTA
Beta 4R-Tag8	tagcACCCTGTGGAGCCACACCCTA
Beta 4R-Tag9	agtcACCCTGTGGAGCCACACCCTA

Unique barcode used to identify each clone is shown in lowercase.

Table 4: Analysis of clones with on- and off-target activity

	R02	**L4-R4**
Total Clones analyzed by T7E1	78	53
% Clones with on- and off-target activity	32.1	13.2
% Clones with on-target activity only	14.1	15.1
% Clones with off-target activity only	6.4	11.3

Clones derived from single cells injected with R02 CRISPR/Cas9 or L4-R4 TALENs were analyzed for on- and off-target activity using the T7E1 mismatch assay. The percentage of clones having on-target indels with and without off-target indels, or off-target indels only is shown in the table for each nuclease.

Chapter 4 Supplementary Information

Table 5: Sequences of primers used to amplify the endogenous genes for T7E1 and TIDE mutation detection assays

Gene	Sequence
HBB-F	AGGCACCGAGCACTTTCTTGCC
HBB-R	ACCCTGTGGAGCCACACCCTA
GRIN3A-F	GTTTCTAAGAGCGGTGGCTCTCA
GRIN3A-R	CTGCCCCATCTATGCTTGGGA

Table 6: Sequences of primers used to amplify the targeted *HBB* gene for the RFLP assay

Gene	Sequence
RFLP-R	GCAATCATTCGTCTGTTTCCCATTC
RFLP-F	CTGGAGACGCAGGAAGAGATCC

HBB EcoRI Donor Sequence

The HBB donor sequence consisting of 1049 bp is shown in uppercase with the EcoRI site highlighted and underlined in Red. The mismatches in the EcoR1 site reduces binding and cleavage by the R02 CRISPR. The remainder of the sequence is the puc18 plasmid DNA.

AGTGCATCAACTTCTTATTTGTGTAATAAGAAAATTGGGAAAACGATCTTCAATATGC

TTACCAAGCTGTGATTCCAAATATTACGTAAATACACTTGCAAAGGAGGATGTTTTTA

GTAGCAATTTGTACTGATGGTATGGGGCCAAGAGATATATCTTAGAGGGAGGGCTG

AGGGTTTGAAGTCCAACTCCTAAGCCAGTGCCAGAAGAGCCAAGGACAGGTACGG

CTGTCATCACTTAGACCTCACCCTGTGGAGCCACACCCTAGGGTTGGCCAATCTAC

TCCCAGGAGCAGGGAGGGCAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAG

CCATCTATTGCTTACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACA

GACACCATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTG**GAATTC**GGG

GCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAAG

GTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAG

ACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTCCCACC

CTTAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGA

TCTGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGA

AAGTGCTCGGTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCAC

CTTTGCCACACTGAGTGAGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACT

TCAGGGTGAGTCTATGGGACGCTTGATGTTTTCTTTCCCCTTCTTTTCTATGGTTAA

GTTCATGTCATAGGAAGGGGATAAGTAACAGGGTACAGTTTAGAATGGGAAACAGA

CGAATGATTGCATCAGTGTGGAAGTCTCAGGATCGTTTTAGTTTCTTTTATTTGCTG

TTCATAACAATTGTTTTCTTTTGTTTAATTCTTGCTTTCgtaatcatggtcatagctgtttcctgtgtga

aattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagc

taactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaa

cgcgcggggagaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcg

gcgagcggtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtg

agcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgac

gagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttccccct

ggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcg

ctttctcaaagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttc

agcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcag

ccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctaca

ctagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaa

caaaccaccgctggtagcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctt

tgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatctt

cacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgct

taatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgata

cgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaat

aaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccg

ggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtc

gtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttag

ctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctctta

ctgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccga

gttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttct

tcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagc

atcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgaca

cggaaatgttgaatactcatactcttccttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttga

atgtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctgacgtc

REFERENCES

1. Weatherall, D.J. The inherited diseases of hemoglobin are an emerging global health burden. *Blood* **115**, 4331-4335 (2010).
2. Modell, B. & Darlison, M. Global epidemiology of haemolgobin disorders and derived service indicators. *Bulletin of the World Health Organization* **86**, 480-487 (2008).
3. Sickle Cell Disease. (Centers for Disease Control and Prevention).
4. Kauf, T.L., Coates, T.D., Huazhi, L., Mody-Patel, N. & Hartzema, A.G. The cost of health care for children and adults with sickle cell disease. *American Journal of Hematology* **84**, 323-327 (2009).
5. Ingram, V.M. Abnormal human haemoglobins III the chemical difference between normal and sickle cell haemoglobins. *Biochimica et Biophysica Acta* **36**, 402-411 (1959).
6. Bridges, K.R. & Pearson, H.A. *Anemias and other red cell disorders*, (McGraw-Hill Companies, New York, NY, 2008).
7. Steinberg, M.H.E. & Embury, S.H.E. *Sickle cell disease basic principles and clinical practice*, (Raven Press, New York, NY, 1994).
8. Sergeant, G.R. & Sergeant, B.E. *Sickle cell disease*, (Oxford Press, New York, NY, 2001).
9. Weatherall, D.J. & Clegg, J.B. *The thalassaemia syndromes*, (Blackwell Science, Malden, MA, 2001).
10. Platt, O.S. *et al.* Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *Journal of Clinical Investigation* **74**, 652-656 (1984).
11. Charache, S. *et al.* Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. *The New England Journal of Medicine* **332**, 1317-1322 (1995).
12. Steinberg, M.H. *et al.* Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia risks and benefits up to 9 years of treatment. *Journal of the American Medical Association* **289**, 1645-1651 (2003).
13. Thomas, E.D. *et al.* Marrow transplantation for thalassaemia. *The Lancet* **320**, 227-229 (1982).
14. Walters, M.C. *et al.* Bone marrow transplantation for sickle cell disease. *The New England Journal of Medicine* **335**, 369-376 (1996).
15. Lucarelli, G. *et al.* Bone marrow transplantation in patients with thalassemia. *The New England Journal of Medicine* **322**, 417-421 (1990).
16. Walters, M.C. *et al.* Impact of bone marrow transplantation for symptomatic sickle cell disease: an interim report. *Blood* **95**, 1918-1924 (2000).
17. Andreani, M. *et al.* Persistence of mixed chimerism in patients transplanted for the treatment of thalassemia. *Blood* **87**, 3494-3499 (1996).
18. Walters, M.C. *et al.* Barriers to bone marrow transplantation for sickle cell anemia. *Biology of Blood and Marrow Transplantation* **2**, 100-104 (1996).
19. Hsieh, M.M., Fitzhugh, C.D. & Tisdale, J.F. Allogenic hematopoietic stem cell transplantation for sickle cell disease the time is now. *Blood* **118**, 1197-1207 (2011).
20. Bone Marrow Transplant. Vol. 2015 (Harvard Medical School).

21. Riviere, I., Dunbar, C.E. & Sadelain, M. Hematopoietic stem cell engineering at a crossroads. *Blood* **119**, 1107-1116 (2012).
22. Paszty, C. *et al.* Transgenic knockout mice with exclusively human sickle hemoglobin and sickle cell disease. *Science* **278**, 876-878 (1997).
23. Trudel, M. *et al.* Towards a transgenic mouse model of sickle cell disease: hemoglobin SAD. *The EMBO Journal* **10**, 3157-3165 (1991).
24. Yang, B. *et al.* A mouse model for beta 0-thalassemia. *Proceedings of the National Academy of Sciences USA* **92**(1995).
25. Skow, L.C. *et al.* A mouse model for β-thalassemia. *Cell* **34**, 1043-1052 (1983).
26. Forrester, W., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G. & Groudine, M. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Research* **15**, 10159-10177 (1987).
27. Tuan, D. & London, I.M. Mapping of DNase I-hypersensitive sites in the upstream DNA of human embryonic epsilon-globin gene in K562 leukemia cells. *Proceedings of the National Academy of Sciences USA* **81**, 2718-2722 (1984).
28. Tuan, D., Solomon, W., Li, Q. & London, I.M. The beta-like-globin gene domain in human erythroid cells. *Proceedings of the National Academy of Sciences USA* **82**, 6384-6388 (1985).
29. Wright, S., Rosenthal, A., Flavell, R. & Grosveld, F. DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. *Cell* **38**(1984).
30. van Assendelft, G.B., Hanscombe, O., Grosveld, F. & Greaves, D.R. The β-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* **56**, 969-977 (1989).
31. Talbot, D. *et al.* A dominant control region from the human β-globin locus conferring integration site-independent gene expression. *Nature: international weekly journal of science* **338**, 352-355 (1989).
32. Grosveld, F., van Assendelft, G.B., Greaves, D.R. & Kollias, G. Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell* **51**, 975-985 (1987).
33. Vigna, E. & Naldini, L. Lentiviral vectors: excellent tools for experimental gene transfer and promising candidates for gene therapy. *The Journal of Gene Medicine* **2**, 308-316 (2000).
34. May, C. *et al.* Therapeutic haemoglobin synthesis in β-thalassaemic mice expressing lentivirus-encoded human β-globin. *Nature* **406**, 82-86 (2000).
35. Imren, S. *et al.* Permanent and panerythroid correction of murine β thalassemia by multiple lentiviral integration in hematopoietic stem cells. *Proceedings From the National Academy of Sciences USA* **99**, 14380-14385 (2002).
36. Person, D.A., Hargrove, P.W., Allay, E.R., Hanawa, H. & Nienhuis, A.W. The degree of phenotypic correction of murine β-thalassemia intermedia following lentiviral-mediated transfer of a human ɣ–globin gene is influenced by chromosomal position effects and vector copy number. *Blood* **101**, 2175-2183 (2003).
37. Rivella, S., May, C., Chadburn, A., Riviere, I. & Sadelain, M. A novel murine model of Cooley anemia and its rescue by lentiviral-mediated human β-globin gene transfer. *Blood* **101**, 2932-2938 (2003).
38. Puthenveetil, G. *et al.* Successful correction of the human β-thalassemia major phenotype using a lentiviral vector. *Blood* **104**, 3445-3453 (2004).
39. Pawliuk, R. *et al.* Correction of sickle cell disease in transgenic mouse models by gene therapy. *Science* **294**, 2368-2371 (2001).

40. Levasseur, D.N., Ryan, T.M., Pawlik, K.M. & Townes, T.M. Correction of a mouse model of sickle cell disease: lentiviral/antisickling β-globin gene transduction of unmobilized, purified hematopoietic stem cells. *Blood* **102**, 4312-4319 (2003).
41. Hacein-Bey-Abina, S. *et al.* A serious adverse event after successful gene therapy for x-linked severe combined immunodeficiency. *The New England Journal of Medicine* **348**, 255-256 (2003).
42. Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669-672 (2000).
43. Hacein-Bey-Abina, S. *et al.* LMO2-associated clonal T-cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415-419 (2003).
44. Imren, S. *et al.* High-level β-globin expression and preferred intragenic integration after lentiviral transduction of human cord blood stem cells. *The Journal of Clinical Investigation* **114**, 953-962 (2004).
45. Abboud, M., Laver, J. & Blau, C.A. Granulocytosis causing sickle-cell crisis. *The Lancet* **351**, 959 (1998).
46. Adler, B.K. *et al.* Fatal sickle cell crisis after granulocyte colony-stimulating factor administration. *Blood* **97**, 3313-3314 (2001).
47. Falzetti, F., Aversa, F. & Minelli, O. Spontaneous rupture of spleen during peripheral blood stem-cell mobilisation in a healthy donor. *Lancet* **353**, 555 (1999).
48. Yannaki, E. *et al.* Hematopoietic stem cell mobilization for gene therapy superior mobilization by the combination of granulocyte-colony stimulating factor plus plerixafor in patients with β-thalassemia major. *Human gene therapy* **24**, 852-860 (2013).
49. Bank, A., Dorazio, R. & Leboulch, P. A phase I/II clinical trial of β-globin gene therapy for β-thalassemia. *Annuals New York Academy of Sciences* **1054**, 308-316 (2005).
50. Payen, E. & Leboulch, P. Advances in stem cell transplantation and gene therapy in the β-hemoglobinopathies. *ASH Education Program Book* **2012**, 276-283 (2012).
51. Cavazzana-Calvo, M. *et al.* Transfusion independence and HMGA2 activation after gene therapy of human [bgr]-thalassaemia. *Nature* **467**, 318-322 (2010).
52. Wang, G.P. *et al.* Dynamics of gene-modified progenitor cells analyzed by tracking retroviral integration sites in a human SCID-X1 gene therapy trial. *Blood* **115**, 4356-4366 (2010).
53. Boztug, K. *et al.* Stem-Cell Gene Therapy for the Wiskott–Aldrich Syndrome. *The New England Journal of Medicine* **363**, 1918-1927 (2010).
54. Ikeda, K., Mason, P.J. & Bessler, M. 3′UTR-truncated Hmga2 cDNA causes MPN-like hematopoiesis by conferring a clonal growth advantage at the level of HSC in mice. *Blood* **117**, 5860-5869 (2011).
55. Negre, O. *et al.* Preclinical Evaluation of Efficacy and Safety of an Improved Lentiviral Vector for the Treatment of β-Thalassemia and Sickle Cell Disease. *Current Gene Therapy* **15**, 64-81 (2015).
56. Wu, L.-C. *et al.* Correction of sickle cell disease by homologous recombination in embryonic stem cells. *Blood* **108**, 1183-1188 (2006).
57. Chang, J.C., Ye, L. & Kan, Y.W. Correction of the sickle cell mutation in embryonic stem cells. *Proceedings From the National Academy of Sciences USA* **103**, 1036-1040 (2006).

58. de Rham, C. & Villard, J. How to cross immunogenetic hurdles to human embryonic stem cell transplantation. *Seminars in Immunopathology* **33**, 525-534 (2011).
59. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-872 (2007).
60. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).
61. Yu, J. *et al.* Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science* **318**, 1917-1920 (2007).
62. Hanna, J. *et al.* Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* **318**, 1920-1923 (2007).
63. Ye, L. *et al.* Induced pluripotent stem cells offer new approach to therapy in thalassemia and sickle cell anemia and option in prenatal diagnosis in genetic diseases. *Proceedings of the National Academy of Sciences* **106**, 9826-9830 (2009).
64. Wang, Y. *et al.* Genetic correction of β-thalassemia patient-specific iPS cells and its use in improving hemoglobin production in irradiated SCID mice. *Cell Research* **22**, 637-648 (2012).
65. Puri, M.C. & Nagy, A. Concise review: Embryonic stem cells verse induced pluripotent stem cells: The game is on. *Stem Cells* **30**, 10-14 (2012).
66. Bollag, R.J., Waldman, A.S. & Liskay, R.M. Homologous recombination in mammalian cells. *Annual Reviews of Genetics* **23**, 199-225 (1989).
67. Shrivastav, M., De Haro, L.P. & Nickoloff, J.A. Regulation of DNA double-strand break repair pathway choice. *Cell Research* **18**, 134-147 (2008).
68. Colleaux, L., D'Auriol, L., Galibert, F. & Dujon, B. Recognition and cleavage site of the intron-encoded omega transposase. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6022-6026 (1988).
69. Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6064-6068 (1994).
70. Smih, F., Rouet, P., Romanienko, P.J. & Jasin, M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Research* **23**, 5012-5019 (1995).
71. Jasin, M. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics* **12**, 224-228 (1996).
72. Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773-782 (2011).
73. Choo, Y. & Klug, A. Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proceedings of the National Academy of Sciences* **91**, 11163-11167 (1994).
74. Kim, Y.G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences* **93**, 1156-1160 (1996).
75. Pavletich, N. & Pabo, C. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809-817 (1991).
76. Smith, J. *et al.* Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Research* **28**, 3361-3369 (2000).

77. Miller, J.C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature biotechnology* **25**, 778-785 (2007).
78. Joung, J.K. & Jeffry, S.D. TALENS: a widely applicable technology for targeted genome editing. *Nature Reviews Molecular Cell Biology* **14**, 49-55 (2013).
79. Bibikova, M., Beumer, K., Trautman, J.K. & Carroll, D. Enhancing gene targeting with designed zinc finger nucleases. *Science* **300**, 764 (2003).
80. Morton, J., Davis, M.W., Jorgensen, E.M. & Carroll, D. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 16370-5 (2006).
81. Meyer, M., de Angelis, M.H., Wurst, W. & Kühn, R. Gene targeting by homologous recombination in mouse zygote mediated by zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 15022-15026 (2010).
82. Meng, X., Noyes, M., Zhu, L.J., Lawson, N.D. & Wolfe, S.A. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nature Biotechnology* **26**, 695-701 (2008).
83. Porteus, M.H. & Baltimore, D. Chimeric nucleases stimulate gene targeting in human cells. *Science* **300**, 763 (2003).
84. Urnov, F.D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646-651 (2005).
85. Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nature Biotechnology* **28**, 839-847 (2010).
86. Tebas, P. *et al.* Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *The New England journal of medicine* **370**, 901-910 (2014).
87. Sebastiano, V. *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* **29**, 1717-1726 (2011).
88. Mussolino, C. & Cathomen, T. TALE nucleases: tailored genome engineering made easy. *Current Opinion in Biotechnology* **23**, 1-7 (2012).
89. Cermak, T. *et al.* Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* (2011).
90. Moscou, M.J. & Bogdanove, A.J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **326**, 1501 (2009).
91. Boch, J. *et al.* Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**, 1509-1512 (2009).
92. Huang, P. *et al.* Heritable gene targeting in zebrafish using customized TALENs. *Nature Biotechnology* **29**, 699-700 (2011).
93. Wood, A.J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307 (2011).
94. Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nature Biotechnology* **29**, 731-734 (2011).
95. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
96. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709-1712 (2007).
97. Sander, J.D. & Joung, J.K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nature Biotechnology* **32**, 347-355 (2014).
98. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819-823 (2013).

99. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research*, 9584-9592 (2013).
100. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotech* **31**, 822-826 (2013).
101. Hsu, P.D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotech* **31**, 827-832 (2013).
102. Lin, Y. *et al.* CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Research* **42**, 7473-7485 (2014).
103. Li, D. *et al.* Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nature Biotechnology* **31**, 681-683 (2013).
104. Kimura, Y., Hisano, Y., Kawahara, A. & Higashijima, S.-i. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Scientific Reports* **4**(2014).
105. Cho, S.W., Lee, J., Carroll, D., Kim, J.-S. & Lee, J. Heritable Gene Knockout in Caenorhabditis elegans by Direct Injection of Cas9–sgRNA Ribonucleoproteins. *Genetics* **195**, 1177-1180 (2013).
106. 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* **31**, 684-686 (2013).
107. Wang, H. *et al.* One-Step Generation of Mice Carrying Mutations in Multiple Genes by CRISPR/Cas-Mediated Genome Engineering. *Cell* **153**, 910-918 (2013).
108. Shan, Q. *et al.* Targeted genome modification of crop plants using a CRISPR-Cas system. *Nature Biotechnology* **31**, 686-688 (2013).
109. Li, J.-F. *et al.* Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. *Nature Biotechnology* **31**, 688-691 (2013).
110. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D.G. & Kamoun, S. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. *Nature Biotechnology* **31**, 691-693 (2013).
111. Choulika, A., Perrin, A., Dujon, B. & Nicolas, J.-F. Induction of homologous recombination in mammalian chromosomes using the I-SceI system of saccharomyces cerevisiae. *Molecular and Cellular Biology* **15**, 1968-1973 (1995).
112. Elliot, B., Richardson, C., Winderbaum, J., Nickoloff, J.A. & Jasin, M. Gene conversion tracts from double strand break repair in mammalian cells. *Molecular and Cellular Biology* **18**, 93-101 (1998).
113. Lombardo, A. *et al.* Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nature Biotechnology* **25**, 1298-1306 (2007).
114. Voit, R.A., Hendel, A., Pruett-Miller, S.M. & Porteus, M.H. Nuclease-mediated gene editing by homologous recombination of the human globin locus. *Nucleic Acids Research*, 1-14 (2013).
115. Sun, N., Liang, J., Abil, Z. & Zhao, H. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Molecular Biosystems* **8**, 1255-1263 (2012).
116. Hendel, A. *et al.* Quantifying Genome-Editing Outcomes at Endogenous Loci with SMRT Sequencing. *Cell Reports* **7**, 293-305 (2014).

117. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nature Biotechnology* **27**, 851-857 (2009).
118. Yusa, K. *et al.* Targeted gene correction of [agr]1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* **478**, 391-394 (2011).
119. Zou, J., Mali, X., Dowey, S.N. & Cheng, L. Site specific gene corection of the a point mutation in human IPS cells derived from an adult patient with sickle disease. *Gene Therapy* **118**, 4599-4608 (2011).
120. Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235-240 (2014).
121. Hoban, M.D. *et al.* Correction of the sickle-cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* (2015).
122. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics* **11**, 636-646 (2010).
123. Szczepek, M. *et al.* Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotech* **25**, 786-793 (2007).
124. Neff, K. *et al.* Mojo Hand, a TALEN design tool for genome editing applications. *BMC Bioinformatics* **14**, 1-7 (2013).
125. Lin, Y. *et al.* SAPTA: a new design tool for improving TALE nuclease activity. *Nucleic Acids Research* **42**, e47 (2014).
126. Sander, J.D. *et al.* ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Research* (2010).
127. Sander, J.D., Zaback, P., Joung, J.K., Voytas, D.F. & Dobbs, D. Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. *Nucleic Acids Research* **35**, W599-W605 (2007).
128. Hodgkins, A. *et al.* WGE: a CRISPR database for genome engineering. *Bioinformatics* (2015).
129. Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Research* **42**, W401-W407 (2014).
130. Xie, S., Shen, B., Zhang, C., Huang, X. & Zhang, Y. sgRNAcas9: A Software Package for Designing CRISPR sgRNA and Evaluating Potential Off-Target Cleavage Sites. *PLoS ONE* **9**, e100448 (2014).
131. Fine, E.J., Cradick, T.J., Zhao, C.L., Lin, Y. & Bao, G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Research* **42**, e42 (2014).
132. Cradick, T.J., Qiu, P., Lee, C.M., Fine, E.J. & Bao, G. COSMID: A Web-based Tool for Identifying and Validating CRISPR/Cas Off-target Sites. *Molecular Therapy. Nucleic Acids* **3**, e214 (2014).
133. Ran, F.A. *et al.* Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**, 1380-1389 (2013).
134. Mali, P. *et al.* RNA-Guided Human Genome Engineering via Cas9. *Science* **339**, 823-826 (2013).
135. Certo, M.T. *et al.* Tracking genome engineering outcome at individual DNA breakpoints. *Nature Methods* **8**, 671-676 (2011).
136. Brinkman, E.K., Chen, T., Amendola, M. & van Steensel, B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research* **42**, e168 (2014).
137. Nightingale, S.J. *et al.* Transient Gene Expression by Nonintegrating Lentiviral Vectors. *Molecular Therapy* **13**, 1121-1132 (2006).

138. Zuris, J.A. *et al.* Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature Biotechnology* **33**, 73-80 (2015).
139. Herman, E.C. & Conley, C.L. Hereditary persistence of fetal hemoglobin: A family study. *The American Journal of Medicine* **29**, 9-17 (1960).
140. Uda, M. *et al.* Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of β-thalassemia. *Proceedings of the National Academy of Sciences* **105**, 1620-1625 (2008).
141. Lettre, G. *et al.* DNA polymorphisms at the BCL11A, HBS1L-MYB, and β-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. *Proceedings of the National Academy of Sciences* **105**, 11869-11874 (2008).
142. Sankaran, V.G. *et al.* Human Fetal Hemoglobin Expression Is Regulated by the Developmental Stage-Specific Repressor BCL11A. *Science* **322**, 1839-1842 (2008).
143. Bauer, D.E. *et al.* An Erythroid Enhancer of BCL11A Subject to Genetic Variation Determines Fetal Hemoglobin Level. *Science* **342**, 253-257 (2013).
144. Ando, D. *et al.* Preclinical Studies for the First Hematopoietic Stem Cell (HSC) Gene Editing Trial: Phase 1 Study of Beta-Thalassemia With Autologous Transplantation of Zinc Finger Nuclease-Treated HSC To Upregulate Fetal Hemoglobin. in *American Society of Gene and Cell Therapy* (New Orleans, 2015).
145. Silva, G. *et al.* Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Current Gene Therapy* **11**, 11-27 (2011).
146. Zhang, Y. & Yu, L.-C. Microinjection as a tool of mechanical delivery. *Current Opinion in Biotechnology* **19**, 506-510 (2008).
147. Gaj, T., Guo, J., Kato, Y., Sirk, S.J. & Barbas III, C.F. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nature Methods* **9**, 805-807 (2012).
148. Tesson, L. *et al.* Knockout rats generated by embryo microinjection of TALENs. *Nature Biotechnology* **29**, 695-696 (2011).
149. Wu, Y. *et al.* Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell* **13**, 659-662 (2013).
150. Zu, Y. *et al.* TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nature Methods* **10**, 329-331 (2013).
151. Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 *in vivo*. *Nature Biotechnology* **28**, 839-847 (2010).
152. Sebastiano, V. *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* **29**, 1717-1726 (2011).
153. Gaj, T., Gersbach, C.A. & Barbas III, C.F. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31**, 397-405 (2013).
154. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting B-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Research* **41**, 9584-9592 (2013).
155. Holkers, M. *et al.* Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Research* **41**, e63 (2013).
156. Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235-240 (2014).

157. Thomas, C.E., Ehrhardt, A. & Kay, M.A. Progress and problems with the use of viral vectors for gene therapy. *Nature Reviews* **4**, 346-358 (2003).
158. Roberg, K., Kagedal, K. & Ollinger, K. Microinjection of cathepsin D induces caspase-dependent apoptosis in fibroblasts. *The American Journal of Pathology* **161**, 89-96 (2002).
159. Davis, B.R. *et al.* Glass needle-mediated microinjection of macromolecules and transgenes into primary human blood stem/progenitor cells. *Blood* **95**, 437-444 (2000).
160. Meyers, M., Ortiz, O., Hrabe de Angelis, M., Wurst, W. & Kuhn, R. Modeling disease mutations by gene targeting in one-cell mouse embryos. *Proceedings in the National Academy of Sciences USA* **109**, 9354-9359 (2012).
161. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).
162. Wefers, B. *et al.* Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proceedings in the National Academy of Sciences USA* **110**, 3782-3787 (2013).
163. Tanaka, R. *et al.* VLA-5 mediated adhesion to fibronectin accelerates hemin-stimulated erythroid differentiation of K562 cells through induction of VLA-4 induction. *Journal of Biological Chemistry* **284**, 19817-19825 (2009).
164. Sive, H.L., Grainger, R.M. & Harland, R.M. Calibration of the injection volume for microinjection of Xenopus oocytes and embryos. *Cold Spring Harbor Laboratory Press*, 1382-3 (2010).
165. Zhang, Y. Microinjection technique and protocol to single cells. *Protocol Exchange* **Published online 2 November 2007**(2007).
166. Lechardeur, D. *et al.* Metabolic instability of plasmid DNA in the cytosol: a potential barrier to gene transfer. *Gene Therapy* **6**, 482-497 (1999).
167. Chenuet, S., Derouazi, M., Hacker, D. & Wurm, F. DNA delivery by microinjection for the generation of recombinant mammalian cell lines. in *Microinjection Methods and Applications* (ed. Carroll, D.J.) 99-112 (Humana Press, Melbourne, FL, 2009).
168. Usuludin, S.B.M., Cao, X. & Lim, M. Co-culture of stromal and erythroleukemia cells in a perfused hollow fiber bioreactor system as an in vitro bone marrow model for myeloid leukemia. *Biotechnology and Bioengineering* **109**, 1248-1258 (2012).
169. Cioe, L. *et al.* Differential expression of the globin genes in human leukemia K562(S) cells induced to differentiate by hemin or butyric acid. *Cancer Research* **41**, 237-243 (1981).
170. Chen, Z. *et al.* Receptor-mediated delivery of engineered nucleases for genome modification. *Nucleic Acids Research*, e182 (2013).
171. Celis, J.E. Microinjection of somatic cells with micropipettes: comparison with other transfer techniques. *The Biochemical Journal* **223**, 281-291 (1984).
172. Schnorf, M., Potrykus, I. & Neuhaus, G. Microinjection technique: routine system for characterization of microcapillaries by bubble pressure measurement. *Experimental Cell Research* **210**, 260-267 (1994).
173. Guschin, D.Y. *et al.* A rapid and general assay for monitoring endogenous gene modification. in *Engineered Zinc Finger Proteins* (eds. Mackay, J.P. & Segal, D.J.) 247-256 (Humana Press, 2010).
174. Kim, H. & Kim, J.-S. A guide to genome engineering with programmable nucleases. *Nat Rev Genet* **15**, 321-334 (2014).
175. Jinek, M. *et al. RNA-programmed genome editing in human cells*, (2013).

176. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.-S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotech* **31**, 230-232 (2013).
177. Li, W., Teng, F., Li, T. & Zhou, Q. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotech* **31**, 684-686 (2013).
178. Van Tendeloo, V.F.I. *et al. Highly efficient gene delivery by mRNA electroporation in human hematopoietic cells: superiority to lipofection and passive pulsing of mRNA and to electroporation of plasmid cDNA for tumor antigen loading of dendritic cells*, 49-56 (2001).
179. Kim, S., Kim, D., Cho, S.W., Kim, J. & Kim, J.-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research* **24**, 1012-1019 (2014).
180. Lin, S., Staahl, B.T., Alla, R.K. & Doudna, J.A. *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery*, (2015).
181. Liu, J., Gaj, T., Patterson, J.T., Sirk, S.J. & Barbas Iii, C.F. Cell-Penetrating Peptide-Mediated Delivery of TALEN Proteins via Bioconjugation for Genome Engineering. *PLoS ONE* **9**, e85755 (2014).
182. Mandal, P.K., Ferreira, L.M.R., Rossi, D.J. & Cowan, C. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. *Cell Stem Cell* **15**, 643-652 (2014).
183. Hamm, A., Krott, N., Breibach, I., Blindt, R. & Bosserhoff, A.K. Efficient transfection method for primary cells. *Tissue Engineering* **8**, 235-245 (2002).
184. von Levetzow, G. *et al.* Nucleofection, an efficient nonviral method to transfer genes into human hematopoietic stem and progenitor cells. *Stem Cells and Development* **15**, 278-284 (2006).
185. Jiang, X.-S. *et al.* Surface-immobilization of adhesion peptides on substrates for ex vivo expansion of cryopreserved umbilical cord blood CD34+ cells. *Biomaterials* **27**, 2723-2732 (2006).
186. Sato, T., Maekawa, T., Watanabe, S., Tsuji, K. & Nakahata, T. Erythroid progenitors differentiate and mature in response to endogenous erythropoietin. *The Journal of Clinical Investigation* **106**, 263-270 (2000).
187. Liu, L. & Gerson, S.L. Targeted modulation of MGMT: Clinical implications. *Molecular Pathways* **12**, 328-331 (2006).
188. Zielske, S.P. & Gerson, S.L. Lentiviral transduction of P140K MGMT into human CD34+ hematopoietic progenitors at low multiplicity of infection confers significant resistance to BG/BCNU and allows selection in vitro. *Molecular Therapy* **5**, 381-387 (2002).
189. Persons, D.A. *et al.* Successful treatment of murine β-thalassemia using in vivo selection of genetically modified, drug-resistant hematopoietic stem cells. *Gene Therapy* **102**, 506-512 (2003).
190. Gohari, N.S. & Helleday, T. Conserative homologous recombination preferentially repairs DNA double stranded breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Research* **32**, 3683-36888 (2004).
191. Liu, P.-Q. *et al.* Generation of Triple-gene knockout mammalian cell line using engineering zinc-finger nucleases. *Biotechnology and Bioengineering* **106**, 97-105 (2010).
192. Fung, H. *et al. Viral Insertion Safety in Patients with Glioblastoma Who Received a Novel Lentiviral MGMT-P140K Gene Therapy to Protect Bone Marrow from*

Chemotherapy: No Dominant Clonal Evolution Observed with Chemo-Selection, 4801-4801 (2014).