CRISPR DIRECTED TARGETED GENE ALTERATION:

MECHANISM TO APPLICATION

by

Natalia Rivera-Torres

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Medical Sciences

Spring 2019

CRISPR DIRECTED TARGETED GENE ALTERATION:

MECHANISM TO APPLICATION

by

Natalia Rivera-Torres

Approved: _____

Esther Biswas-Fiss, Ph.D.
Chair of the Department of Medical and Molecular Sciences

Approved: _____

Kathleen S. Matt, Ph.D.
Dean of the College of Heath Sciences

Approved: _____

Douglas J. Doren, Ph.D.
Interim Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Esther Biswas-Fiss, Ph.D.
Chair of the Department of Medical and Molecular Sciences

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Eric B. Kmiec, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Nicholas Petrelli, M.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Edward Anders Kolb, M.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Kathleen S. Matt, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

Immeasurable appreciation and deepest gratitude for the help and support are extended to the following persons who in one way or another have contributed in making this study possible.

First of all, I would like to wholeheartedly thank my advisor, Dr. Kmiec, for his mentoring, patience, encouragement, and valuable insight on my research over the last six years. His dynamic approach has generated inspiration and confidence in me to peruse my Ph.D. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for research was contagious and motivational for me, even during tough times in the Ph.D. pursuit. I am also thankful for the excellent example he has provided as a successful molecular biologist and professor. The members of the Kmiec Lab have contributed immensely to my personal and professional time in the lab. The group has been a source of friendships as well as good advice and collaboration. I am especially grateful for the special company of the original Kmiec Lab members who stuck it out in grad school with me: Pawel Bialk. Your expertise, support, assistance and guidance has made my time here great and of delightful company. To my committee, Dr. Esther Biswas-Fiss, Dr. Kathy Matt and Dr. Nicholas Petrelli and Dr. Andy Kolb, I am extremely grateful for your assistance and suggestions throughout my project.

I would like to thank the Andrew McDonough B+ Foundation in specific its president Joe McDonough for selecting me as their first Graduate Student Fellow

which provided financial support throughout my Ph.D. studies. I hope I have been able to give back through my research in pediatric Leukemia and help expand the knowledge on how to better treat patients with this devastating disease.

I thank Dr. Sylvain Le Marchand in the Bio-Imaging Center for assistance with all the confocal microscopy imaging. LSM880 confocal microscope**:** Microscopy equipment was acquired with a shared instrumentation grant (S10 OD016361) and access was supported by the NIH-NIGMS (P20 GM103446), the NSF (IIA-1301765) and the State of Delaware. Microscopy access was supported by grants from the NIH-NIGMS (P20 GM103446), the NSF (IIA-1301765) and the State of Delaware.

I thank Dr. Lynn Opdenaker at the CTCR Flow Cytometry Core Facility for running, analyzing, and sorting cells. Along with Lynn am also very appreciative to Dr. Shirin Mourdarai for all the help of both in teaching me and explaining how to do the cell staining.

I thank my friend and program peer Meera Patel for helping do all the protein computational analysis as well as being great support during the tough times. I hope to see you defend soon.

I dedicate my thesis to my family and friends. A special gratitude to my loving parents. Their words of encouragement and push for tenacity will always be present. My sister Solymar, we have chosen different career paths but you have taught me to keep going and never give up until you finish the race for your set goals. It doesn’t

matter if it rains or shines it's that drive for the finish line that should always stay alive. In the loving memory of my aunt Titi. She was not able see me finish my degree but throughout my life she always did what she could to see us reach our goals. Thank you for your unconditional love and support. I also dedicate this thesis to a few of my friends who have supported me throughout the process. I will always appreciate all they have done. Especially Lizette and Luis thank you for all the advice and being my strong rock through the good and hard days. My home away from home.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFF1	AF4/FMR2 Family Member 1
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
bp	Base Pairs
Cas9	CRISPR associated protein 9
Cas12a	CRISPR associated protein 12a
Con A	Concanavilin A
CPT	Camptothecin
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
crRNA	CRISPR RNA
DNA	Deoxyribonucleic Acids
Dsb	Double Stranded Break
eGFP	Enhanced Green Fluorescent Protein
FACS	Fluorescent Activated Cell Sorting
FAH	Fumarylacetoacetate hydrolase
FDA	U.S. Food and Drug Administration
FLT3	Fms-related tyrosine kinase 3
HDR	Homology Directed Repair
HR	Homologous Recombination
ITD	Internal Tandem Duplication

KMT2A	Lysine Methyl Transferase 2A
MLL	Mixed Lineage Leukemia
NHEJ	Non-Homologous End Joining
NT	Non-Transcribed Or Non-Template
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PI	Propidium Iodide
PM	Point Mutation
RGEN	RNA Guided Endonucleases
RPP	Reduced Proliferation Phenotype
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
RT-PCR	Reverse Transcriptase PCR
RTK	Receptor Tyrosine Kinase
RVD	Repeat variable di-residues
SCD	Sickle Cell Disease
sgRNA	Singel Guide RNA
SSA	Single Strand Annealing
ssODN	Single stranded DNA oligonucleotide
T	Transcribed or template
TALEN	Transcription Activator-Like Effector Nuclease
TALE	Transcription Activator-Like Effector
TKI	Tyrosine Kinase Inhibitor

tracrRNA	Trans-activating crRNA
WGA	Wheat Germ Agglutinin
ZFN	Zinc Finger Nuclease

ABSTRACT

Single base mutations can be repaired by introducing single stranded DNA oligonucleotides (ssODN) into a target cell. The frequency at which this occurs is dependent on several of factors: the length of ssODN, the position of the cell in its proliferative cycle, and the presence of double-stranded DNA breaks in the host genome. Genome editing offers a promising strategy for gene repair and correction by overcoming difficulties associated with lack of precision. CRISPR/Cas has increased the pace and lowered the cost of research, allowing the genetic manipulation even in organisms that have historically been difficult to modify. Furthermore, the combinatorial approach uniting ssODNs and CRISPR/Cas9 has emerged as a feasible therapeutic approach. In the work presented in this dissertation I focused on the mechanism and application of gene editing utilizing CRISPR systems. I tested combinatorial approach of utilizing CRISPR/Cas9 system along with ssODN to promote single base pair correction and demonstrate it is now possible to direct single nucleotide exchange in efficient manner. We find that both insertions and deletions accompany single base repair as result from allelic analysis of clonally expanded cell populations. CRISPR/Cas9 and single-stranded oligonucleotide donor DNA molecules working in tandem can lead to the precise repair of the point mutation in the eGFP gene, and led to propose a new model for the repair of point mutations, a process we have termed ExACT. The relationship between transfection efficiency and gene editing activity was tested and analyzed based on experimental and visual data and

found that there is no direct correlation between efficient cellular uptake and genome modification directed by an RNP. By understanding the mechanisms by which CRISPR/Cas executes gene editing in human cells, a more efficacious and potential approach to drug development could be undertaken. The application of the CRISPR gene editing system in two different approaches to study pediatric Leukemia was explored. (1) pediatric patient specific ALL chromosomal translocation (4:11)(q21:q23) was re-created by utilizing the CRISPR/Cas9 system in HEK293 cells. This led to the development of a convenient platform for rapid modeling of cancer-related genetic mutations in vitro. (2) Implemented the use of a novel gene editing approach to create expression vectors that harbor patient specific mutations that were tested against TKI. We have developed a diagnostic system to monitor the impact of mutant FLT3 ITDs on the progression of oncogenesis and to evaluate the efficacy of novel AML drugs.

Chapter 1

INTRODUCTION

Gene Editing

Targeted gene alteration is a molecular strategy that aims to modify the genome at a precise predetermined site. This genome engineering approach could correct single base mutations responsible for genetic diseases. The traditional approach to gene therapy requires a virus to deliver a functional copy of a gene, which would compensate for the lack of activity from the dysfunctional gene. This approach has not been as successful in its early attempts since it was not regulated properly, often causing random insertions generally disrupting other genes and often driving the cells toward oncogenesis [1,2]. The current gene-editing concept involves using single stranded DNA oligonucleotides to direct a nucleotide exchange at a precise genomic site [3]. Therefore, single base mutations have the potential to be repaired by introducing single stranded DNA oligonucleotides (ssODN) into a target cell [3–5]. The frequency of this corrective activity is dependent on a set of factors: the length of ssODN, the position of the cell in its proliferative cycle [6,7], and double-stranded DNA breaks in the host genome [8,9].

Three major agents are used to catalyze specific double stranded DNA breaks: Zinc-Finger Nucleases (ZFNs), Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 and Transcription Activator-Like Effector Nucleases (TALENs),(see Figure 1) [10–15]. All three technologies catalyze double strand

cleavage of chromosomal DNA in a site-specific manner, triggering endogenous DNA repair systems to engage in targeted genome modification. ZFNs are programmable nucleases composed of a FokI non-specific DNA cleavage catalytic domain and a zinc-finger protein DNA-binding domain. TALENs are proteins with a TALE binding domain and the same FokI nuclease domain [16–18]. CRISPR/Cas relies on a guide RNA and the Cas nuclease that will cleave DNA upon target recognition and activation[19]. These so-called "programmable nucleases" [20] could enable more efficient use of the double stranded break as a stimulatory factor in reactions designed to correct single base mutations. Each of these nucleases has it unique advantages and limitations summarized in Table 1.

Two distinct pathways repair double stranded breaks: error prone non-homologous end joining (NHEJ) or Homology Directed Repair (HDR) (see Figure 2). Shortly after the cleavage the cell attempts to repair the damage through NHEJ unless a suitable donor is present. If no donor is present, the gene may be rendered inactive through deletions and/or frame-shift mutations (others called indels) at the site of the break; events that can lead to knockout of the gene. If a suitable donor is present (ssODN for example), the cell may repair the damage through the HDR repair pathway. The donor serves as a template to guide the repair and generate specific nucleotide exchange.

Gene Editing and Genetic Disorders

Gene editing can generate inheritable nucleotide changes and as such could be used in the treatment of genetic disorders caused by point mutations. Sickle cell anemia occurs in approximately 1 out of every 500 African American births and 1 out of every 36,000 Hispanic American births; over 2 million Americans have sickle cell

trait. At the genetic level, a mutation in the second position of the sixth codon in the β-globin gene(s) represents the genotype of SCD. The mutant codon encodes Glutamic acid (GAG) instead of the normal Valine (GTG) and results in the production of hemoglobin Hb^S which differs from the normal Hb^A in that it tends to polymerize into long strands that deform the erythrocyte [21]. Hydroxyurea and chronic transfusions are treatments; these therapies are wrought with short and long term side effects that limit efficacy[22,23]. Gene editing aims to correct the single base mutation, changing the sixth codon to Glutamic Acid and permitting the cells to produce Hb^A [21,24].

Promising results in the study of other genetic disorders caused by point mutations such as Tyrosinemia with gene editing approaches demonstrate the possibility of gene editing as a clinical therapy for this genetic disorders and others in a near future. Hao et al.[25] corrected the point mutation that causes a Tyrosinemia in an essential gene involved in liver metabolism using a mouse model system [25]. Adult mice carrying the mutated form of the FAH enzyme, were treated with delivered RNA guide strands, the Cas9 gene and a 199-nucleotide DNA repair template that includes the correct sequence of the mutated FAH gene [25]. Other diseases that could be treated using a gene editing approach include: AIDs, Hemophilia, Huntington's Diseases and Combined Immunodeficiency Syndrome [26,27]. Genome editing offers a promising strategy to overcome difficulties associated with lack of precision when inserting new genetic material and the potential effects of viral vectors that have limited the success of *in vivo* gene therapy. However, in the recent years gene therapy there has been successful in correcting an inherited eye genetic mutation that leads to blindness [28]. By 2017 it became the third gene therapy approved by the FDA for use

in the U.S. for children and adults with retinal dystrophy due to a mutation of the RPE65 gene [29].

The CRISPR-Cas9 system is being employed to edit the *CFTR* gene to repair mutations that lead to cystic fibrosis and in the dystrophin gene, where mutations lead to Duchenne and Becker muscular dystrophy[30]. Epigenomic editing is a potential strategy for diseases of epigenetic dysregulation, such as cancers. This could be achieved using a Cas9 protein modified to deliver an epigenetic modification to a target site rather than to cut the genome[31]. Cas9 might also be altered, or related enzymes may be employed, to cleave different forms of RNA, with potential application to the removal of infectious RNA viruses (e.g. rotavirus, Ebola and Zika) or in the recognition of eukaryotic RNA carrying modifications such as methylation[32,33].

Single stranded oligonucleotide (ssODN) Directed Gene Editing

Single-stranded oligonucleotides (ssODNs) are short, single stranded DNA synthetic molecules, produced by solid-phase chemical synthesis of any specified sequence and length. It was one of the first tools used for gene editing. The ssODN is designed in a way that it is complementary to the target gene at all bases except one, which creates a mismatch at the base intended to be changed [34]. Single stranded DNA is highly susceptible to intracellular degradation; shorter ssODNs are highly susceptible to nulcleolytic degradation. Phosphorothioate linkages are often used between the terminal 5’ and 3’ to protect the ssODNs from nucleolytic degradation. common for bases ends to have 3 phosphorothioate bonds [34]. Different lengths of ssODNs (35 to 150 bases) have been used for gene editing experiments.

Another parameter that may vary in the design of the ssODN is the strand to which it will be complementary too, transcribed strand (T) or non-transcribed (NT). A strand bias has been attributed to the increased accessibility of the non-transcribed (NT) strand for ssODN annealing during transcription having a beneficial role on ssODN-mediated gene targeting [34]. Most successful studies use NT ssODNs [35–37]. These reaction parameters must be considered when designing the best conditions for obtaining high correction efficiencies. Single base changes directed by ssODNs have been reported at frequencies ranging from 0.3% and 1% routinely [3–5].

Previous data reveals that double-stranded breaks introduced by including anti-cancer drugs in the reactions, enhances the frequency of this repair. The anti-cancer drugs are typically DNA breakage more often at the replication forks, which can activate the DNA repair mechanisms [8,9,38]. Pre-treatment of cells, targeted for gene editing by ssODNs, with camptothecin (CPT) enhanced gene editing activity 5-10 fold [39]. These double stranded breaks were not site specific, so they could disrupt normal cell function by random cleavage (thus limiting their potential for clinical application). Although the double stranded breaks are not specific, they enhance gene editing activity by providing an entry point for the ssODN to align in homologous register with the target region (see Figure 3). Some have suggested that once aligned the ssODN could provide a 3'OH for extension and act as a "quasi Okazaki fragment" [3,37]. During elongation phased DNA replication in gene editing directed by ssODNs can be divided into three distinct phases [3,4]. First, initiation, involves the alignment of the ssODN in homologous register with the target site. Next, the step of correction comprises the actual nucleotide exchange at the single mismatch base and last, recovery is the phase in which the cell resumes its normal metabolic activities [37].

Impact of Cell Cycle on Gene Editing

The importance of the cell cycle stage for gene editing has been demonstrated repeatedly [6,37,38,40–42] . Studies have centered on the effect of S phase when the cells are undergoing DNA replication and the chromatin is in a non-condensed state. This chromatin structure may permit unfolding allowing the gene to be more accessible to interact with the ssODN searching for homology. A model has emerged in which reversal of genotype takes place most often through incorporating the ssODN into a newly synthesized DNA strand [3]. This incorporation appears possible when DNA replication takes place during the gene editing reaction [3,6,7,38]. The mechanism of action involves the incorporation of the ssODN into a growing replication fork [4,5,12] (see Figure 3), which probably disrupts the chromatin structure, reduces steric hindrance and permits ssODN access to the target site.

In past studies, the level of ssODN required to activate the reaction is so high that the corrected cells cease to proliferate; a phenomenon termed, Reduced Proliferation Phenotype (RPP) [37]. It has also been reported that gene correction takes place most effectively during mid S phase: synchronized and released cells (~4hrs) are most amenable (Figure 4). These observations establish the importance of DNA replication in the gene editing reaction [3,6,7,38]. The level of gene repair is enhanced dramatically when cells are targeted during S phase and, specifically, when slowed in their progression through S phase [38,40,41]. Gene editing directed by a 72-mer (NT), a 72bp ssODN that hybridizes to the non-transcribed strand, takes place at an approximate level of 0.7% targeting unsynchronized cells [37]. When synchronized and released cells are targeted, the correction frequencies approach 2% [38,43]. Synchronization and release of these cells enables higher levels of targeting since more cells are traversing through S phase, which may shift the balance toward HR or

HDR and away from NHEJ. Double stranded DNA breakage (random) arrests cell cycle progression and elevates gene-editing frequency significantly. The large number of free single stranded ends from ssODNs induces the activation of the DNA damage response pathway and RPP begins.

Gene Editing Model System

HCT 116-19 cells are a well-established gene editing model system used to elucidate the mechanism of action [21,34,40] . This cell line (HCT 116-19) is derived from human colorectal carcinoma with a single copy of the mutant eGFP gene integrated into the genome. This model system enables the correlation between genotypic and phenotypic changes with functional protein activity. The integrated single copy eGFP gene (illustrated in Figure 5) contains a nonsense mutation at position +67, which results in the expression of nonfunctional eGFP. The eGFP gene was mutated near the 5' end of the coding sequence creating a stop codon (TAG) in place of a tyrosine (TAC). The presence of the stop codon (TAG) truncates the production of eGFP resulting in the absence of green fluoresce in the cell. This system is an established model for analyzing the mechanism of gene editing in human cells [21,34,37,40,44,45].

In most applications, 72NT ssODN has been used in optimization studies for delivery and the response of cell and genomic DNA in the gene editing reaction [46]. Once the reaction or the repair of the TA**G**→TA**C** has been facilitated, the population of cells is analyzed by Flow Cytometry or Fluorescence-activated cell sorting (FACS) and the percentage of live green fluorescent cells within that population is presented as a frequency of gene correction. Sorted $eGFP^+$ cells are easily quantified and genotype can be verified by direct DNA sequencing. Thus, genotype and phenotype, expression

of a functional protein, is assessed in a valid, simple way, a critical component of reaction optimization or characterization studies.

Transcription Activator Like-Effector Nuclease (TALEN)

Transcription Activator-Like Effectors (TALEs) are naturally occurring proteins found in pathogenic bacteria *Xanthomanas*. TALEs consist of DNA binding domains composed of 33-35 amino acids each recognizing a single base of DNA. Their specificity is enabled by two hyper-variable amino acids at positions 12 and 13 positions, known as repeat variable di-residues (RVD) and notated as NI, NG, HD, and NN, which specifically recognize Adenine, Thymine, Cytosine and Guanine. Single base recognition by TALEs gives them greater flexibility in design than the Zinc Finger proteins, which bind to DNA in triplets recognizing codons (Figure 6).

TALEs have been fused with the nuclease FokI, a type IIS restriction enzyme structured by a separable DNA-binding domain and a nuclease domain. The nuclease domain from Fok1 is fused to the carboxyl termini of the TALE protein and this fusion protein is known as Transcription Activator-Like Effector Nuclease (TALEN). TALENs will only cut when dimerized, and dimerization will occur only when two TALEN arms or subunits that recognize a specific sequence in the genome associate. Each arm binds to one of the two strands of DNA; the Left arm (L) is the TALEN subunit the will bind 5'-3' strand and the Right arm (R) is the TALEN that will bind to the 3'-5' strand. TALEN arms bound to their respective target sites must have a spacer of 13-30 bases for an effective dimerization. The spacer region is the distance between the site where Fok1 binds and where the last RVD binds to on each TALEN arm. The cut sites are predicted to occur at the center of the spacer region for each TALEN pair. While unnecessary the preferred specific sequence recognized by each

arm is best proceeded by a thymine. The flexibility for the design of custom TALENs allows for the creation of a site-specific endonuclease that can act anywhere the genome.

TALENs were designed as a tool as site-specific endonucleases for selective genome engineering. The site-specific cleavage creates a double strand break in the chromosomal DNA acting together with ssODN to execute gene editing (Figure 7). First, a precise entry point for the ssODN, which can function as primer through the incorporation with the 3' end serving as a "quasi Okazaki fragment" is created by the break. Second, the 5' site undergoes error-prone NHEJ as the preferred pathway to repair double stranded breaks. The site-specific chromosomal double strand break increases the efficiency of genome modification. Current activity involving TALENs center on the generation of gene knockouts in cells or animals [47]. But the combination of TALENs and ssODNs to direct gene editing does show great promise particularly because this combination increases correction efficiencies [20,48–52]. Recently, Yang et al. [20] published an elegant study focusing on cleavage/target site location and several reaction parameters including ssODN length. They found that proximal cleavage within 50-100 bases of the target base produced the highest level of gene editing and there is an optimal length for the ssODN in driving the reaction.

TALENs may help the gene editing reaction take place by reducing steric hindrance of the chromatin structure blocking access of the ssODN to the target sequence and providing an entry point for the ssODN to begin the search for homology. The formation of a plectonemic joint is enabled between one end of the duplex and the ssODN permitting alignment in a homologous register. TALENs also reduce the level of ssODNs needed for nucleotide exchange, eliminating the onset of

the RPP [37]. The site specific double stranded break increases the efficiency of the gene editing reaction by lowering the levels of ssODN required to direct gene editing. As such, these conditions decrease the possibility of activation of the DNA damage response pathway eliminating the factors that induce the cells to go through a RPP. Liu et al. [53] suggest that the ssODNs may reduce the number of NHEJ events tipping the balance away from the potentially mutagenic activity of NHEJ. The efficiency of gene editing has been dependent on including both TALEN arms, the specific ssODN designed to align with the target site and an optimized TALEN:ssODN (μg/μg) ratio in the reaction. TALENs appear to be dose dependent with imbalanced levels of TALEN and ssODN leading to a decrease in the reaction efficiency [37].

Clustered Regularly Interspace Short Palindromic Repeats Associated with Cas9

While TALENs have been effective, the design, construction and effective expression of these molecules in mammalian cells is challenging. This ease-of-use issue has enabled the rapid development and acceptance of a new technology, a new tool for gene editing; Clustered, regularly interspace, short palindromic repeats (CRISPR) molecules associated with Cas9 endonucleases (CRISPR/Cas9). This system functions naturally in the adaptive immunity pathway of bacteria. Through repurposing, CRISPR/Cas9 has emerged as the preferred system to catalyze site-specific DNA cleavage [12,13,54–57]. Cas9 assembles with a designed crRNA to create a complex that can cut at any desired site in a chromosome. A protospacer adjacent motif (PAM) is required for Cas9 binding, the target must be upstream of a 5’-NGG-3’ site (with SpCas9). A major advantage of CRISPR/Cas9 is the simplicity with which the vectors expressing the reaction components can be created and utilized.

While there are some notable sequence restrictions, i.e. the proto-spacer adjacent motif (PAM), these RNA- guided engineered nucleases (RGENs) have become main stream in this field. Once present in mammalian cells, the Cas9/gRNA complex migrates to the nucleus and scans the genome for the PAM sequence. When the Cas9/gRNA binds to a PAM, it then partly unwinds the DNA upstream of the PAM; if the target region of the gRNA can bind with a 20:20 match, the Cas9 nuclease then cuts both strands of DNA 3 bp upstream of the PAM[58]. Once Cas9/gRNA has created a DSB, the same DNA repair pathways exploited by ZFNs and TALENs are activated.

In response cells employ one of two pathways to repair the damage: either through non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Figure 2) [17]. NHEJ can occur through canonical NHEJ (C-NHEJ), which ligates the broken ends back together. There is an alternative end joining pathway (alt-NHEJ), in which one strand of the DNA on either side of the break is resected to repair the lesion [59,60]. Both of these repair methods are error-prone, meaning that the lesion is repaired imperfectly, resulting in insertions or deletions. If there is a nearby DNA molecule with homology to the region around the double-strand break, then the homologous DNA can be a template to repair the break through the homology-directed repair (HDR) pathway.

CRISPR-directed Genome Editing

CRISPR-directed genome editing has far been in scientific research. However, the potential applications of these techniques are much wider than just research. Given that genome editing has the potential to alter any DNA sequence, whether in a bacterium, plant, animal or human, CRISPR/Cas has an almost limitless range. Areas of research and possible applications include:

- **Crops and livestock** (e.g. increasing yield, introducing resistance to disease and pests, tolerance of different environmental conditions) [61,62]
- **Industrial biotechnology** (e.g. developing 'third generation' biofuels and producing chemicals, materials and pharmaceuticals) [63,64]
- **Biomedicine** (e.g. pharmaceutical development, xenotransplantation, gene and cell-based therapies, control of insect-borne diseases) [64,65]
- **Reproduction** (e.g. preventing the inheritance of a disease trait) [66,67]

CRISPR- Cas9 system, have increased the pace and lowered the cost of research, expanding the possibilities and allowing the genetic manipulation of cells and organisms that have historically been difficult to modify[68]. A major challenge is changing a DNA sequence to define gene function, rather than to delete the gene function which results in total dysfunction [69,70]. Global research and development companies have used CRISPR/Cas 9 for novel drug screening and to treat several life-threatening medical conditions, including Sickle-Cell Anemia and cancer. The first trials involve injecting genetically modified cells directly into tissues or taking cells out, engineering them in the lab, and replacing them [71,72]. Cell based therapies involve transfusion or transplantation of cell populations edited expanded and prepared in the laboratory. Genome editing techniques can generate cell lines with specific characteristics to provide disease models and investigate underlying pathology, and to screen potential medicines by evaluating their toxicity before they are considered for humans. Many animal models are highly inbred, offering near defined genetic backgrounds for analysis of the consequences of specific mutation. A longstanding limitation with certain human cells (e.g. induced pluripotent stem cells –

iPS cells) or outbred animals used to model disease is that the healthy controls (to which the disease model is compared) may have multiple genetic differences compared to the disease model[73]. Combined with other technologies (e.g. iPS cell production), genome editing can develop cells whose genetic background is identical (isogenic) to that of the disease model. Editing isogenic genomes introduces a change so the cell line differs only regarding that specific change. This gives greater certainty about the effect of the precise, known difference between the disease variant and the control. CRISPR gene editing will make it more likely that research will diversify into modelling a greater variety of diseases, including individually 'rare' diseases [74,75]· These are a growing focus as more disease-causing mutations are discovered, which are potentially more tractable to the available technology than complex polygenic diseases [76,77]. A prospect is the development of 'personalized' mutant animals that model a disease variant affecting a particular human family or individual [78–80].

Greater use of genome editing in various biological systems can also be expected to lead to greater understanding and refinement of the technique. In genome editing, a new generation of Cas9 protein has been engineered to be so efficient that no off- target cutting is detectable across the whole genome [81,82]. The technique has also been extended for the use to overcome limitations to the visualization of multiple genomic loci by using 'nuclease-dead' Cas9 to bind to cells with up to seven distinct fluorescent markers. This allows researchers to track the location of genes in a chromosome in living cells, which is important in understanding what happens (and what can go wrong) in cellular development [83]

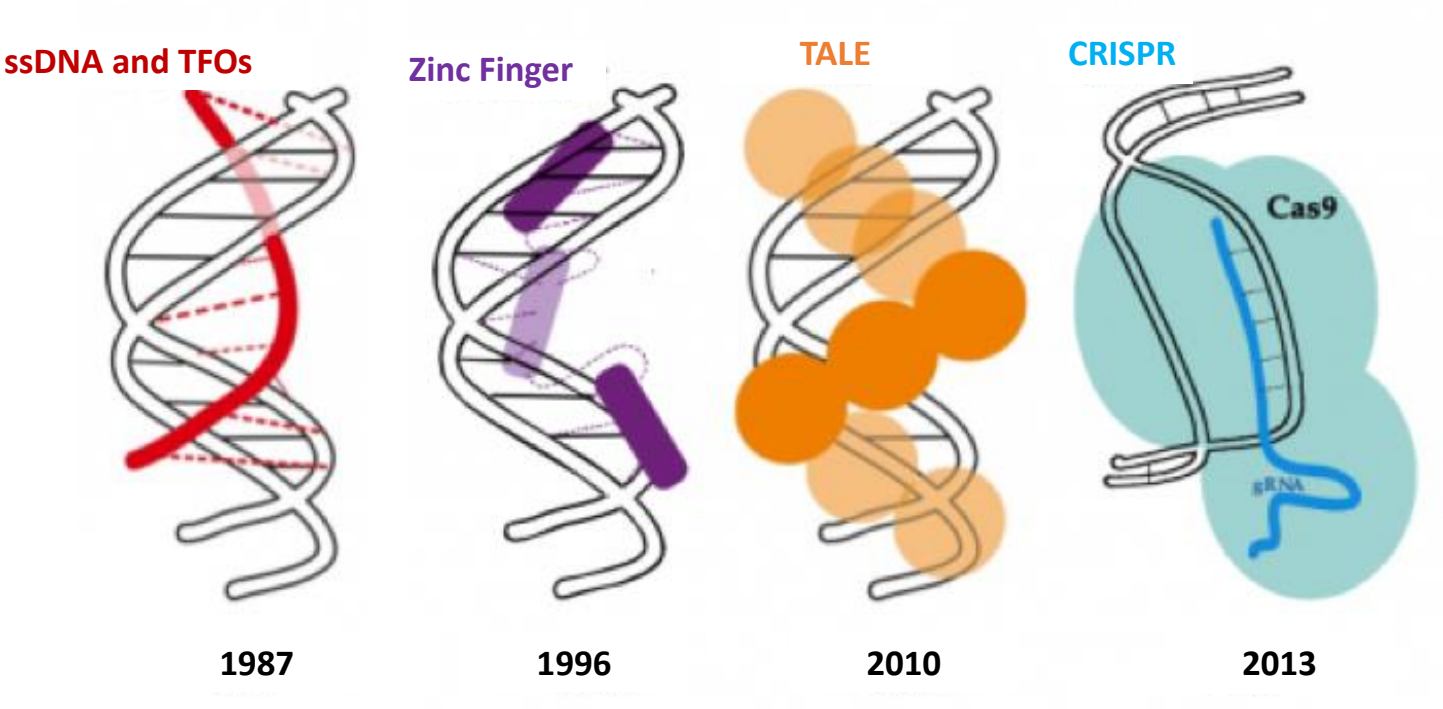


Figure 1. Gene Editing Tools and Technologies.
Schematic representation of the three major agents are currently used to catalyze specific double stranded DNA breaks: Zinc-Finger Nucleases (ZFNs), Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 and Transcription Activator-Like Effector Nucleases (TALENs) as well as the initial tool to carry out gene editing the single-stranded oligodeoxynulceotide (ssODN). Adapted from: [84].

Table 1. Comparison of different programmable nucleases

Variable	ZFN	TALEN	CRISPR
DNA-recognition moiety	Protein	Protein	RNA
Target site size	18-36	30-40	22
Nuclease	FokI	Fok1	Cas
Nuclease recognition site	-	-	PAM
Cytotoxicity	Variable to high	Low	Low
Design complexity	Highly complex	Complex	Low
Construction Time	3-6 months	1 month	2-7 days
Multiplexing capability	Low	Low	High

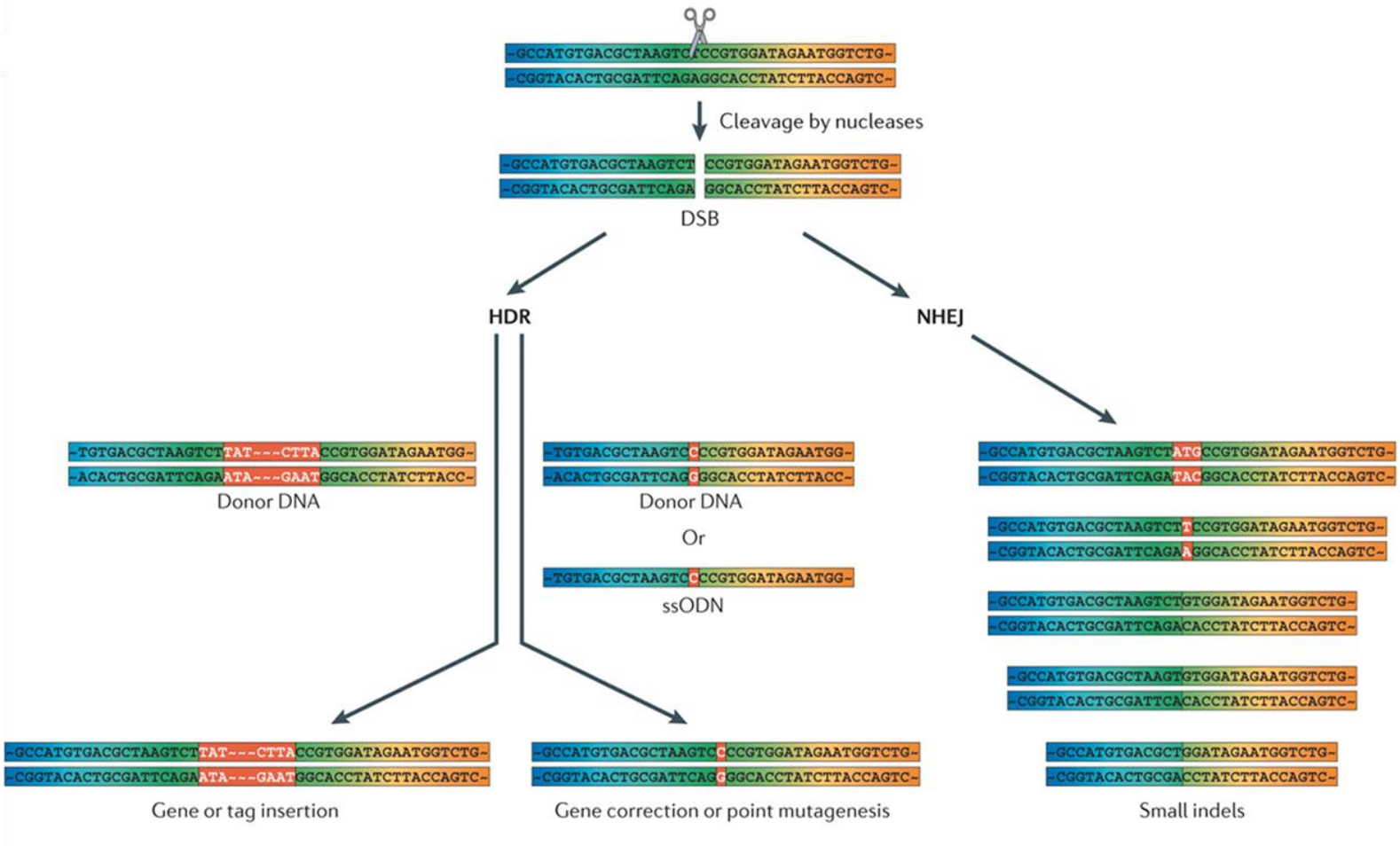


Figure 2. DNA Repair Pathways for Gene Editing with Programmable Nucleases.
Nuclease-induced double strand breaks (DSBs) can lead to sequence insertion, nucleotide correction or change (red box) through homology-directed repair (HDR) in the presence of a donor DNA or a single-stranded oligodeoxynulceotide

(ssODN), both of which contain homology arms. DSBs can also be repaired through error-prone non-homologous end joining (NHEJ), which does not require donor DNA or ssODN and consequently often leads to small insertions and deletions (indels). Source: [17].

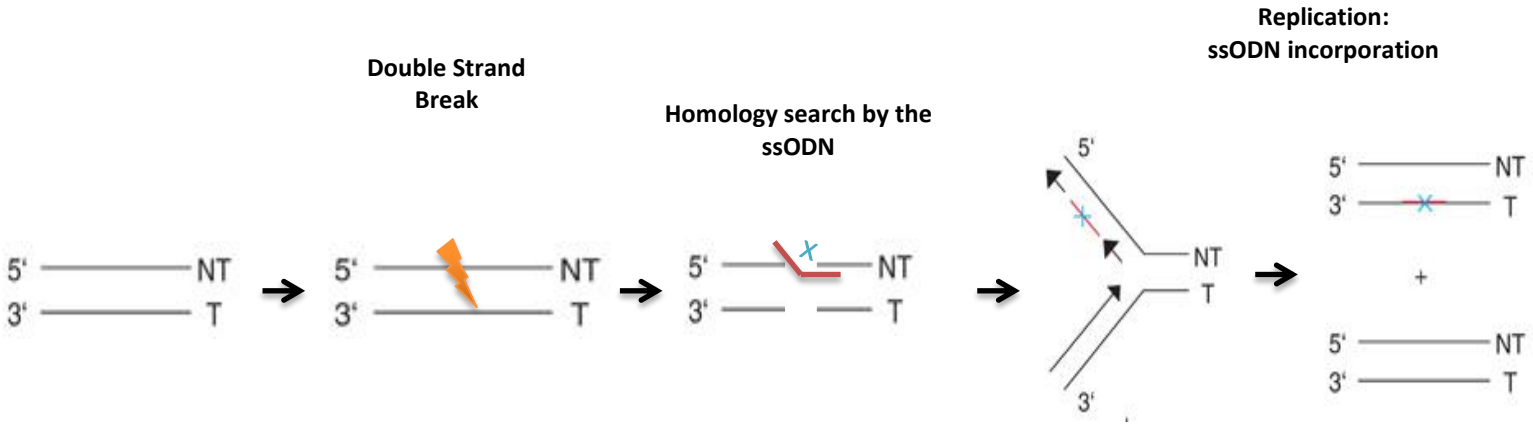


Figure 3. Replication Dependent Model.
SsODN serves as a primer for DNA synthesis during replication and becomes integrated into the newly synthesized DNA strand. Adapted from: [5]

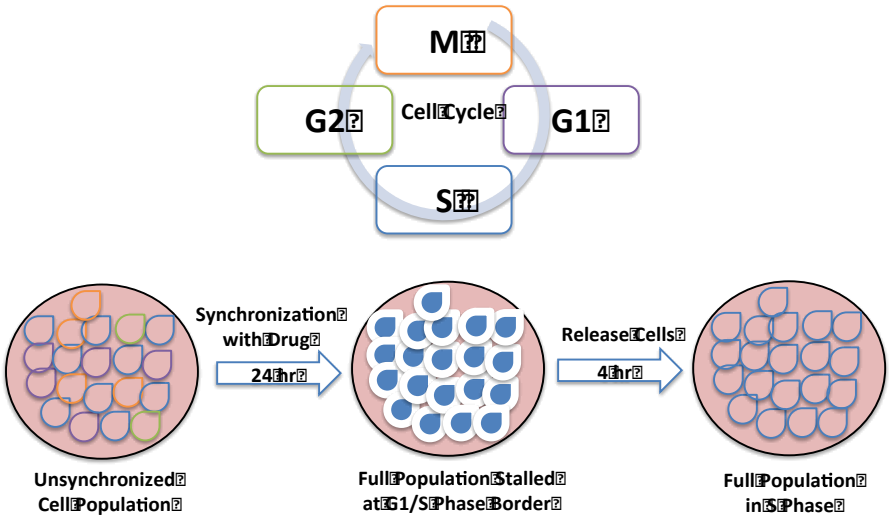


Figure 4. Cell Cycle Synchronization.
The frequency of gene editing can be raised if the cells targeted are progressing through S phase. An unsynchronized cell population contains cells at multiple stages of the cell cycle. A 24hrs drug treatment stalls all the cells at the G1/S phase border; it is this step that synchronizes the cell population. Releasing the cells for 4 hours after washing out the drug permits the cells to progress through S phase and it is at this time that they are targeted to raise the gene editing frequency.

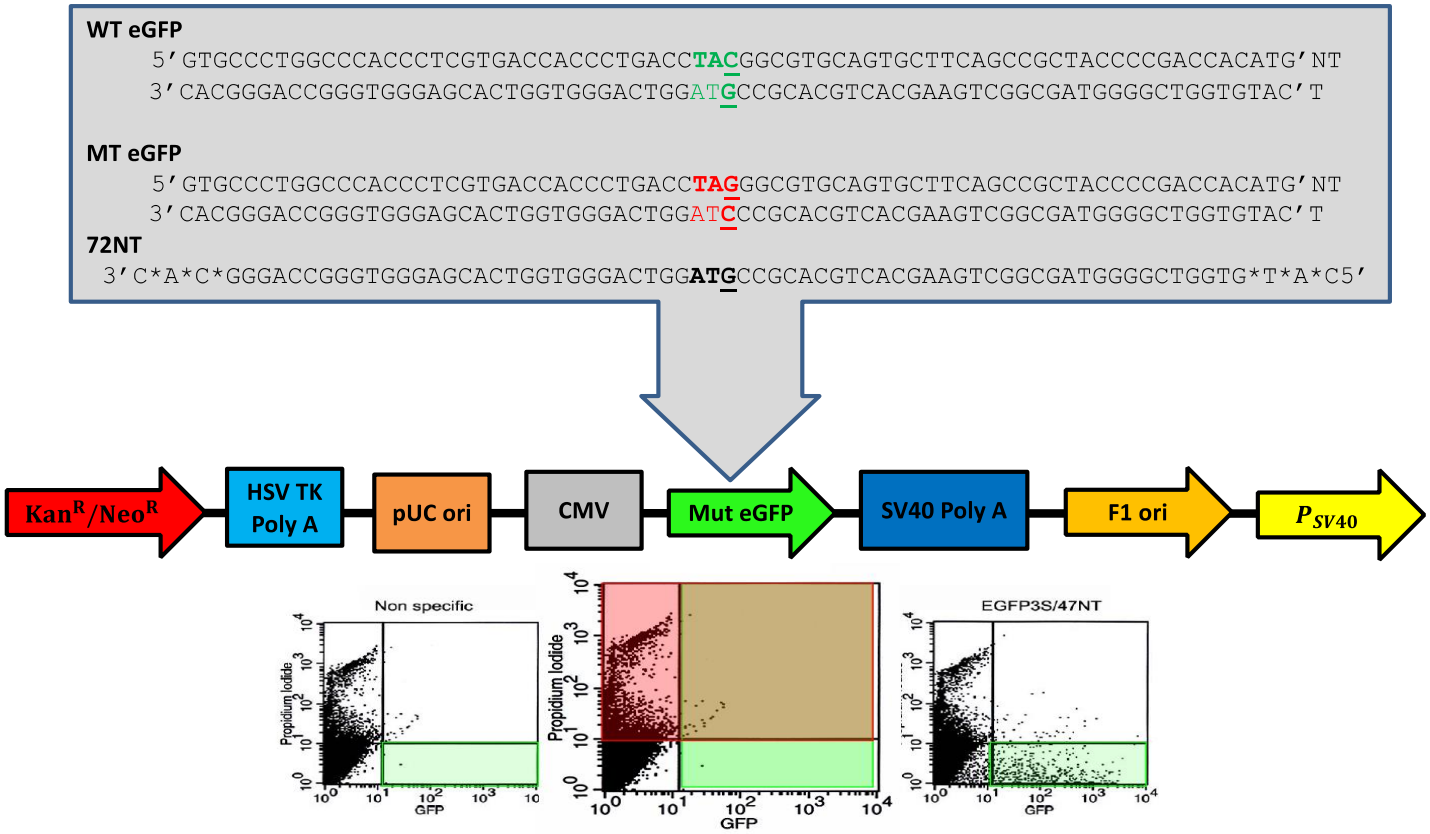


Figure 5. HCT 116-19 Gene Editing Model System.
The wild type and mutated eGFP gene segments with the target codon located in the center of the sequences are displayed in green and red respectively. The integration vector used to insert a single copy of the gene driven by a CMV promoter in HCT116 cells is showed in the middle. The population of cells that have undergone the gene editing reaction are analyzed by Flow Cytometry or Fluorescence-activated cell sorting (FACS) and the percentage of live green fluorescent cells within that population is presented as a frequency of gene correction. Adapted from: [37]

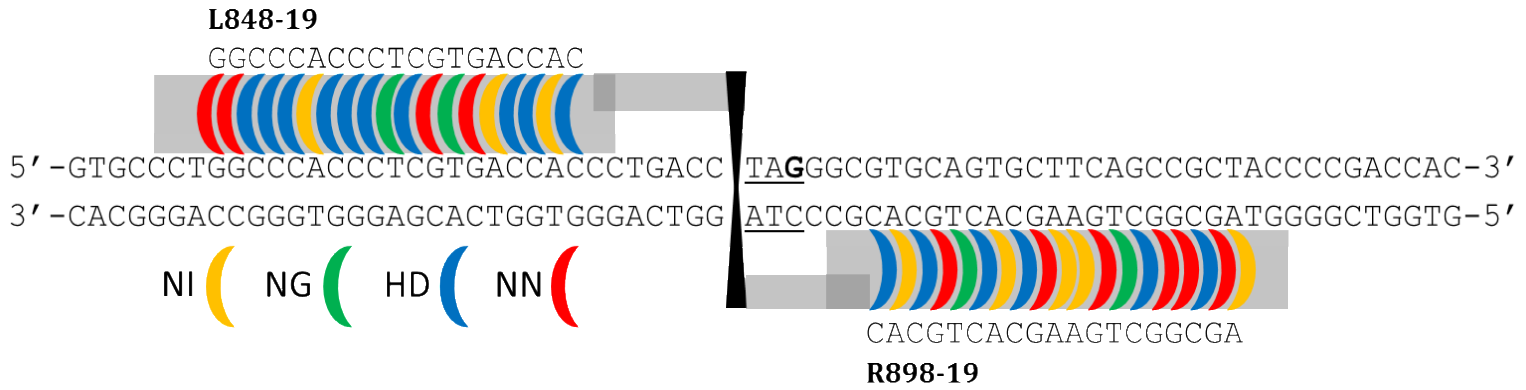


Figure 6. TALEN Binding Site Model.
The TALEN pair induces a double stranded break preceding the mutant codon. RVDs are shown as color coded binding blocks next to their respective base, yellow NI: A, green NG:T, blue HD: C and red NN: G. Fok1 domains are represented by the black and are positions at the predicted cut site. Source: [37].

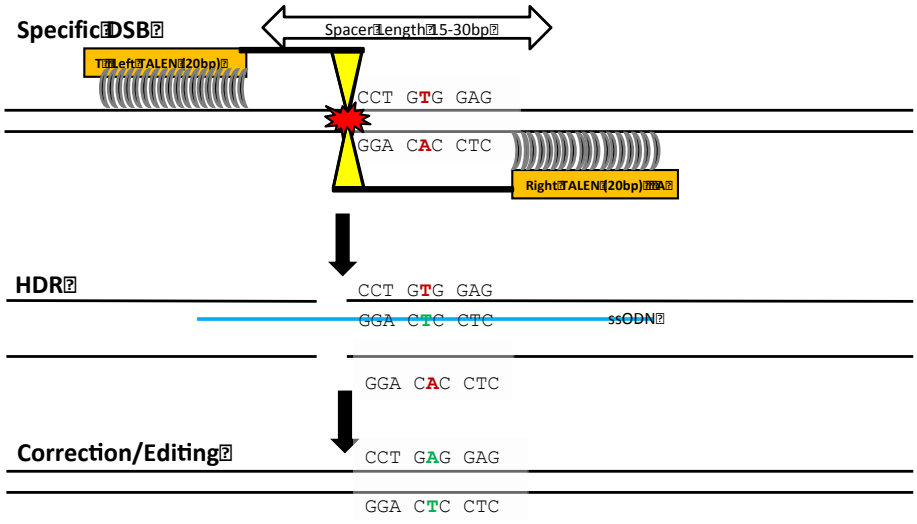


Figure 7. Gene Editing Combining TALENS and ssODNs.
TALENs produce a specific double stranded break at a specific site near the target base intended to be edited. The ssODN has a specific entry point and can serve as a template in homology directed repair (HDR).

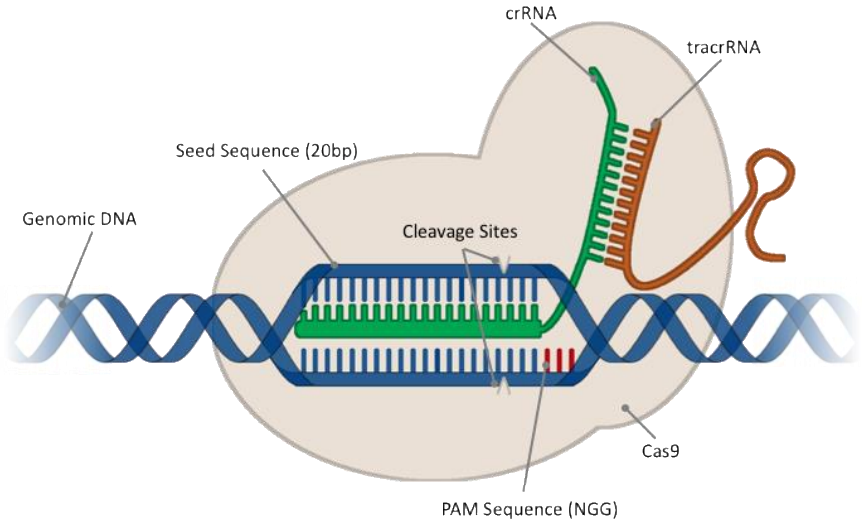


Figure 8. Schematic of the CRISPR/Cas9 Binding Site Model.
The Protospacer Adjacent Motif (PAM) NGG bases of the target DNA strand are shown in Red. If the target DNA is complementary to the guide RNA strand, the two

strands will base pair. This will allow the target DNA to unzip, as the bases flip up and bind the guide RNA. Complete annealing of the guide RNA to the target DNA allows the HNH and RuvC nucleases to cleave their respective strands.

Chapter 2

REGULATION OF GENE EDTITING ACTIVITY DIRECTED BY SINGLE-STRANDED OLIGONUCLEOTIDES AND CRISPR/CAS9 SYSTEM

Introduction

Single base mutations can be repaired by introducing single stranded DNA oligonucleotides (ssODN) into a target cell [3–5]. The frequency of this corrective activity is dependent on a set of factors: the length and polarity of ssODN [44], the position of the cell in its proliferative cycle [6,7], and the presence of double-stranded DNA breaks in the host genome [8,9]. The design, construction and effective expression of nucleases that can produce site-specific double-stranded DNA breaks in mammalian cells is challenging. Overcoming these limitations has been the predominant reason the rapid development and acceptance of a new tool for gene editing; Clustered, Regularly Interspace, Short Palindromic Repeats (CRISPR) molecules associated with Cas9 endonucleases (CRISPR/Cas9).

We tested the ease-of-use of the CRISPR/Cas9 system for its combinatorial capacity to promote site directed gene editing with ssODNs under previously determined optimal reaction parameters. This was done by evaluating the CRISPR/Cas9-driven activity at the same genetic sites successfully edited or repaired previously by the combinatorial approach of TALENs and ssODNs [44]. Editing activity was detected in reactions utilizing the CRISPR/Cas9 constructs than TALENs at several specific sites but both appear to be influenced by similar reaction parameters. First, ssODNs that hybridize to the non-transcribed strand direct a higher

level of gene repair than those that hybridize to the transcribed strand. Second, cleavage must be proximal to the targeted mutant base to enable higher levels of gene editing. Third, DNA cleavage enables a higher level of gene editing activity as compared to single-stranded DNA nicks, created by modified Cas9 (Nickases). Fourth, the hybridization potential and free energy levels of ssODNs that are complementary to the guide RNA sequences of CRISPRs were calculated. This study determined a correlation between free energy potential and the capacity of single-stranded oligonucleotides to inhibit specific DNA cleavage activity, indirectly reducing gene editing activity.

Methods and materials

Cell Line and Culture Conditions

HCT116 cells were acquired from ATCC (American Type Cell Culture, Manassas, VA). The HCT116-19 was created by integrating a pEGFP-N3 vector (Clontech, Palo Alto, CA) containing a mutated eGFP gene. The mutated eGFP gene has a nonsense mutation at position +67 resulting in a nonfunctional eGFP protein[85]. For these experiments, HCT116 (-19) cells were cultured in McCoy's 5A Modified medium (Thermo Scientific, Pittsburgh, PA) supplemented with 10% fetal bovine serum, 2mM L-Glutamine, and 1% Penicillin/Streptomycin. Cells were maintained at 37°C and 5% CO_2. Custom designed oligonucleotides, 72NT, 72T and 72NT PM were synthesized from IDT (Integrated DNA Technologies, Coralville, IA).

CRISPR Design and Construction

The mutant eGFP gene sequence was entered into the Zhang Lab's online generator (http:// crispr.mit.edu/) and the five CRISPR guide sequences which bind

upstream and downstream with close proximity to target (TAG = 0) were chosen. The CRISPRs were constructed using standard cloning methods following the latest oligo annealing and backbone cloning protocol with single-step digestion-ligation[54]. The five CRISPR guide sequences were cloned into the pX330 backbone vector (Addgene plasmid 42230), a human codon-optimized SpCas9 and chimeric guide RNA expression plasmid, and into pX460 backbone vector (Addgene plasmid 48873) which is a D10A nickase mutant human codon-optimized SpCas9 and chimeric guide RNA expression plasmid. pX458 (Addgene plasmid 48138) was a gift from Feng Zhang and is a human codon optimized pSpCas9 and chimeric guide RNA expression plasmid with a 2A- eGFP. All plasmids were purchased through Addgene (https://www.addgene.org). Following construction, clones were verified by DNA sequencing by Genewiz Incorporated (South Plain- field, NJ).

Transfection of HCT116-19 Cells and Experimental Approach

For experiments utilizing synchronized cells, HCT116-19 cells were seeded at 2.5 x 10^6 cells in a 100mm dish and synchronized with 6μM aphidicolin for 24 hours before targeting. Cells were released for 4 hours before trypsinization and transfection by washing with PBS (-/-) and adding complete growth media. Synchronized and unsynchronized HCT116-19 cells were simultaneously transfected at a concentration of 5 x 10^5 cells/100μl in 4mm gap cuvette (BioEx- press, Kaysville, UT). Single-stranded oligonucleotides and CRISPR or Nickase plasmid constructs were electroporated (250V, LV, 13ms pulse length, 2 pulses, 1s interval) using a Bio- Rad Gene Pulser X Cell Electroporation System (Bio-Rad Laboratories, Hercules, CA). Cells were then recovered in 6-well plates with complete growth media at 37°C for 48 hours before analysis.

Analysis of Guide RNA and DNA Oligo Hybridization

Each guide RNA sequence and the 72NT oligo sequence were aligned and analyzed for base pairing and maximum ΔG values utilizing Oligo Analyzer 3.1 (https://www.idtdna.com/ analyzer/Applications/OligoAnalyzer/). The ΔG is calculated by the longest stretch of complementary bases between the DNA and RNA structures the maximum ΔG value is determined as the free energy of the RNA sequence binding to its complement.

Analysis of Gene Edited Cells and Transfection Efficiency

Fluorescence (eGFP+) was measured by a Guava Easy Cyte 5HT Flow Cytometer (Millipore, Temecula, CA). Cells were harvested by trypsinization, washed once with 1x PBS (-/-) and re-suspended in buffer (0.5% BSA, 2mM EDTA, 2μg/mL Propidium Iodide (PI) in PBS-/-). Propidium iodide was used to measure cell viability viable cells stain negative for PI (uptake). Correction efficiency was calculated as the percentage of the total live eGFP positive cells over the total live cells in each sample. Error bars are produced from two sets of data points generated over two separate experiments using basic calculations of Standard Error. Statistical significance was performed by using two-sample unequal variance students T-test distribution to compare the value. $p<0.05$ Sequence confirmation of ssODN/CRISPR edited cells was carried out by fluorescence-activated cell sorting of eGFP+ cells using a BD FACSAria II sorter- 488nm (100mw) (BD Biosciences, San Jose, CA). 1.35μg 72NT and 2μg CRISPR 2C transfected cells were sorted at 72 hours post electroporation. Immediately, DNA was isolated from each sample was using the DNeasy Blood and Tissue kit (Qiagen, Hilden, Germany). The targeted site was amplified via PCR using forward primer, 5'CTGGACGGCGACGTAAACGGC and reverse primer, 5' ACC

ATGTGATCGCGCTTCTCG. PCR cleanup was performed using the QIAquick. PCR purification kit (Qiagen, Hilden, Germany) and the purified samples were sent for sequencing to Gene- wiz Incorporated (South Plainfield, NJ). Unsynchronized HCT116-19 cells were harvested and electroporated at a concentration of $5x10^5$cells/100µl with 2µg of the indicated CRISPR/Cas9 (2C, 3C, 5C and empty pX458 vector) plus 1.35µg of either 72NT or 72 PM. Following electroporation, transfection efficiency is determined after 24 hours of incubation by the percentage of total viable eGFP+ cells in the population. The normalized correction efficiency was determined after 48 hours of incubation as the percentage of total viable eGFP+ cells in the population divided by the transfection efficiency.

RFLP Analysis of CRISPR/Cas9 Cleavage Activity

HCT116-19 test samples were electroporated at a concentration of 5×10^5 cells/100µl in 4mm gap cuvette (BioExpress, Kaysville, UT) with 2µg of CRISPR/Cas9 constructs 2C and 3C and 2µg 2C + 1.35µg 72NT PM and 2µg 3C + 1.35µg 72NT PM. Cells were then recovered in 6-well plates with complete growth media at 37°C for 72 hours. DNA was isolated using the Blood and Tissue DNeasy kit (Qiagen, Hilden, Germany). RFLP analysis was performed on 181bp amplicons that were created using forward primer, 5'GAGGGCGATGCCACCTACG GC and reverse primer, 5'GGACGTAGCCTTCGGGCATGGC. PCR samples were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) and treated with the AvrII restriction enzyme following the manufactures protocol. Digested samples were loaded along with NEB 2-log DNA ladder (NEB, Ipswich, MA) into a 2% TBE agarose gel for analysis. SYBR Gold (Invitrogen, Carlsbad, CA) was used to stain the gel and images were acquired by the Gel Doc EZ System (BioRad, Hercules, CA) to

create an electrophoregram. Using Bio Rad's Image Lab software, automated lane detection was performed, followed by selecting bands. Using the software, the concentration of each band represented by a peak on the electrophoregram was derived from the area of each peak as a percent of the total lane peak area.

SURVEYOR Analysis of CRISPR/Cas9 Cleavage Activity

HCT116-19 cells were electroporated at a concentration of 5 x 10^5 cells/100µl in 4mm gap cuvette (BioExpress, Kaysville, UT) with 2µg of each CRISPR/Cas9 construct. Cells were then re- covered in 6-well plates with complete growth media at 37°C for 72 hours. Genomic DNA was extracted and purified using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer's protocol. The genomic region surrounding the mutant eGFP gene locus targeted by each gRNA and an untreated sample was PCR amplified using Phusion High-Fidelity PCR Master Mix with HF Buffer (Thermo Scientific). 200ng of each PCR product was mixed with 200ng of PCR product from the untreated sample and subjected to a heteroduplex formation: 95°C for 10 minutes, 95°C to 85°C with a ramp rate of -2°C/s, 85°C for 1 minute to 75°C at-.1°C/s, 75°C for 1 minute to 65°C at-.1°C/s, 65°C for 1 minute to 55°C at-.1°C/s, 55°C for 1 minute to 45°C at-.1°C/s, 45°C for 1 minute to 35°C to 25°C at-.1°C/s, 25°C for 1 minute. After duplex formation products were treated with SURVEYOR Nuclease S and SURVEYOR Enhancer S (IDT Technologies) for 30 minutes at 42°C, gel electrophoresed and stained with SYBR Safe DNA stain (Life Technologies). Gels were imaged with a Gel Doc EZ Imager (Bio-Rad) and densitometry was performed by measuring the area under the curves of each band, using the Image Lab software (Bio-Rad). Calculations were based on the following formula:

% cleaved= sum of cleaved products/sum of cleavage products + parent band

Results

What are the parameters for CRISPR/Cas9 systems with ssODNs for gene editing?

The ease-of-use of the CRISPR/Cas9 system for its combinatorial capacity to promote site directed gene editing with ssODNs under previously determined optimal reaction parameters was tested by evaluating the CRISPR/Cas9-driven activity at the same genetic sites successfully edited or repaired previously by the combinatorial approach of TALENs and ssODNs [44].

The CRISPR and Nickase constructs and single-stranded oligonucleotides were designed to target the mutated eGFP gene in HCT 116 cells (Figure 9A). Cells were transfected through electroporation with a specific CRISPR /Cas9 construct or the Nickase construct and single-stranded ssODNs. The specific design and locations of all the CRISPR/Cas9 complex cleavage and the appropriate, specific single-stranded oligonucleotides used in targeting are presented in Figure 9B and Figure 10. Each arrow indicates the position of cleavage by the appropriate CRISPR/Cas9 with the guide RNA sequences depicted in blue. The guide RNA sequence is also aligned above or below based on its complementarity to the gene. Above and below the target sequence are the 72-base long single-stranded oligonucleotides used in the reaction (72T, 72NT and 72NT-PM). The 72NT ssODN is complementary to the top (NT) strand and therefore complementary to a guide RNA sequence that hybridizes to the bottom/T strand. 72T has the opposite polarity being complimentary to the bottom or T strand as shown in Figure 9B. CRISPRs 2C and 3C were designed to cut at the same site with a difference in the strand each guide RNA is complimentary to. Since

complimentary guide sequences can hybridize to the single-stranded oligonucleotide with opposite polarity, but bearing complementarity, the guide and ssODN were aligned to calculate the free energy measurements of the heterodimer formation. This data is useful to better hypothesize the expected efficiency of the gene editing reaction at the target site for all combinations of interest. 3C has a ΔG of hybridization of -37.6 kcal/mole and thus has a significant probability of becoming annealed whereas 2C and 5C have ΔGs of -10 kcal/mole (Figure 11). 4C has a ΔG of approximately -12 kcal/mole because there are only six bases in common whereas the guide sequence in CRISPR 3C is completely complementary therefore it is expected to anneal completely to the ssODN.

Experiments where a DSB has been introduced by a programmed nuclease, TALENs, the cleavage 5' (upstream) relative to the target nucleotide (here the G of TAG) was the optimal site for increasing gene editing activity. However when the DSB was introduced 3' to the target base the gene editing activity is reduced [44]. This 5' cleavage for optimal gene editing was tested using the different CRISPR constructs described above. Cells were synchronized for 24 hours and then released for four hours before electroporation. Synchronized and unsynchronized cells were electroporated with ssODNs and the appropriate CRISPR/Cas9 complex. After 48–72 hours of recovery, the cells were analyzed by flow cytometry to measure the correction of the eGFP gene (Figure 12) and by RFLP to confirm specific CRISPR/Cas9 cleavage activity at the target site (Figure 13). In Figure 12 the left side bar graphs the 72NT was used combined with each individual CRISPR/Cas9 complexes. Gene editing levels vary widely among the five complexes used in the experiment. CRISPR 2C and 5C promote maximum activity. This is due to their

proximal cleavage to the targeted G nucleotide. 5C cleaves one base 3' to the targeted base and 2C cleaves 3 bases up-stream. The CRISPR/Cas9 3C does show some gene editing activity but at lower levels although its cut site is the same as the 2C construct just annealing on opposite strands. On the right side of the graph, the 72T ssODN that hybridizes to the transcribed strand of the gene was used combined with the same CRISPR/Cas9 complexes used in the previous set. Consistent with previous data [5,6,34,36,44,86–90] the 72T enables a lower level of gene editing activity, independent of the CRISPR/Cas9 complex used. In accordance to the optimal parameters for gene editing the CRISPR/Cas9 complexes that cleaved near the targeted nucleotide exhibit the highest level of gene editing activity. When distal cuts sites to the target base were produced by either CRISPR 4C or 1C, no gene editing was achieved independent of which ssODN is used.

Another parameter that has been identified to enhance the gene editing reaction in the presence of a DSB and ssODN is the cell cycle stage in which the cells are positioned at the time of targeting. When cells are synchronized at the G1/S border and released for four hours before the introduction of the editing tools the frequency of HDR is higher in comparison to unsynchronized cells that have undergone the same gene editing reaction [7,36,40,86,91–93]. To examine the effect of cell synchronization and release on CRISPR/Cas9 targeting, driven by ssODNs, cells were synchronized at the G1/S border for 24 hours and then released for four hours before CRISPR/Cas9 complexes 2C, 3C and 5C and the 72NT ssODN were introduced. Gene editing efficiency was measured 48 hours post electroporation. The results compiled in Table 2, support the need to have the cells synchronized and released prior to the addition of the gene editing reaction components, in this case CRISPR/Cas9 complex

and ssODN, in order to achieve enhanced optimal levels of gene editing. The results in Table 2 indicate that the CRISPR/Cas9 complex, 3C, catalysis gene editing events at a 4 to 5-fold lower frequency than its counterpart CRISPR/Cas9 complex 2C. The guide sequence of 2C anneals to the NT strand of the eGFP gene. In contrast the guide sequence of 3C anneals to the T strand (see Figure 9B). As described before the 72 NT is complementary to the guide sequence of 3C but is not complimentary to the guide sequence of 2C. Thus, in agreement with the data in Figure 10 where the maximum Δ G of 3C and the 72 NT ssODN is -37.6, the low frequency of gene editing activity is most likely due to the annealing of the guide sequence of 3C with the 72 NT ssODN due to its complementarity.

To better determine the hybridization of the 3C guide RNA to the 72 NT ssODN and its effect on the gene editing efficiency outcome, a reaction utilizing the 2C and 3C CRISPR/Cas9 complexes in the presence or absence of a perfect match single-stranded oligonucleotide complementary to the non-transcribed strand (72 NT–PM) was tested. The 72 NT–PM is complementary to the NT strand as the 72 NT ssODN with the difference of not containing a mismatch to the mutant eGFP. The use of the matched oligonucleotide will prevent the loss of the site for the restriction enzyme AvrII that would be modified by the in the presence of the 72NT ssODN. After 72 hours of reaction time, genomic DNA was isolated and a 181 base pair fragment, spanning the target site, was amplified by PCR. The amplified fragment was digested by the restriction enzyme AvrII which cleaves at the target region and thus can be an indicator of double strand DNA cleavage activity[44]. The resultant mixture of DNA fragments were electrophoresed through agarose gel and the results are presented in Figure 13A. The uncut fragment migrates to a position consistent with its

size of 181 bases whereas the AvrII-treated DNA is cleaved to completion (90 and 91 bases, respectively). In the reactions containing either 2C or 3C in the absence of the ssODN the results show high levels of CRISPR/Cas9 cleavage activity. Densitometry tracing place the level of resistant bands at 33% for 2C and 32% for 3C respectively. In the presence of the 72 NT–PM ssODN the reaction mixture with the 2C complex, the level of cleavage activity remains identical to reaction mixtures that included only 2C. A barely detectable level of a resistant band is seen in reaction mixtures which contain 3C and 72 NT–PM. These results follow the notion that the 72 NT ssODN used in the gene editing reactions is hybridizing to the 3C CRISPR/Cas9 complex inhibiting its potential to cut at the target site. In contrast the 2C guide RNA will not hybridize to the 72 NT ssODN given its lack of complementarity, maintaining its full capacity to full fill the gene editing reaction observed by its high gene editing efficiency (Figure 12). To further test the CRISPR/Cas9 activity the SURVEYOR cleavage assay[94] was also performed (Figure 13B). All five CRISPR/Cas9 constructs tested showed to have DNA brake activity by showing cleavage. The purpose of this assay was not to develop correlations between the level of gene editing and DNA cleavage since this assay is sensitive to some variability and activity can depend on background signals when polymorphisms are present in the target genome[95]. The SURVEYOR assay was employed to support as an indicator of the CRISPR/Cas9 general cleavage activity of the expression constructs.

While CRISPR/Cas9 complexes can induce efficient cleavage and promote enhanced levels of gene editing, directed by single-stranded oligonucleotides, concern has been raised that intact CRISPR/Cas9 molecules promote high levels of off-site mutagenesis. In response a significant effort has been put into constructing a variation

of the CRISPR/Cas9 developed into single Nickase enzymes that cleave only one strand of the dsDNA, which have been reported to reduce off-site mutagenesis[54,96]. To test the Nickase activity in the catalysis of gene editing on the mutant eGFP gene complexes consisting of a CRISPR/Cas9 in which only one nuclease domain was functional were used. The single strand cut is depicted and described in Fig 9B. In the top panel of Figure 14, the red arrows indicate the position of an individual Nickase (N) cleavage. The 4N cleaves upstream from the target base but on the T strand whereas Nickase 3N cleaves the T strand, but only a few bases up- stream from the target base. 2N, 5N and 1N cleave the NT strand alone at the indicated positions. Each of these Nickase complexes were in combination of the 72 NT ssODN and the level of correction efficiency monitored by FACS at 48hrs. In line with the previous data described above the cleavage proximal to the target site provides the highest level of gene editing activity. Cleavage at distal sites promotes lower activity. However, the editing efficiency is 10 to 15-fold lower in the presence of a nick, dsb induced by the complete CRISPR/Cas9 complex is used. Although the intended use of the Nickase is to avoid potential off-site mutagenesis, the level of gene editing using ssODN is greatly diminished when a single-strand cut on only one of the two strands of the DNA is made. To further study the potential of Nickase enzymatic activity these were used to reconstitute the dsb by combining two complexes. Interestingly, the combination of Nickase did not promote sufficient activity to resurrect the same level of gene editing efficiency as produced by the intact CRISPR/Cas9 complexes. Most of these combinations seen in Figure 15, including 3N/5N, where some gene editing activity was able to be measured in reactions that produced a dsb at or 5’ to the target base which are key parameters for efficient gene editing to occur. It is possible, as

described above; Thus some combinations do not work due to the 3N and 4N guide RNA sequences that can be hybridize to the ssODN, impeding their function, while synergistic activity is observed when 2N and 5N are used.

In previous work the boundaries of efficient gene editing directed by TALENs and single-stranded oligonucleotides have been defined as a region spanning the target nucleotide [44]. It has also been established that moderate gene editing (approximately 1%) diminishes upstream when the distance is expanded from -8 /-9 to -28 respectively. The distance downstream appears to be more restricted with editing levels reduced between +6\+7 and +8\+10[44]. Although data set is admittedly low, significant levels of gene editing at position -4 (upstream) are observed while no activity is seen at position -33 (upstream). The data generated in the present work is overlaid with the previous TALEN/ssODN data (Figure 16) while statistical analysis is presented in Table 3. In direct comparison, CRISPR/Cas9 complexes that act at proximal regions also approximate the highest level of activity as the levels promoted by the TALEN and ssODN combination that act in the same region. Both targeted nucleases support the notion that creating a dsb near the targeted nucleotide, destined for change, promotes the highest level of gene editing activity. When the mutant eGFP gene editing system for studying the mechanism of action was developed, the gene correction frequency was defined as the percent of eGFP+ cells within the entire viable population of cells treated[3,39,40]. The rationale was that this frequency will more adequately represent correction presumably attainable in primary cells, where it will not be practical to only select transfected cells to measure actual correction levels. An alternative for calculating gene editing frequencies based on the percent of cells that actually receive the CRISPR/Cas9 construct is by utilizing a CRISPR/Cas9

plasmid that exhibits fluorescence, hence transfected cells can be identified. Table 4, compiles the gene editing efficiencies for the three most active CRISPR/Cas9 complexes (vectors lacking the eGFP marker) that best catalyze gene editing adjusted to the transfection levels of the pX458 eGFP expression CRISPR construct in the presence of the 72NT oligonucleotide. The transfection efficiency was based on the number of cells exhibiting fluorescence. An obvious increase in correction frequency is observed even in unsynchronized cells, which is the true target population type for in vivo human therapeutic application. To confirm gene editing at the genomic level, we isolated eGFP positive cells were isolated by cell sorting followed by submitted the direct DNA sequence analysis of the samples. A typical and reproducible sequence result is presented in Figure 17. The upper panel represents a DNA sequence from a control mutant eGFP gene while the lower panel represents the DNA sequence of a corrected cell from the sorted population. The TAG to TAC base conversion is readily observable to be complete and thus these data confirm and correlate the observed phenotype with genotypic analysis.

Discussion

Single-stranded oligonucleotides and CRISPR/Cas9 complexes can be used in combination to direct nucleotide exchange in a precise and efficient fashion. The mutant eGFP gene was used as the target for gene editing reactions because it enables genotypic and phenotypic readout. The system provides a framework upon which mechanistic studies to define reaction parameters can be built. Introducing an expression construct containing a CRISPR/ Cas9 cleavage system and the appropriate single-stranded oligonucleotide leads to the correction of the mutant eGFP gene in the production of functional fluorescent protein. CRISPR/Cas9 complexes that vary in

position of cleavage, within the mutant eGFP gene, were found to produce a wide range of editing activities. Using an ssODN that is complementary to the NT strand in the editing reaction promotes a higher level of correction compared to the T strand. Such strand bias has been identified in many previous reports[6,34,36,44,86–90]. Synchronization at the G1/S border with subsequent release produces an enriched population of cells undergoing DNA replication. This manipulation of the cell population produces a more amenable environment for gene editing activity[6,38,41,97]. The phenomenon of strand bias, however is complex and likely to be determined or influenced by the strand serving as the template for lagging strand synthesis rather than the transcription template. The data from early studies led to the fundamental model outlined in 2007[3] expanded upon and reconfirmed in 2011[5]. These models are based on the foundational assumption that replication activity can modulate the mechanics of gene editing and the extent to which gene editing takes place when driven by single-stranded DNA. Lin et al[93] produced elegant data with CRISPRs suggesting that HDR is also occurring as the cells transit S-phase and possibly into G2, aligning with this notion. Ran et al (2013)[98] showed that when single-stranded or double-stranded donor DNA was introduced into a targeted genomic site, with the objective of DNA insertion, no strand bias was observed.

The objective of the experimental design throughout this work was to fix or repair a single base mutation without the need of the insertion of a donor fragment. This approach leads the reaction to be more dependent on DNA replication as well as a restriction on strand preference, as stated above. Different objectives using CRISPR and various donor DNA templates could take alternative routes and be governed by different reaction requirements. In example, Davis and Maizels [99] reported HDR is

more active on the transcribed strand as compared to the non-transcribed strand in their system. These data were obtained from experiments in which the outcome of HDR reaction was measured after being initiated by a nick on one of the two DNA strands. There was no indication as to which of these two strands served as a template for lagging strand synthesis, a key aspect of the strand bias phenomena observed in the gene editing reactions presented through our lab. In this and previous work, the same strand bias independent of DNA cleavage activity promoted by programmable nucleases is observed.

CRISPR/Cas9 complexes that cleave at proximal positions relative to the target base are more efficient in promoting the reaction, directed by ssODN. This result aligns with recent data from our lab and others [20,44,49] that suggest the double strand break enables integration of the oligonucleotide more efficiently at the proper site. Using modified CRISPR/Cas9 complexes, redesigned as single-strand endonucleases (Nickases), showed to promote gene editing activity but at a level roughly 90% less than the wild type Cas9 enzyme for applications involving single nucleotide exchange. The combination of two Nickases (double nicking) does not recapitulate the level of gene editing activity seen with the intact complex. The activity from the combination of 2N and 5N reveal an interesting data set. Since these two Nickases act on the same DNA strand, creating a gap that is large enough to promote hybridization of the ssODN. The single stranded character of such gapped DNA molecules can engage single-stranded DNA if the complementarity zone can be maintained for a certain period. By double nicking on the same strand, a long enough section of complimentary DNA may be available for productive annealing of the ssODN. However, gene editing execution is most efficient in the presence of a dsb

produced by a functional CRISPR/Cas9 complex. The offset lengths in Figure 15 represent the distance in base pairs between the PAM-distal (5') ends of the guide sequences of a pair of guide RNAs as defined by Ran et al. (2013)[96]. Following this convention, the enzymes used in this study range from +33bp to -96bp, producing nicks in a way which result in 3' overhangs. It has been reported that only sgRNA pairs creating 5' overhangs with offsets greater than -8bp between the guide sequences mediated detectable indels[96]. The detected levels of correction could be due to the result of indel formation.

The differing levels of gene editing activity promoted by CRISPR/Cas9 complex 2C in comparison to 3C are striking. Both cleave the gene at the same site, 5' upstream to the target base, and yet nucleotide exchange promoted by 2C occurs at a level 5-fold higher than the nucleotide exchange promoted by 3C. The difference in the constructs lies in their hybridization potential of the guide RNA sequence with the 72 NT ssODN. The free energy of pairing of 3C with 72 NT was much lower than that of 2C. The entire guide sequence of 3C can hybridized with a section of 72 NT. On the contrary the 2C CRISPR/Cas9 complex and 72 NT share the same polarity, therefore they cannot hybridize productively. This observation is reflected in the higher-level ΔG as seen in Figure 10. This suggests that the 72 NT ssODN is titrating the guide RNA sequence of 3C, reducing the overall effectiveness of 3C in binding properly to the target site and promoting efficient DNA cleavage. This hypothesis was directly tested and the data are presented in Figure 12A. Efficient cleavage activity of 2C and 3C should be easily identified by losing a restriction site in the DNA through the creation of a deleted DNA sequence. Here, the recognition site for AvrII would be lost if the DNA had been modified at or near the cleavage site. If 72 NT was titrating the

guide RNA of either 2C or 3C, it would reduce AvrII cleavage since CRISPR/Cas9 complex activity would have been inhibited by 72 NT. The results indicate that 72 NT inhibited only the cleavage activity of 3C, supporting the presented hypothesis. While the majority of CRISPR/Cas9 activity is directed toward gene knockout, when this genetic tool is used combined with single-stranded oligonucleotides, to direct single nucleotide exchange, several additional reaction parameters need to be considered. With the objective to enable single base repair, the data in the study suggest gene editing to be most effective with cleavage 5' or at a proximal position relative to the mutant bases intended to be corrected. These data align completely with published work when the combination of TALENs and single-stranded DNA oligonucleotides was used to direct gene correction[44,49].

Consistent with many other reports, oligonucleotides that are complementary to the NT strand of the gene are most effective in promoting gene editing[36,86]. There are several theories as to why such strand bias is observed[35,53,100], the non-transcribed strand of the target gene is also the lagging strand in DNA replication facilitating the incorporation of the oligonucleotide more easily into the growing replication fork[86]. When programmable nucleases are used to promote gene editing, the creation of an entry point, proximal to the target base, may provide a significant advantage in the gene editing reaction.

The work here has permitted us to identify an important difference in using TALENs versus RGEN technologies in gene editing. Since TALENs comprise binding domains coupled to a functional nuclease, the single-stranded oligonucleotide used to direct the nucleotide exchange will not affect TALEN cleavage nor TALEN activity since there is no guide RNA. With the data analyzed in this study, a series of

guidelines to assure efficient nucleotide exchange and gene correction can be developed. First, the oligonucleotide should be complementary to the non-transcribed strand of the gene. Second, the cleavage by the programmable nuclease should be within 20 to 50 bases, preferably upstream, of the nucleotide designated for change. Third, the nuclease system should be designed so the guide RNA sequence is of the same polarity as the non-transcribed strand to avoid hybridization to the single-stranded oligonucleotide which, usually will be complementary to the non-transcribed strand. And fourth, double-stranded DNA cleavage is more efficient than single-stranded cleavage in providing an amenable target (entry point) for gene editing by single-stranded oligonucleotides.

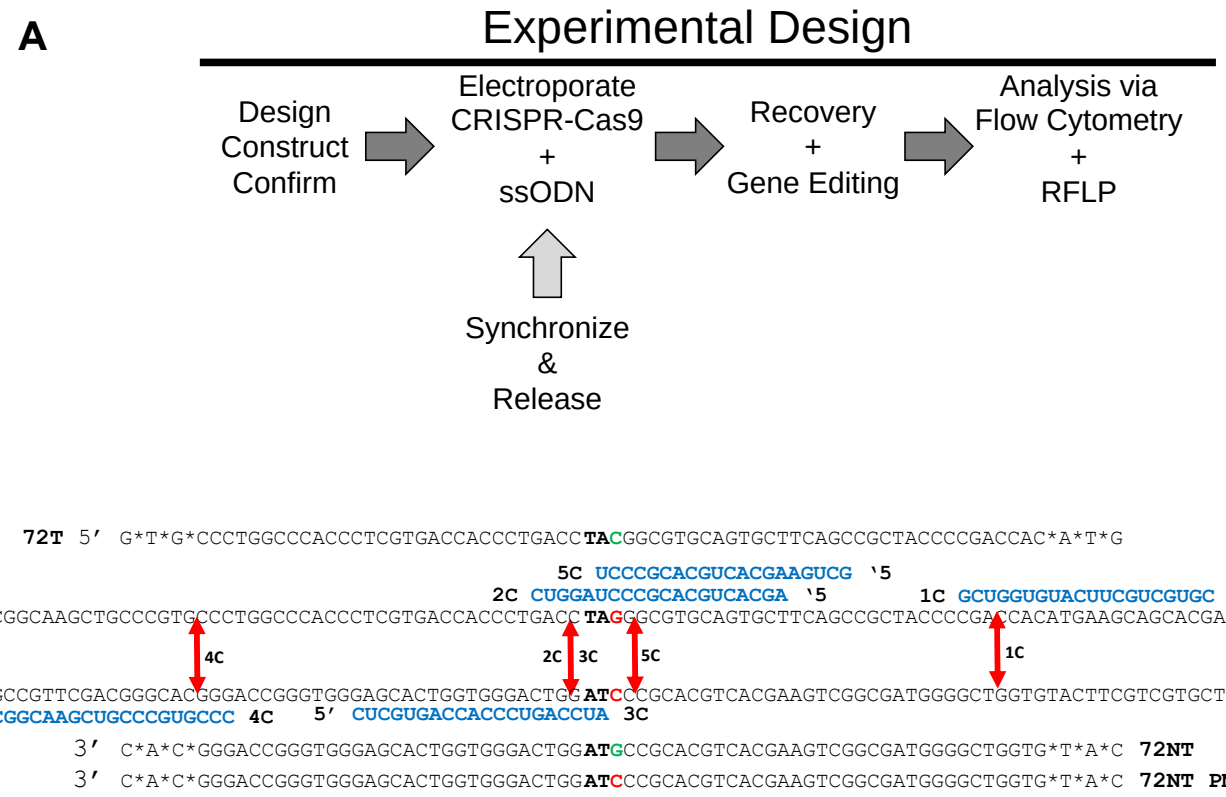


Figure 9. Experimental Design and Mutant eGFP Gene Editing System.
(A) CRISPR/Cas9 constructs were designed and built following published guidelines and protocols[98]. Unsynchronized or synchronized and released cells were transfected via electroporation with CRISPR/Cas9 construct and ssODNs and allowed to recover before analysis. Gene editing was assessed by flow cytometry and CRISPR/Cas9 activity was measured by RFLP using restriction enzyme AvrII. (B) A segment of the mutant eGFP gene is shown with the three ssODNs, 72NT, 72T and 72NT PM respectively aligned above or below their respective binding sites. Phosphorothioate modified and protected ends

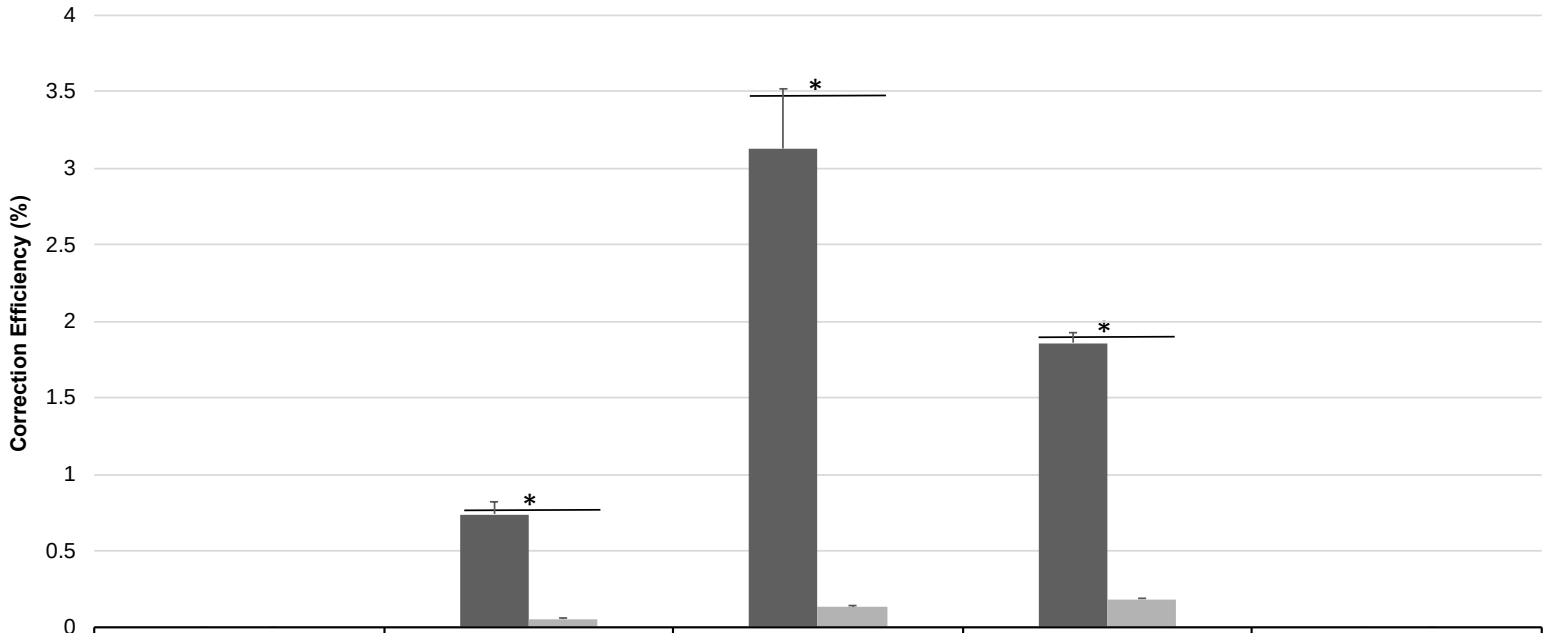
are denoted with *. The five arbitrarily named (1C-5C) custom CRISPR/ Cas9s RNA guide sequences are depicted in blue with their predicted cleavage sites shown by the red arrows. The effected codon is shown in bold with the mutant base to be edited in red. The base driving the gene editing conversion is shown in green.

CRISPR #	Score	Potential off Targets	# in genes	RNA Guide Binding Strand	Cut Site	CRISPR Binding Element(5'-3)	OLIGOs
1C	87	82	28	NT	+30	**CGACCACATGAAGCAGCACG** GCTGGTGTACTTCGTCGTGC	5'CACC G CGTGCTGCTTCATGTGGTCG C GCACGACGAAGTACACCAGC CAAA
2C	87	80	13	NT	-4	**GACCTAGGGCGTGCAGTGCT** CTGGATCCCGCACGTCACGA	5'CACC G AGCACTGCACGCCCTAGGTC C TCGTGACGTGCGGGATCCAG CAAA
3C	78	101	9	T	-4	CTCGTGACCACCCTGACCTA **GAGCACTGGTGGGACTGGAT**	5'CACC G CTCGTGACCACCCTGACCTA 3' C GAGCACTGGTGGGACTGGAT CAAA
4C	72	197	54	T	-33	CCGGCAAGCTGCCCGTGCCC **GGCCGTTCGACGGGCACGGG**	5'CACC G CCGGCAAGCTGCCCGTGCCC C GGCCGTTCGACGGGCACGGG CAAA
5C	81	139	27	NT	+2	**AGGGCGTGCAGTGCTTCAGC** TCCCGCACGTCACGAAGTCG	5'CACC GCTGAAGCACTGCACGCCCT 3' CGACTTCGTGACGTGCGGGA CAAA

Figure 10. Summary of Constructed CRISPR/Cas9 gRNAs.
Each designed guide RNA was generated from the crispr.mit.edu online algorithm. CRISPR # designates the gRNA name used in this study. Score represents the likelihood of the gRNA binding and causing unwanted mutations (score of 100 is the best possible gRNA). Possible off-target effects are the total number of individual loci across the genome that could be cleaved with the number of those within in genes (exons) listed in the following column. RNA guide binding strand denotes which stand the guide RNA will target (NT = non-transcribed, T = transcribed). Cut site is where the gRNA will direct the DSB break to be made relative to TAG = 0. CRISPR binding element shows the segment of the eGFP gene that the guide will bind with the actual bound strand in bold. Oligos used for construction with correct linkages are shown in the final column.

CRISPR gRNA	Sequence Alignment	Base Pairs	Maximum ΔG (kcal/mole)
1C	5' CGUGCUGCUUCAUGUGGUCG 3' CACGGGACCGGGTGGGAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTAC	7	-12.16
2C	5' AGCACUGCACGCCCUAGGUC 3' CACGGGACCGGGTGGGAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTAC	5	-10.88
3C	5' CUCGUGACCACCCUGACCUA 3' CACGGGACCGGGTGGGAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTAC	20	-37.6
4C	5' CCGGCAAGCUGCCCGUGCCC 3' CACGGGACCGGGTGGGAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTAC	6	-12.57
5C	5' GCUGAAGCACUGCACGCCCU 3' CACGGGACCGGGTGGGAGCACTGGTGGGACTGGATGCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTAC	5	-10.88

Figure 11. Free Energy Heterodimer Values for CRISPR Guide RNA and ssODN Combinations.
Each gRNA sequence was aligned and analyzed for base pairing and maximum free energy (ΔG) values utilizing the IDT heterodimer calculator and measured in kcals/mole. The highest ΔG value for each gRNA/ 72NT ssODN pairing is shown. Solid lines represent the longest stretch of direct base pairing while dotted lines represent additional complimentary bases (not calculated in ΔG). A more negative ΔG value represents a stronger binding capacity.



	4C	3C	2C	5C	1C
72 NT	0	0.74	3.13	1.86	0
72 T	0	0.05	0.13	0.19	0
P-value		0.0134	0.0165	0.0018	

Figure 12. CRISPR/Cas9 and ssODN Gene Editing Activity.
Unsynchronized HCT116-19 cells were harvested and electroporated at a concentration of $5x10^5$cells/100µl with 2µg of the indicated CRISPR/Cas9 plus 1.35µg of either 72NT or 72T. Following electroporation, cells were allowed to incubate for 48 hours and correction efficiency was determined by the percentage of total viable eGFP+ cells in the population. 1C– 5C CRISPR/Cas9 complexes and are listed left to right relative to their cut site. The numbers below indicate the average correction efficiency. Error bars represent standard error.

Table 2. S-Phase Increases CRISPR/Cas9 and ssODN Directed Gene Editing

CRISPR/Cas9	Unsynchronized Cells CE(%) ± SD	Synchronized Cells CE(%) ± SD	P-value
2C	3.13 ± 0.55	4.59 ± 1.39	0.3013
3C	0.74 ± 0.11	0.92 ± 0.03	0.1552
5C	1.67 ± 0.10	5.62 ± 0.23	*0.0019

HCT116-19 cells were seeded at $2.5x10^6$ cells in a 100mm dish and synchronized for 24 hours with 6µM aphidicolin then released for 4 hours. Synchronized and unsynchronized cells were electroporated at a concentration of $5x10^5$ cells/100µl with CRISPR/Cas9 and 72NT ssODN under the standard reaction conditions. Following electroporation, cells were seeded in 6-well plates and allowed to recover for 48 hours before flow cytometry analysis was carried out. Correction efficiency (%) was determined by the number of viable eGFP+ cells. Each sample set was performed in duplicate and ± represent calculated standard deviation per sample. The unsynchronized data is the same shown in Figure 11. Statistical analysis was performed using two-sample unequal variance students T-test distribution to compare the value of correction efficiency between synchronized and un-synchronized cells when treated with CRISPR/Cas9. *$p<0.05$

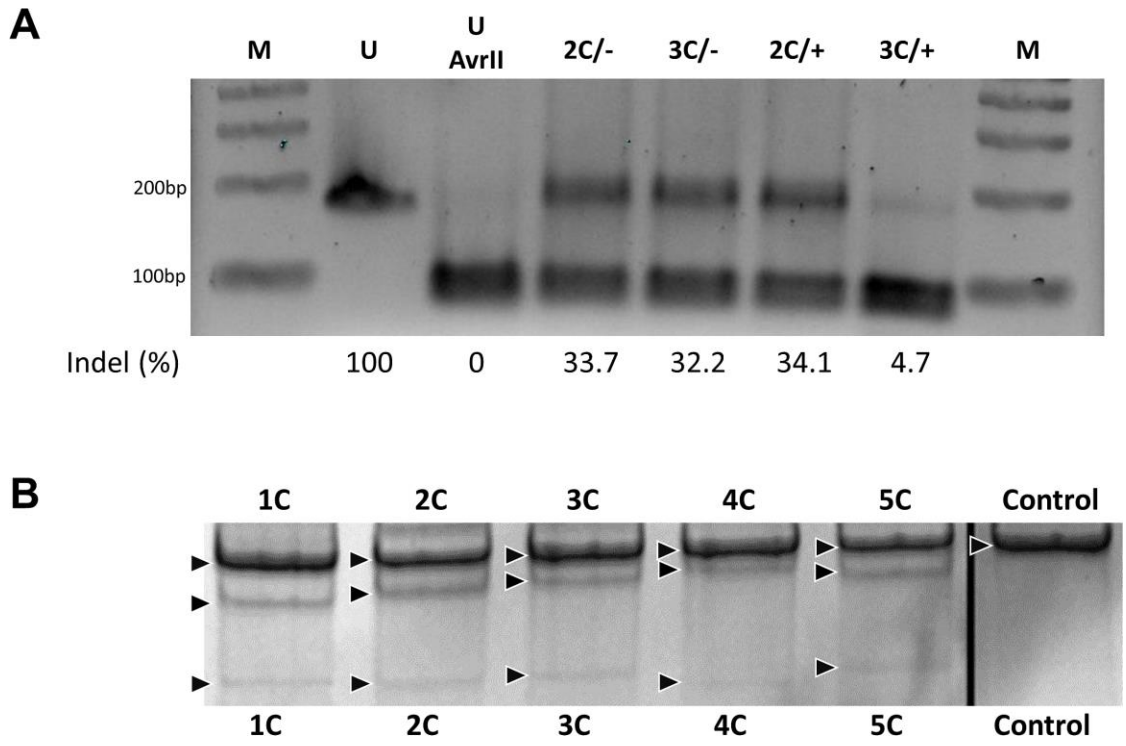


Figure 13. CRISPR/Cas9 Cleavage Activity.
(A) 2% TBE agarose gel analysis of cleavage products generated by 2C and 3C CRISPR/Cas9 complexes at the eGFP gene target site. AvrII restriction enzyme was used to digest the amplified region of the eGFP gene. 2C/-, CRISPR 2C with no ssODN; 2C/+, CRISPR 2C with 72NT etc. Untreated 181bp (U) and Untreated + AvrII 181bp PCR products were used as internal controls with AvrII digested 2C, 3C and 2C + 72NT PM and 3C 72NT PM 181bp PCR samples. Densitometry was performed on all samples and percent cleavage (181bp band) is indicated below each sample. (B) SURVEYOR assay comparing the cleavage efficacy of each gRNA as the percent of indel formation. Arrowheads indicate parental bands and cleaved products.

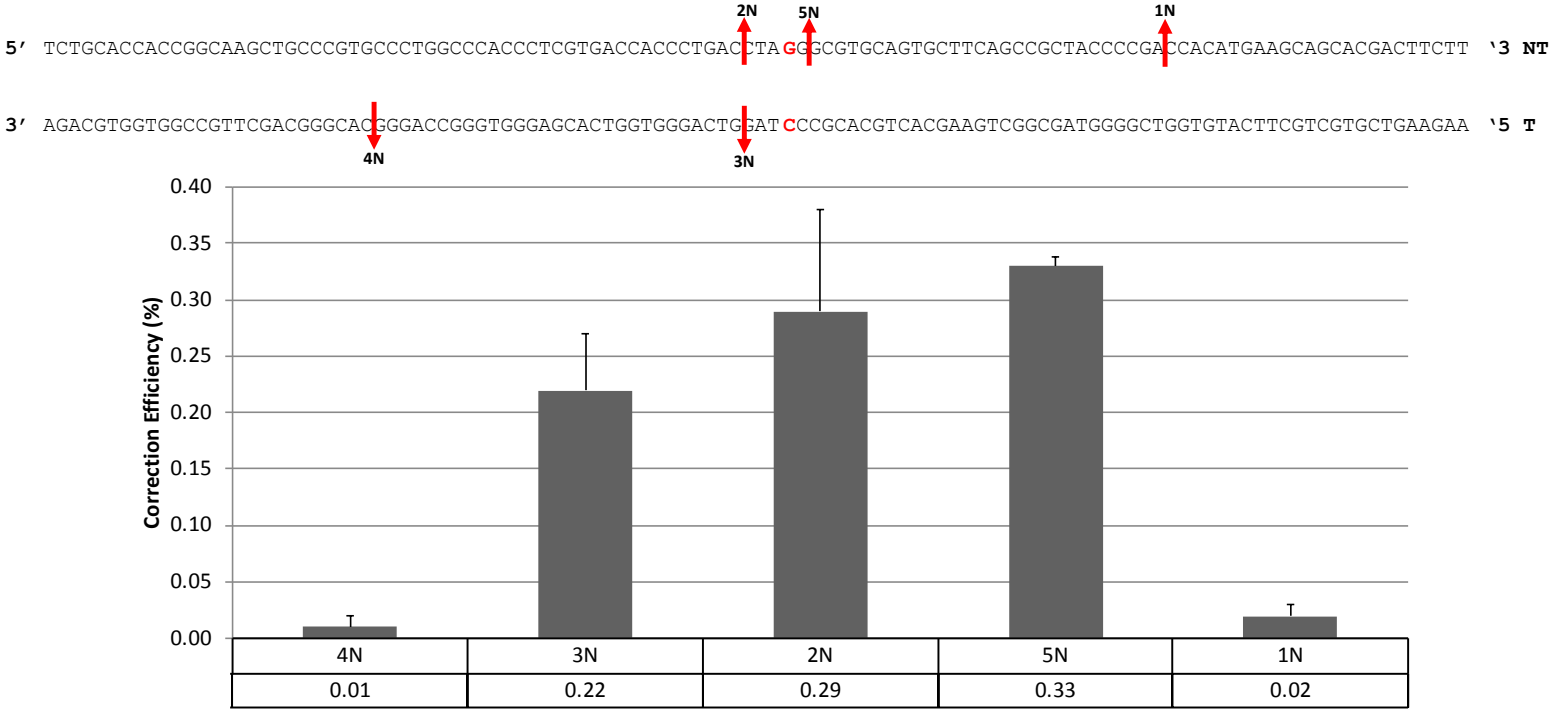


Figure 14. Gene Editing Activity with CRISPR/Cas9 Nickases and ssODNs.
Unsynchronized HCT116-19 cells were electroporated at a concentration of $5x10^5$ cells/100µl with 2µg of the indicated CRISPR/Cas9 Nickase (1N, 2N, 3N, 4N, 5N) plus 1.35µg of 72NT. Following electroporation, cells were allowed to incubate for 48 hours. Correction efficiency was determined by the percentage of total viable eGFP+ cells in the population as described previously. Each treatment was performed in duplicate and error bars represent standard error.

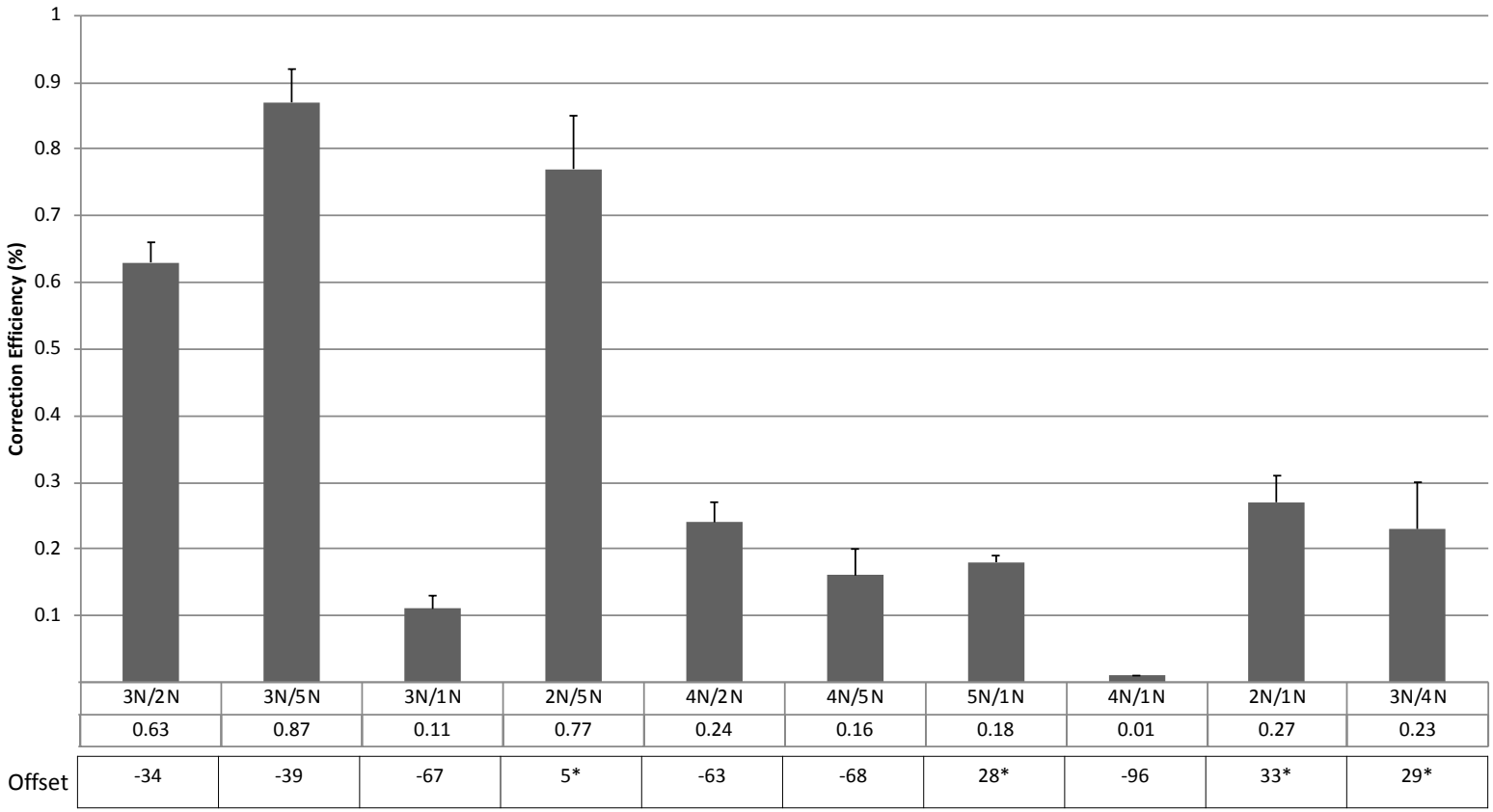


Figure 15. Double Nicking Nuclease Array of Gene Editing.
Unsynchronized HCT116-19 cells were electroporated with 1ug of each of the indicated combinations of CRISPR/Cas9 nickases (1N, 2N, 3N, 4N, 5N) plus 1.35µg of 72NT. Offsets denoted with a star (*) represent nicking pairs which induce nicks on the same strand. Following electroporation, cells were allowed to incubate for 48 hours. Correction efficiency was determined by the percentage of total viable eGFP+ cells in the population as described previously. Each treatment was performed in duplicate and error bars represent standard error.

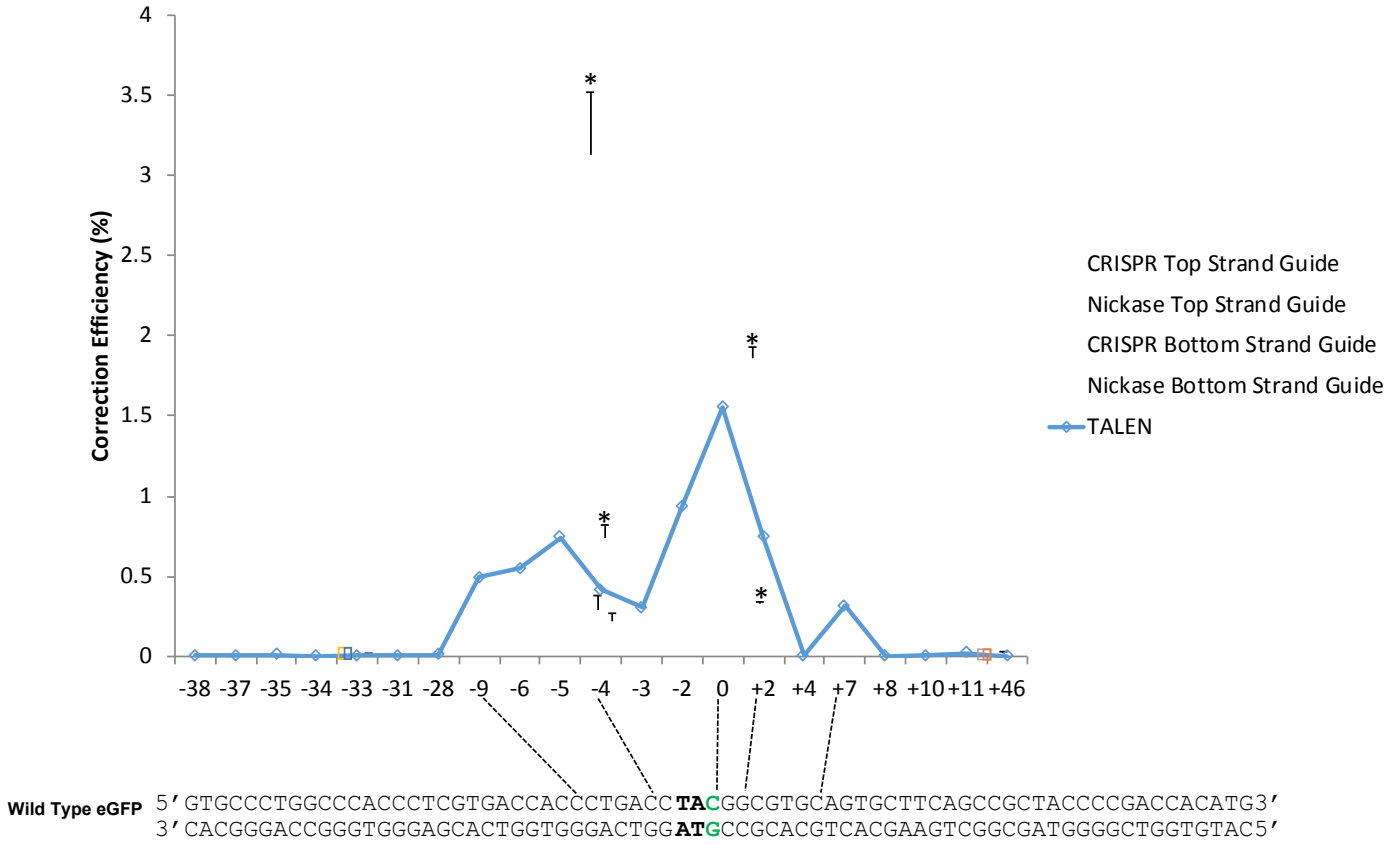


Figure 16. Activity Profile of Gene Editing catalyzed by TALENs, CRISPRs or Nickases at the target eGFP gene. ssODN directed gene editing activity utilizing TALENs, CRISPR/Cas9s or CRISRP/Cas9 nickases was compiled and plotted within the region of the target eGFP gene. Cytosine of the corrected tyrosine codon is designated as base 0. TALEN data was derived from previous work[44]. Statistical analysis was performed using two-sample unequal variance students T-test distribution. *$p<0.05$ (see Table 2)

Table 3. Comparing Activity of Gene Editing catalyzed by TALENs, CRISPRs or Nickases at the target eGFP gene

TALEN	CRISPR/Nickase	P-value (0.05)
-4	3C	*0.0198
-4	2C	*0.0517
-4	3N	0.2941
-4	2N	0.0628
-4	5C	*0.0106
-4	5N	*0.0473

Significance was determined using a T-test to compare the TALEN value to the CRISPR/Nickase value for each set presented in the table. A * was marked on graph to show those that were statistically significant

Table 4. Normalized Correction Efficiencies

CRISPR/Cas9	Dose	ssODN	Transfection (%)	Normalized Correction Efficiency (%)	SE
pX458	2µg	-	67.92	-	1.28
pX458	2µg	1.35µg 72 PM	45.96	-	0.42
2C	2µg	1.35µg 72NT	45.96	6.81	0.39
3C	2µg	1.35µg 72NT	45.96	1.49	0.08
5C	2µg	1.35µg 72NT	45.96	4.04	0.07

Unsynchronized HCT116-19 cells were harvested and electroporated at a concentration of $5x10^5$cells/100µl with 2ug of the indicated CRISPR/Cas9 (2C, 3C, 5C and empty pX458 vector) plus 1.35µg of either 72NT or 72 PM. Following electroporation, transfection efficiency was determined after 24 hours of incubation by the percentage of total viable eGFP+ cells in the population. The normalized correction efficiency was determined after 48 hours of incubation as the percentage of total viable eGFP+ cells in the population divided by the transfection efficiency. Standard Error was calculated from two sets of data points generated over two separate experiments.

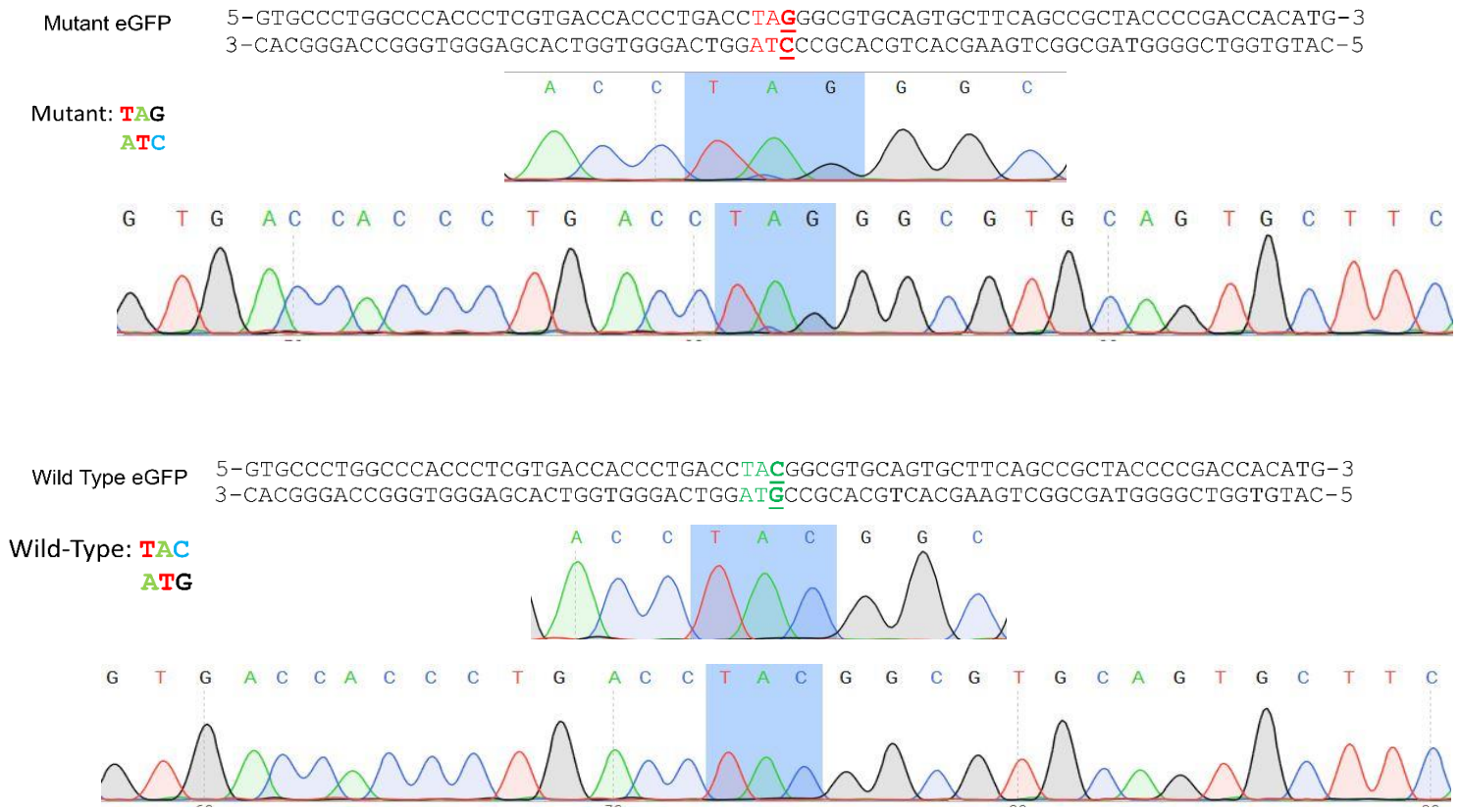


Figure 17. Sequence confirmation of ssODN/CRISPR edited cells.
Unsynchronized HCT116-19 cells were electroporated under the following conditions; 2µg CRISPR 2C and 1.35µg 72NT at $5x10^5$ cells/100µl. Cells were then sorted for GFP+ at 72 hours post electroporation. Immediately following cell sorting, DNA was isolated and the region surrounding the target base was amplified via PCR. Samples were submitted to Genewiz (South Plainfield, NJ) for sequencing analysis.

53

Chapter 3

ANALYSES OF POINT MUTATION REPAIR AND ALLELIC HETEROGENEITY GENERATED BY CRISPR/CAS9 AND SINGLE-STRANDED DNA OLIGONUCLEOTIDES

Introduction

CRISPR/Cas9 executes double-stranded DNA cleavage efficiently, closure of the broken chromosomes is dynamic, as varying degrees of heterogeneity of the cleavage products appear to accompany the emergence of a corrected base. Therefore, it has been of interest to analyze allelic variance surrounding the target site. In this study, the objective was to examine how successful single base gene editing correlates with CRISPR/Cas9 cleavage activity in a reaction where the goal was to correct a single base mutation in a well-established model gene editing system where both genotypic and phenotypic readout have been validated [101]. The product of CRISPR/Cas9 activity is often a heterogeneous population of chromosomal ends created through DNA resection promoted by non-homologous end joining (NHEJ), therefore the relationship between point mutation repair and DNA cleavage activity needs to be defined. DNA cleavage assays have often used as a benchmark to measure the level of successful gene editing in mammalian cells, especially regarding genetic knockout. These correlations are likely to be valid for studies where the goal is to disable the gene. In this study the hypothesis that gene editing of a single base mutation could be accompanied by variable DNA changes at the target site is tested by utilizing ssODNs to coordinately execute gene editing while reducing the

heterogeneity of chromosomal ends generated by CRISPR/Cas9 cleavage. Still more studies are needed to understand the heterogeneous mix of sequence alterations surrounding the target site as a result of the combinatorial action of oligonucleotide and CRISPR/Cas9.

Methods and materials

Cell Line and Culture Conditions

HCT116 cells were acquired from ATCC (American Type Cell Culture, Manassas, VA). The HCT116-19 was created by integrating a pEGFP-N3 vector (Clontech, Palo Alto, CA) containing a mutated eGFP gene. The mutated eGFP gene has a nonsense mutation at position +67 resulting in a nonfunctional eGFP protein[85]. For these experiments, HCT116 (-19) cells were cultured in McCoy's 5A Modified medium (Thermo Scientific, Pittsburgh, PA) supplemented with 10% fetal bovine serum, 2mM L-Glutamine, and 1% Penicillin/Streptomycin. Cells were maintained at 37°C and 5% CO_2. The eGFP targeting custom designed 72-mer oligonucleotide was synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

CRISPR Design and Construction

The guide RNA and CRISPR/Cas9 for the eGFP system used for gene editing in this system was described previously[101]. CRISPR/Cas9 were constructed using standard cloning methods following the latest oligo annealing and backbone cloning protocol with single-step digestion-ligation[54]. The eGFP target gRNA was cloned into pX330 backbone vector (Addgene plasmid 42230), a human codon-optimized

SpCas9 and chimeric guide RNA expression plasmid. The pX330 was purchased through Addgene (https://www.addgene.org). Following construction, clones were verified by DNA sequencing by Genewiz Incorporated (South Plainfield, NJ).

Experimental Strategy

For experiments utilizing synchronized cells, HCT116-19 cells were seeded at 2.5×10^6 cells in a 100 mm dish and synchronized with 6 μM aphidicolin for 24 hours before targeting. Cells were released for 4 hours before transfection by washing with PBS (− /−) and adding complete growth media. Synchronized and unsynchronized HCT116-19 cells were simultaneously transfected at a concentration of 5×10^5 cells/100 μL in 4 mm gap cuvette (BioExpress, Kaysville, UT). 0.6 μM of single-stranded oligonucleotide and/or the appropriate dose of pX330 constructs were electroporated (250 V, LV, 13 ms pulse length, 2 pulses, 1 s interval) using a Bio-Rad Gene Pulser XCellTM Electroporation System (Bio-Rad Laboratories, Hercules, CA). Cells were then recovered in 6-well plates with complete growth media at 37 °C for 24–48 hours before analysis.

Analysis of Gene Edited Cells

HCT116-19 cell fluorescence (eGFP+) was measured by a Guava EasyCyte 5HT Flow Cytometer (Millipore, Temecula, CA). Cells were harvested by trypsinization, washed once with 1x PBS (− /−) and resuspended in buffer (0.5% BSA, 2 mM EDTA, 2 μg/mL Propidium Iodide (PI) in PBS (− /−)). Propidium iodide was used to measure cell viability viable cells stain negative for PI (uptake). Correction efficiency was calculated as the percentage of the total live eGFP positive cells over the total live cells in each sample.

SURVEYOR Analysis of CRISPR/Cas9 Cleavage Activity

HCT116-19 cells were electroporated at a concentration of 5×10^5 cells/100µl in 4 mm gap cuvette (BioExpress, Kaysville, UT) with pX330 or pX330 and 1.35µg of the 72 mer ssODN. Cells were then recovered in 6-well plates with complete growth media at 37 °C for 72 hours. DNA was isolated using the Blood and Tissue DNeasy kit (Qiagen, Hilden, Germany). The Surveyor assay was performed on 605 bp amplicons that were created using forward primer, 5′ CTGGACGGCGACGTAAACGGC and reverse primer, 5′ ACCATGTGATCGCGCTTCTCG. PCR samples were purified using the QIAquick PCR purification kit (Qiagen, Hilden, Germany). 200 ng of each PCR product was mixed with 200 ng of PCR product from the untreated sample and subjected to a heteroduplex formation: 95 °C for 10 minutes, 95 °C to 85 °C with a ramp rate of − 2 °C/s, 85 °C for 1 minute to 75 °C at − 0.1 °C/s, 75 °C for 1 minute to 65 °C at − 0.1 °C/s, 65 °C for 1 minute to 55 °C at − 0.1 °C/s, 55 °C for 1 minute to 45 °C at − 0.1 °C/s, 45 °C for 1 minute to 35 °C to 25 °C at − 0.1 °C/s, 25 °C for 1 minute. After duplex formation products were treated with SURVEYOR Nuclease S and SURVEYOR Enhancer S (IDT Technologies) for 30 minutes at 42 °C, gel electrophoresed and stained with SYBR Safe DNA stain (Life Technologies). Gels were imaged with a Gel Doc EZ Imager (Bio-Rad) and densitometry was performed by measuring the area under the curves of each band, using the Image Lab software (Bio-Rad). Calculations were based on the following formulas:

$$\boldsymbol{Indel\ \%} = \mathbf{100} \times \left(\mathbf{1} - \left(\mathbf{1} - \frac{\boldsymbol{b} + \boldsymbol{c}}{\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}} \right)^{\frac{\mathbf{1}}{\mathbf{2}}} \right)$$

$$a = Intensity\ of\ a\ the\ undigested\ product\ (parent)$$

$$b + c = Intensity\ of\ cleveage\ products$$

Cell Cycle Analysis

For experiments utilizing synchronized cells, HCT116-19 cells were seeded at 2.5 x 10^6 cells in a 100mm dish and synchronized with 6μM aphidicolin for 24 hours. Cells were released for 4 hours prior to trypsinization and transfection by washing with PBS (-/-) and adding complete growth media. Synchronized and unsynchronized HCT116-19 cells were simultaneously transfected at a concentration of 5 x 10^5 cells/100μl in 4mm gap cuvette (BioExpress, Kaysville, UT). Single-stranded oligonucleotide and/or CRISPR plasmid constructs were electroporated at the indicated concentration (250V, LV, 13ms pulse length, 2 pulses, 1s interval) using a Bio- Rad Gene Pulser XCell Electroporation System (Bio-Rad Laboratories, Hercules, CA). Cells were then recovered in 6-well plates with complete growth media at 37°C for 24 hours. For cell cycle analysis, cells were harvested and washed in PBS. They were then spun at 450g for 5 minutes, re-suspended in 500μl cold PBS and fixed by adding 5ml of 70% cold ethanol while vertexing followed by an overnight incubation at 4°C. After fixation, the cells were centrifuged for 5 minutes at 450g, washed once with PBS and then re-suspended in 200μl of PBS and 200 μl of pre-warmed Guava ® Cell Cycle Reagent. The cells were incubated away from light at 37°C for 30min followed by analysis through flow cytometry. DNA content was analyzed by a Guava EasyCyte 5HT Flow Cytometer (Millipore, Temecula, CA) using the Guava EasyCyte Cell Cycle Software Module. Cell cycle modeling was performed and the percent of cells in G0–G1, S and G2-M phase was calculated using Modfit (Verity Software House, Topsham, ME.). Cell cycle modeling was performed using the auto analysis feature of the Modfit software.

Results

What are the on-target effects of gene editing with CRISPR/ Cas9 and ssODNs?

The variability of DNA changes at the target site as an outcome of gene editing of a single base mutation was tested by utilizing ssODNs to coordinately execute gene editing while reducing the heterogeneity of chromosomal ends generated by CRISPR/Cas9 cleavage.

Repair of a single point mutation in mutant eGFP gene has been executed by the combined action of a specifically designed ssODN and the appropriate CRISPR/Cas9 system in cells that are either unsynchronized or synchronized and released [4,101–103] (as shown in Figure 18A). Figure 18 B illustrates the alignment of the ssODN with the mutant eGFP gene. The gene editing activity was evaluated by the combined action of the ssODN, and a CRISPR/Cas9 system designed for the same target[101,102] and CRISPR/Cas9 cleavage activity was concurrently measured using the well-known Surveyor endonuclease assay[104]. Since this is a reporter gene, off-target effects are minimized, yet in Table 5 the top ten predicted off target sites are presented. The 72 base ssODN (72-mer) and the CRISPR/Cas9 expression construct were introduced by electroporation and the corrected eGFP expression measured 48 hours later by FACS. The amount of ssODN was fixed and the level of CRISPR/Cas9 expression construct was increased in a stepwise fashion.

Figure 19 illustrates the level of gene editing activity obtained from a population of cells either treated in an unsynchronized state or synchronized and released before the addition of the ssODN and CRISPR/Cas9. Gene editing activity is dose-dependent, exhibiting higher levels when synchronized and released cells are targeted as compared to cells targeted in the unsynchronized state. Consistent with

previous findings, synchronization and release increases the percentage of cells transiting S phase during DNA addition, a reaction condition shown to increase the frequency of gene editing [6,38,40,92,103]. Activity is predictably reduced when higher levels of the expression constructs are added due to extensive DNA cleavage activity and cellular toxicity [37,101]. The optimal length of ssODN molecules for gene editing in the eGFP system was confirmed; ranging within 50–100 bases [40,103] (see inset in Figure 19).

Next, the Surveyor endonuclease was used to determine the DNA cleavage activity generated through the action of the CRISPR/Cas9 system at dosages that support significant levels of gene editing. The data are presented in Figure 20 and Figure 21. The Surveyor assay detected CRISPR/Cas9 activity in the absence of the ssODN; a predictable rise in activity was observed as a function of the amount of expression vector present in the reaction. When cleavage activity is measured in reaction mixtures with the ssODN, the Surveyor endonuclease activity is increased (orange line vs yellow line), particularly at the two doses where gene editing activity is near maximal. While this may appear to be counterintuitive, the Surveyor endonuclease assay prefers DNA duplexes with small but definitive indels. DNA duplexes bear small insertions, deletions or single base changes which are more appropriate as substrates for cleavage [104]. This may suggest a significant degree of heterogeneity at the target site in conjunction with point mutation repair or perhaps independent of it. A similar observation was made by Schumann et al. using a T7 Endonuclease I (T7E1) recognition assay to measure CRISPR/Cas9 cleavage activity [105]. In that system, however, the objective was different; to insert a segment of DNA. The correlation between the extent of single base gene editing and

heterogeneity detected by the Surveyor assay prompted us to examine the DNA sequence both at and surrounding the target site, as a function of point mutation repair. This led to the testable hypothesis that point mutation repair could come with various degrees of onsite mutagenesis, resulting in a population of heterogeneous re-ligated DNA ends, the resection products.

The Kmiec lab and others [8,91,92,106,107] have reported previously that the process of single base gene editing in mammalian cells leads to small but detectable slowing of the progression of targeted cells through S phase, perhaps due to the activation of the DNA damage response pathway. Interestingly this collateral effect enhances gene editing activity as the longer period of time spent in S phase coordinately extends the open conformation of the chromatin thereby increasing accessibility of the target for the oligonucleotide; otherwise the cells exhibit a standard cell cycle profile. The profile of the targeted cell population under conditions that exhibit both significant levels of gene editing activity as well as CRISPR/Cas9 cleavage activity was analyzed next. Interest was centered on S phase events to determine if there would be an extension of S phase induced by the presence of CRISPR/Cas9 at increasing concentrations of plasmid. Analysis was given to the three specific parameters associated with S phase: diploid (%), percent of all events that are associated with a single cycling population; diploid: S phase (%), percent of all cells in the diploid cycle; and Total S Phase (%) (Average S Phase), the sum of all S Phase areas as a percentage of the total area for cycling cells of all populations. Debris (%) was also analyzed to determine the quality of the analyzed data. As seen in Table 6 (see also Figure 22), no significant damage to cell cycle progression is observed but an increase in the number of cells progressing through S phase is seen when all of the

reaction components are present. Some increase in the amount of debris generated under these conditions is also apparent which may be due chromosomal degradation as a function of repetitive cleavage activity.

Discussion

This study investigated the relationship between DNA cleavage activities, as measured by two established assays, and a gene editing reaction wherein an ssODN and CRISPR/Cas9 jointly execute single base repair. This reaction differs from the broadly used approach where single-stranded DNA is used as a template for DNA insertion. The presence of the ssODN enhances the activity of the Surveyor endonuclease assay by increasing the size of the population of suitable cleavage substrates but reduces the activity of AvrII in the RFLP assay by catalyzing single base exchange through the destruction of the restriction site. These data suggest that positive (or negative) correlations between DNA cleavage and gene editing should be made cautiously, since these relationships are dependent on the type of genetic or biochemical readout used to measure respective activities. The data are in general agreement with observations reported by Schumann *et al.* [105] although the mechanism by which gene editing takes place is likely quite different since the approach of the presented study involved genetic surgery [3,86,108–110] and not DNA addition through insertion. This difference is most apparent in the restriction on the polarity of the oligonucleotide used to direct single base exchange and gene repair. While other strategies are being tested to improve the efficiency of precise genome editing by inhibiting NHEJ [111], our data suggest that ssODN themselves might also be useful in reducing the heterogeneity of DNA ends created through the activity of RNA-guided engineered nucleases and nonhomologous end joining. Taken together,

these results indicate that single point mutation repair, catalyzed by the combination of a single-stranded oligonucleotide and a specific CRISPR/Cas9, can generate a population of cells with corrected alleles and a heterogeneous mix of sequence alterations at or surrounding the target site. These results could also suggest that onsite mutagenesis in gene editing systems where point mutation repair is the objective, could affect the translation of the technology into a more relevant clinical setting, perhaps for Sickle Cell Disease. The importance of these data will affect gene editing studies with a central focus of identifying and analyzing converted cells but not evaluating the population of cells not displaying the desired phenotype; the uncorrected population. Deep sequencing of offsite mutagenesis remains a central focus of most analytical approaches for gene editing, but our data suggest that onsite mutagenesis may be significant. The repair of a point mutation can be facilitated by combined activity of a single-stranded oligonucleotide and a CRISPR/Cas9 system. While the mechanism of action of combinatorial gene editing remains to be elucidated, the regulatory circuitry of nucleotide exchange executed by oligonucleotides alone has been largely defined. The appropriate CRISPR/Cas9 system leads to an enhancement in the frequency of gene editing directed by single-stranded DNA oligonucleotides. Still more studies need to be put in place to further understand the heterogeneous mix of sequence alterations at or surrounding the target site as a result of the combinatorial use of oligonucleotide and a specific CRISPR/Cas9 to correct a point base mutation.

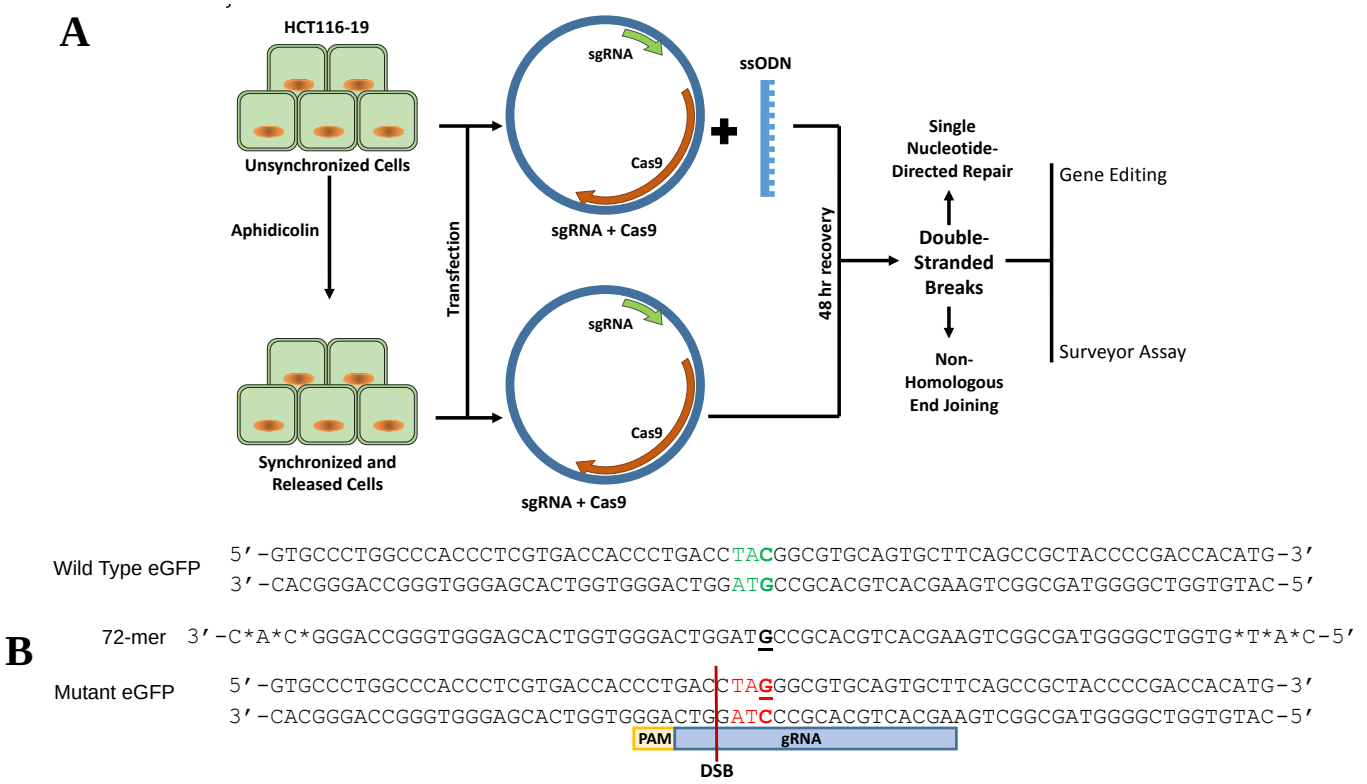


Figure 18. Experimental Design for HCT116-19 Gene Editing.
(A) Experimental workflow. HCT116-19 cells are either unsynchronized or synchronized and released, then transfected with a CRISPR/Cas9 expression vector (pX330) with or without ssODN, then after 48 hours analyzed for gene editing activity and Surveyor endonuclease digestion. **(B) Gene editing model system and ssODNs.** The wild-type and mutated eGFP gene segments with the target codon located in the center of the sequences are displayed in green and red, respectively. The nucleotide targeted for exchange is emphasized in bold and underlined. The gRNA and protospacer adjacent motif (PAM) shown indicate the CRISPR/Cas9 target site and the location of the resulting double-stranded break (DSB). The phosphorothioate modified, end protected 72-mer which is used to target the mutated eGFP gene is shown.

Table 5. Top 10 Potential Predicted Off-target sites of the mutant eGFP Targeting gRNA in the Human Genome

Mutant eGFP targeting gRNA - AGCACTGCACGCCCTAGGTCAGG				
Top 10 potential genome-wide off-target sites				
Sequence	Score	Mismatches	UCSC gene	Locus
AGCACTGCCCGCCCTAGGCCAGG	1.7	2MMs [9:19]	NM_001199642	chr3:+123071187
AGCTCTGCAGGCCCTAGGTGGAG	1.3	3MMs [4:10:20]		chr11:-795678
AGAGCTGCCTGCCCTAGGTCTAG	0.8	4MMs [3:4:9:10]		chr1:-6697733
AGCTCTGCACTACCTAGGTCAAG	0.7	3MMs [4:11:12]		chr5:+141485541
AACAGTGCATGCCCTAGGTACAG	0.6	4MMs [2:5:10:20]		chr1:+89214834
TGCACTGCAAGCCCTCGGTCAAG	0.5	3MMs [1:10:16]	NM_030576	chr17:-61773589
AGAAATGCCCTCCCTAGGTCCAG	0.5	4MMs [3:5:9:11]	NM_207404	chr3:+42958472
AGCCCTGGATGCCCTAGGCCAAG	0.5	4MMs [4:8:10:19]		chr4:-4374838
AGCCCTGCACGCCCTAGGGAAAG	0.4	3MMs [4:19:20]	NM_139027	chr9:-136323253
AGCCCTGCCTGCCCTAGGTGGAG	0.4	4MMs [4:9:10:20]		chr9:-137677669

The top 10 potential off-target sites are listed with the off-target sequence, off-target hit scores, the number of mismatches and their location in the seed sequence, the UCSC gene IDs (if off-target occurs in a gene), and the genomic locus of the off-target site.

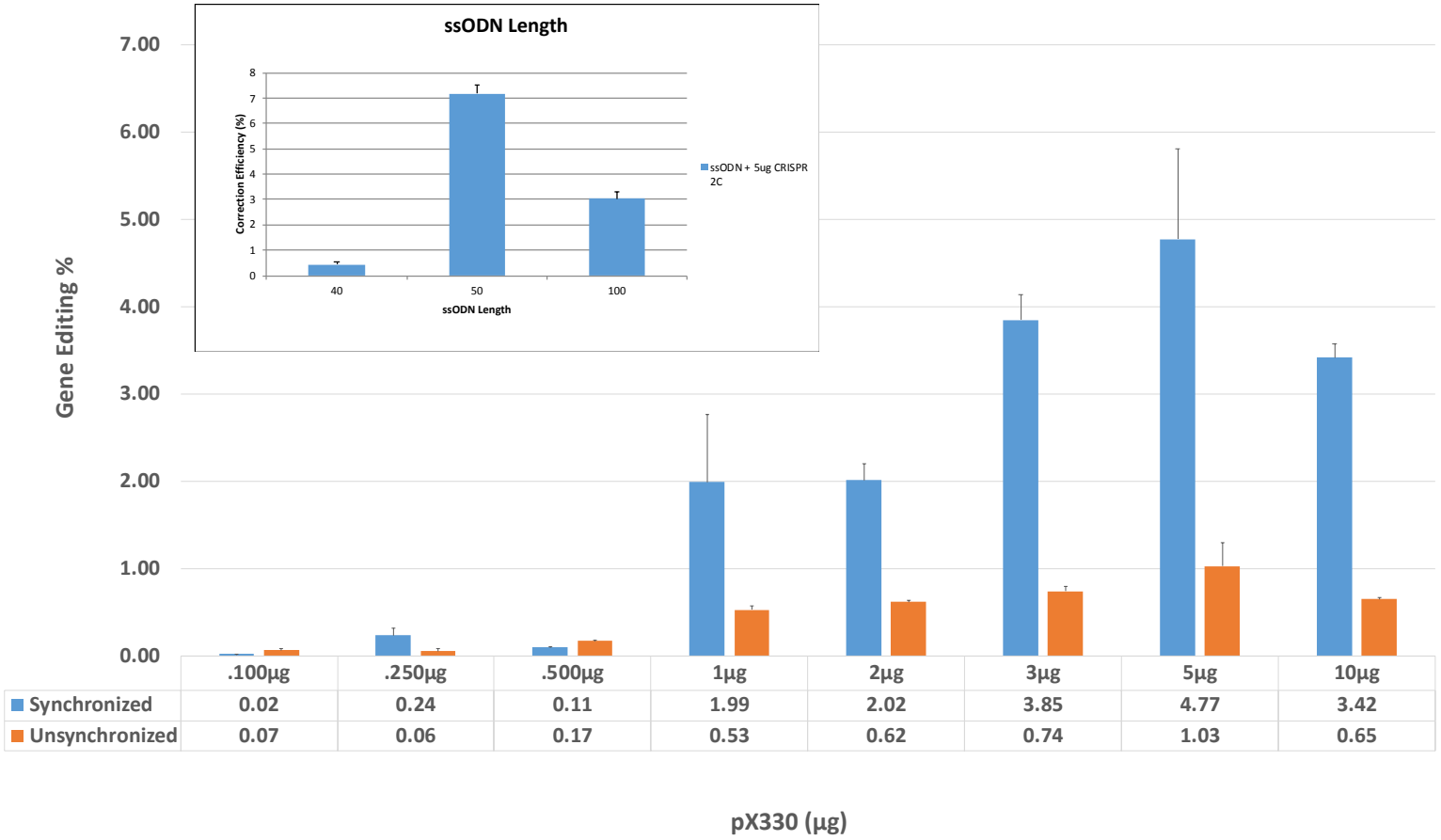
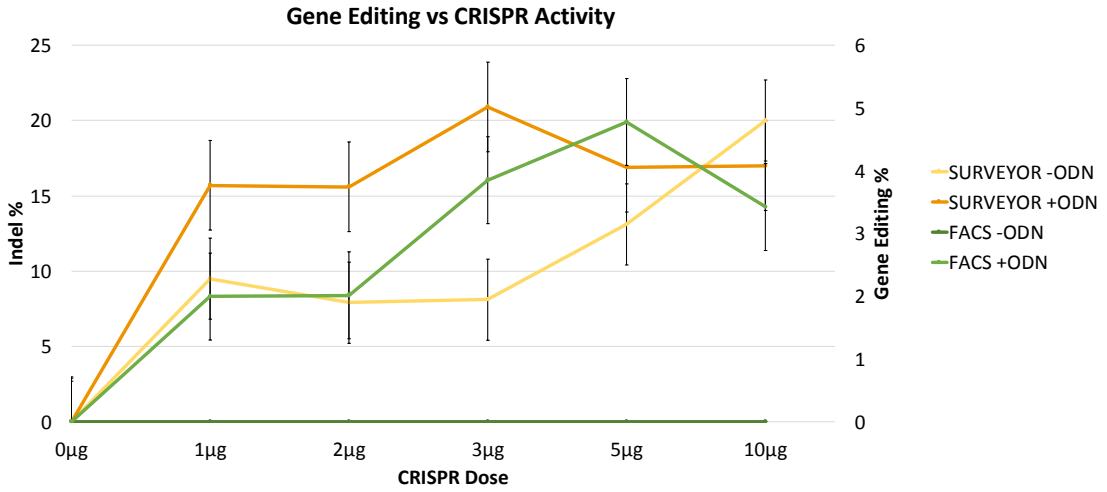


Figure 19. Gene editing dose curve using in synchronized and unsynchronized cells.
Synchronized (blue) and unsynchronized (orange) HCT116-19 cells were electroporated with 0.1–10.0 µg of pX330 and 1.35 µg of 72mer. After a 48-hour recovery period, gene editing activity was measured using a Guava Easy Cyte 5HT Flow Cytometer. Gene editing is displayed as correction efficiency (%), determined by the number of viable eGFP positive cells

divided by the total number of viable cells in the population. Each treatment was performed in triplicate and standard error is illustrated with accompanying bars. Statistical analysis was performed using two sample unequal variance students T-test distribution to compare the value of correction efficiency between synchronized and un-synchronized cells when treated with CRISPR/Cas9. *$p < 0.05$. Inset. Correction efficiency using varying lengths of ssODN at equimolar concentrations with 5 ug of pX330, 48 hours after electroporation.



CRISPR Dose	SURVEYOR -ODN	SURVEYOR +ODN	FACS -ODN	FACS +ODN	FACS Error
0µg	0	0	0	0	0
1µg	9.5	15.7	0.0	2.0	0.77
2µg	7.9	15.6	0.0	2.0	0.19
3µg	8.1	20.9	0.0	3.8	0.29
5µg	13.1	16.9	0.0	4.8	1.04
10µg	20.0	17	0.0	3.4	0.15

Figure 20. Correlation between CRISPR/Cas9 cleavage and gene editing activity.
Synchronized and released HCT116-19 cells were electroporated with 0–10µg of pX330 and with (+ ODN) or without (− ODN) 1.35µg of 72-mer. CRISPR/Cas9 cleavage activity measured by Surveyor endonuclease assay (orange and yellow) as well as gene editing (dark green and light green) activity measured by FACS are shown. Standard error is represented by the bars on each data point.

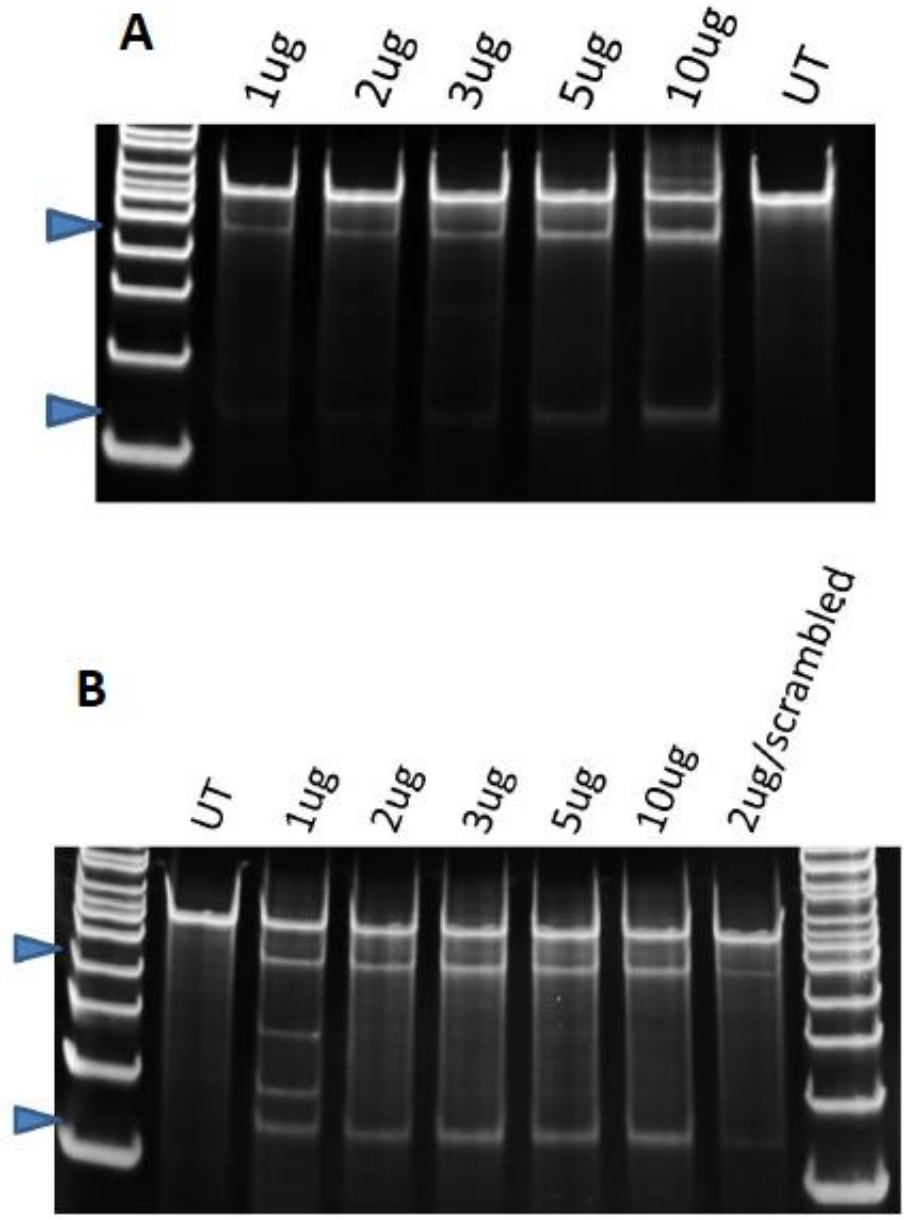


Figure 21. Surveyor Analysis.
(A) Surveyor analysis of synchronized HCT116-19 cells that were transfected with increasing concentrations of pX330 containing the eGFP targeting gRNA. Cleavage products are indicated by the small arrows. (B) Surveyor analysis of synchronized HCT116-19 cells that were transfected with increasing concentrations of pX330 containing the eGFP targeting gRNA and 1.35µg of the 72-mer ssODN or 2µg of pX330 with a 72-mer scrambled oligo. Cleavage products are indicated by the small arrows.

Table 6. Cell Cycle Analysis

Cell Cycle Analysis				
Treatment	**Diploid (%)**	**Diploid: S phase (%)**	**Total S phase (%)**	**Debris (%)**
Unsynchronized	100	41.59	41.59	17.24
Synchronized Un-treated	90.44	32.52	29.41	11.66
72-mer ssODN	98.09	29.85	29.28	12.18
72-mer ssODN/ CRISPR-Cas9	97.90	35.80	35.04	19.38

HCT116-19 cells were synchronized and released in culturing medium. Synchronized and unsynchronized HCT116-19 cells were transfected with 72-mer and CRISPR/Cas9 plasmid construct px330. Cells were allowed to recover in complete growth media for 24hr. Cell cycle profiles represent the DNA content distributions of the cells at 24hr post transfection. Cell cycle modeling was performed using the Auto analysis feature of the Modfit software. S phase events were recorded for three specific parameters associated with s-phase. The diploid (%), percent of all events that are associated with a single cycling population. Diploid: S phase (%), percent of all cells in the diploid cycle. And, Total S Phase (%) (Average S Phase), the sum of all S Phase areas as a percentage of the total area. The Debris (%) was also analyzed to determine the quality of the analyzed data.

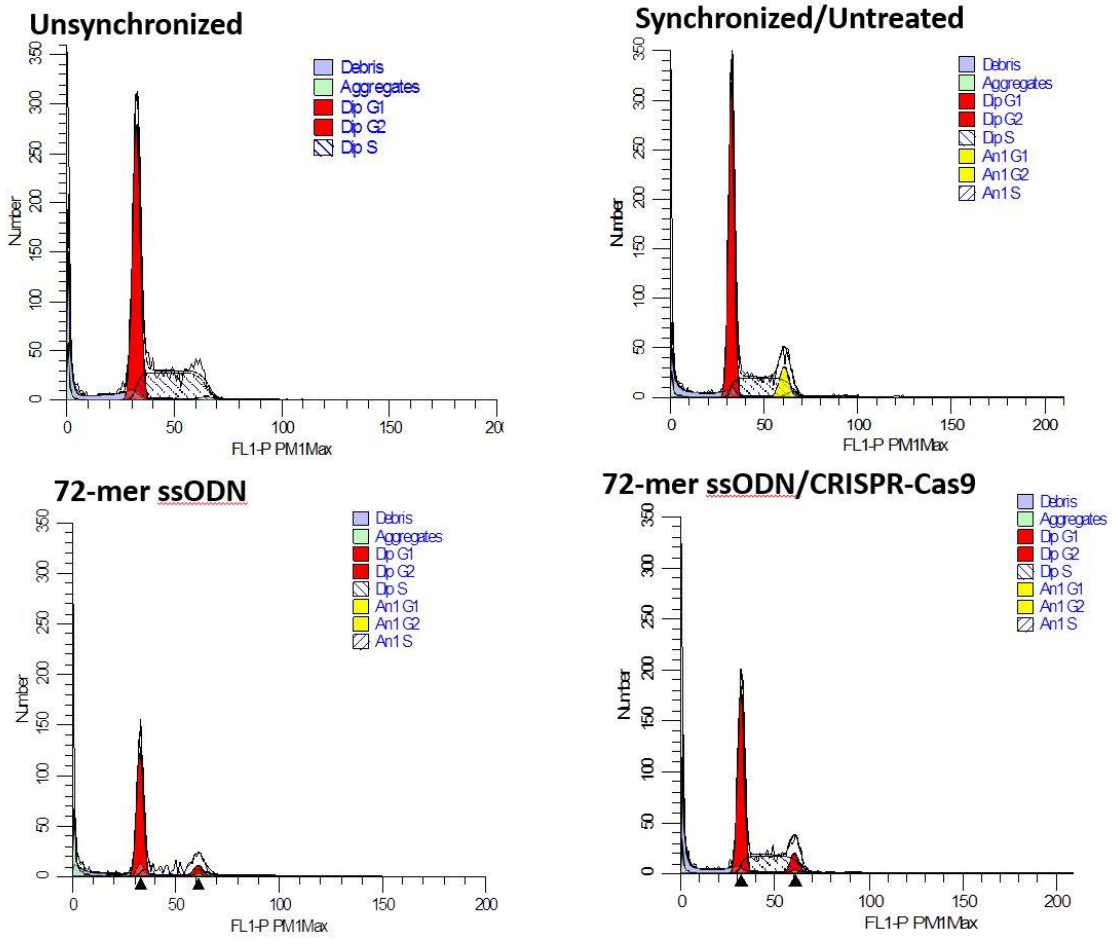


Figure 22. Cell Cycle profiles of gene editing reactions.
HCT116-19 cells were synchronized and released in culture medium. Synchronized and unsynchronized HCT116-19 cells were transfected with 72-mer and CRISPR/Cas9 plasmid construct px330. Cells were allowed to recover in complete growth media for 24hr. Cell Cycle profiles represent the DNA content distributions of the cells at 24hr post transfection with modeling performed using the Auto analysis feature of the Modfit software. S phase extension was by Diploid (%), Diploid: S phase (%), Total S Phase (%) (Average S Phase), and Debris (%) was also analyzed to determine the quality of the analyzed data.

Chapter 4

INSERTIONAL MUTAGENESIS BY CRISPR/CAS9 RIBONUCLEOPROTEIN GENE EDITING IN CELLS TARGETED FOR POINT MUTATION REPAIR DIRECTED BY SHORT SINGLE-STRANDED DNA OLIGONUCLEOTIDES

Introduction

CRISPR/Cas9 and single-stranded DNA oligonucleotides (ssODNs) have been used to direct the repair of a single base mutation in human genes. This study examines a method designed to increase the precision of RNA guided genome editing in human cells by utilizing a CRISPR/Cas9 ribonucleoprotein (RNP) complex to initiate DNA cleavage. It is of great interest to better understand the outcomes of gene editing when employed by different technologies and delivery methods, i.e. RNP vs. plasmid. The RNP is assembled in vitro and induces a double stranded break at a specific site surrounding the mutant base designated for correction by the ssODN. Significant gene correction activity is promoted by the RNP and single-stranded DNA oligonucleotide validated by genotypic and phenotypic readout. This was achieved in the presence of all gene editing components. The genotype of individually sorted, corrected and uncorrected clonally expanded cell populations for the mutagenic footprint left by the action of gene editing tools was examined in detail. While the DNA sequence of the corrected population is exact with no adjacent sequence modification, the uncorrected population exhibits heterogeneous mutagenicity with a wide variety of deletions and insertions surrounding the target site. This type of DNA aberration was designated as on-site mutagenicity. Analyses of two clonal populations

bearing specific DNA insertions surrounding the target site, indicate that point mutation repair has occurred at the level of the gene. The phenotype, however, is not rescued because a section of the single-stranded oligonucleotide has been inserted altering the reading frame and generating truncated proteins. These data illustrate the importance of analyzing mutagenicity in uncorrected cells. These results also form the basis of a simple model and standardized methodology for point mutation repair directed by a short single-stranded DNA oligonucleotide and CRISPR/Cas9 ribonucleoprotein complex.

Methods and materials

Cell Line and Culture Conditions

HCT116 cells were acquired from ATCC (American Type Cell Culture, Manassas, VA). The HCT116-19 was created by integrating a pEGFP-N3 vector (Clontech, Palo Alto, CA) containing a mutated eGFP gene. The mutated eGFP gene has a nonsense mutation at position +67 resulting in a nonfunctional eGFP protein[85]. For these experiments, HCT116 (-19) cells were cultured in McCoy's 5A Modified medium (Thermo Scientific, Pittsburgh, PA) supplemented with 10% fetal bovine serum, 2mM L-Glutamine, and 1% Penicillin/Streptomycin. Cells were maintained at 37°C and 5% CO_2. The eGFP targeting custom designed 72-mer oligonucleotide was synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

CRISPR/Cas9 RNP Design and Complexing

The mutant eGFP gene sequence was entered into the Zhang Lab's online generator (http:// crispr.mit.edu/) and the CRISPR guide sequences which binds with

close proximity to target (TAG = 0) was chosen. crRNA, tracrRNA and Cas9 protein were kind gifts from Integrated DNA Technologies (Coralville, Iowa) and stored and utilized according to their suggestions. RNP assembly was performed by mixing RNA oligos (crRNA and tracrRNA) in equimolar concentrations to a final duplex concentration of45μM. For the RNA to duplex the mix was heated at 95˚C for 5 minutes and allowed to cool to room temperature (15–25˚C). For each sample crRNA:tracrRNA (45μM working solution) and Cas9 protein (60μM stock solution) were diluted in their respective buffers to a final volume of 5μL each to achieve the desired treatment concentration. Prior to mixing with cells crRNA:tracrRNA duplex and Cas9 protein we mixed and set to incubate at room temperature for 15 minutes. The same annealing conditions and reactions were carried out in the assembly of the mutant eGFP or B-globin gene crRNA RNP.

Experimental Strategy

For all experiments, HCT 116–19 cells were synchronized for 24 hours with Aphidicholin at the G1/S border prior to introducing the Cas9 ribonucleoprotein (RNP) complex or CRISPR/ Cas9 generated from an expression construct. The CRISPR expression plasmid was constructed using standard cloning methods following the latest oligo annealing and backbone cloning protocol with single-step digestion-ligation. The CRISPR guide sequences were cloned into the pX330 backbone vector (Addgene plasmid 42230), a human codon-optimized SpCas9 and chimeric guide RNA expression plasmid. Single-stranded DNA oligonucleotides used in this study are 72 base pairs in length and designed as depicted in Figure 23. RNP assembly was performed by mixing RNA oligos (crRNA and tracrRNA) in equimolar concentrations to a final duplex concentration of 45μM. For the RNA to duplex the

mix was heated at 95˚C for 5 minutes and allowed to cool to room temperature (15–25˚C). For each sample crRNA:tracrRNA (45μM working solution) and Cas9 protein (60μM stock solution) were diluted in their respective buffers to a final volume of 5μL each to achieve the desired treatment concentration (24-120pmol). Prior to mixing with cells crRNA:tracrRNA duplex and Cas9 protein we mixed and set to incubate at room temperature for 15 minutes. Electroporation transfection was performed by mixing cells at concentration of $5x10^5$ cells/100 microliters along with the RNP and ssODNs in a 4mm gap cuvette (BioExpress, Kaysville, UT) (250V, LV, 13ms pulse length, 2 pulses, 1s interval) using a Bio-Rad Gene PulserTM XCell Electroporation System (Bio-Rad Laboratories, Hercules, CA). Cells were then recovered in 6-well plates with complete growth media at 37˚C for 72 hours prior to analysis.

Analysis of Gene Edited Cells and Transfection Efficiency

HCT 116–19 cell fluorescence (eGFP+) was measured by a BD FACSAria II (BD Biosciences, San Jose, CA). Cells were harvested by trypsinization, washed once with 1x PBS (-/-) and resus- pended in buffer (0.5% BSA, 2mM EDTA, 2μg/mL Propidium Iodide in PBS -/-). Propidium iodide was used to measure cell viability as such, viable cells stain negative for PI (uptake). Correction efficiency was calculated as the percentage of the total live eGFP positive cells over the total live cells in each sample. Error bars are produced from three sets of data points generated over three separate experiments using basic calculations of Standard Error.

RNP in Vitro Activity

Cellular gDNA was isolated from pellets of 1×10^6 untreated HCT 116–19 cells using Qiagen DNAEasy Blood and Tissue Kit (Cat. ID 69506, Valencia, CA).

PCR was performed using AmpliTaq (Thermo-Scientific, Waltham, MA) on 200ng of isolated gDNA, with amplification parameters optimized for an amplicon size of 605bp with forward primer 5'-CTGGACGGCGACGTAAACGGC-3'and reverse primer, 5'-ACCATGTGATCGCGCTTCTCG-3'. Amplicon size was verified on 1% agarose gel and PCR samples were cleaned up using the QIAquick PCR purification kit (Qiagen, Hilden, Germany). After purification, 300ng of PCR sample was combined with Buffer 3.1 and 25pmols or 50pmols of RNP complex. The mix was incubated for 40 minutes at 37˚C then 1 microliter of proteinase K was added to the mix and incubated for 15 minutes. Samples were loaded along with NEB 2-log DNA ladder (NEB, Ipswich, MA) and analyzed on a 2% TBE agarose gel. Cellular gDNA was isolated from pellets of 1–2 x 10^6 K562 cells using the Qiagen DNEasy Blood and Tissue Kit (Cat. ID 69506, Valencia, CA). PCR was performed using Phusion High-Fidelity PCR Master Mix with HF Buffer (Thermo-Scientific, Waltham, MA) on isolated gDNA, with amplification parameters optimized for an amplicon size of 345bp with forward primer 5'- TCCTAAGCCAGTGCCAGAAGAG -3' and reverse Primer 5'- CTATTGGTCTCCTTAAACCT-3'. Amplicon size was verified on 1% agarose gel. PCR samples were cleaned up using the QIAquick PCR purification kit (Qiagen, Hilden, Germany) and treated with DdeI restriction enzyme (NEB, Ipswich, MA) following the manufacturer's protocol or RNP following the method described above. Digested samples were loaded along with NEB 2-log DNA ladder (NEB, Ipswich, MA) and analyzed on a 2% TBE agarose gel.

DNA Sequence Analysis

Synchronized and released HCT 116–19 cells were harvested and electroporated at a concentration of 5 x 10^5 cells/100µLwith RNP complex at

100pmols and 72NT ODN at 2.0 μM. Following electroporation, cells were placed in 6-well plates and allowed to recover for 72 hours. Cells were then individually sorted by a BD FACSAria II sorter-488nm (100mw) (BD Biosciences, San Jose, CA) for eGFP+/- into 96-well plates. Cells were expanded over 6 weeks and harvested. Cellular gDNA was isolated using Qiagen DNEasy Blood and Tissue Kit (Cat. ID 69506, Valencia, CA) and the region surrounding the target base was amplified via PCR (718bp, forward primer 5'-ATGGTGAGCAAGGGCGAGGA-3' and reverse primer 5'-ACTTGTACAGCTCGTCCATGC-3'). Samples were submitted to Eton Bio Incorporated (Union, NJ) for sequencing analysis.

Results

Can the precision of RNA guided genome editing in human cells be increased?

Gene editing using CRISPR/Cas9 and ssODNs aims to rescue a mutated eGFP by converting the G base to a C and restoring the normal tyrosine codon (TAC) (see Figure 23). For this study a specific combination of CRISPR/Cas9 and ssODN that has been shown previously to be optimal for RGEN-directed correction was used (described in Figure 23)[101]. In the experiments reported herein, the 72NT oligonucleotide was used since previous data have established that targeting the non-transcribed strand at this ssODN length leads to a higher level of gene editing[44]. In addition, the use of the complementary oligonucleotide, targeting the transcribed strand, leads to artefactual annealing to the sgRNA component of the CRISPR/Cas9 complex, reducing overall activity[101]. Successful correction of the point mutation leads to the production of a functional eGFP which can be detected and quantified by FACS. In previous studies, both components (sgRNA and Cas9) were generated from

a plasmid expression vector[101]. In this study, however, the CRISPR/Cas9 were used as a ribonucleoprotein complex that is preassembled prior to introduction into the cells. Figure 24A provides a schematic of the RNP assembly process. The crRNA and the tracrRNA are identical in sequence to the longer sgRNA used previously although they are used as separate RNA molecules in this protocol. The crRNA and tracrRNA are reannealed by mixing RNA oligos (crRNA and tracrRNA) in equimolar concentrations with subsequent addition of purified Cas9 protein. To measure inherent activity of the RNP with regard to its capacity to cleave DNA, an in-vitro reaction was carried out to assess the capacity of this particular RNP complex to induce double strand DNA cleavage in a specific fragment of DNA. The fragment was created by PCR amplification across the mutant eGFP target site generating a 605 base pair template containing the target site for the RNP. The preassembled RNP was mixed with this fragment at various concentrations for 40 minutes followed by deproteinization by Proteinase K. The digestion fragments were visualized after gel electrophoresis and the data are presented in Figure 24B. As predicted, the RNP efficiently catalyzes double strand DNA cleavage of the specific fragment but not of a fragment lacking the target site. These results support the notion that the RNP complex assembled under these reaction conditions contains the appropriate level of activity and specificity for inducing double strand DNA cleavage. For the cell-based gene editing reaction, the RNP was combined with 72NT at a prescribed molar ratio of 1:2.5 and immediately electroporated into HCT 116–19 cells. Cells are allowed to incubate for 72 hours after which time they are processed for FACS analysis. A dose curve was carried out with increasing concentrations of the preassembled RNP, while maintaining a constant ratio of ssODN (72NT) to RNP. The data are presented in

Figure 25A and exhibit a steady increase in correction efficiency, rising steadily from the initial level of 24 pmols of RNP to a high level when 120 pmols are used in the reaction. In contrast, single agent gene editing using only the 72NT produces a much lower level of gene editing, as previously reported (see inset Figure 25A)[3,5]. These data suggest that the RNP particle used in combination with the 72NT oligonucleotide can promote gene correction at a level approaching 10 to 12% reproducibly. The next experiment addresses the question of the importance of each reaction component. A complete reaction mixture containing 100pmol of RNP complex and 2.0μM of 72NT respectively was utilized. As shown in Figure 25B, elimination of one or two of the essential reaction components eliminates gene editing activity completely. In addition, a complete reaction mixture for activity after replacing the specific RNP complex with one that targets the beta globin gene was also tested. No reproducible levels of gene editing were observed emphasizing the requirement for the specific RNP particle coupled to the single-stranded oligonucleotide to direct correction of the point mutation in the eGFP gene. The reaction mixture containing all of the relevant components, however, promotes correction efficiency of approximately 10%, consistent with the previous data (Figure 25A). This level of gene editing directed by the RNP/72NT complex can therefore produce a sufficient level of eGFP positive cells separated from eGFP negative cells to enable robust single cell sorting by FACS and subsequent clonal expansion. Figure 26A presents a side scatter plot of a complete gene editing reaction on a population of HCT 116–19 cells. The segmented plot illustrates a distinct percentage of cells in the P2 quadrant, representing eGFP positive cells that can be distributed as individual cells into a single well of a 96 well plate. In a similar fashion, uncorrected cells from the population displayed in quadrant P3 can

also be isolated. Figure 26B displays the experimental flow following transfection and sorting, enrichment and finally clonal expansion prior to DNA harvesting, extraction and sequence analysis. Using this experimental strategy, it was possible to interrogate the allelic composition of corrected and uncorrected single cells specifically measuring the degree of DNA heterogeneity, or on-site mutagenesis, accompanying successful and unsuccessful gene editing activity. After sorting, isolation and expansion of corrected and uncorrected single cells, the DNA sequence of multiple clones was analyzed using direct Sanger sequencing following PCR amplification of a 718bp long PCR fragment. As shown in Figure 27A, precise conversion of the TAG codon, to TAC, (light blue highlighted area) confirms phenotypic expression in the eGFP positive clones, at the DNA level. Sixteen eGFP positive clones were expanded in the same fashion and all contained the converted DNA sequence as presented in Figure 27A. No sequence alterations or onsite mutagenesis was observed within the 718 base pair DNA region in these experiments. The DNA sequence readout found in all the clones is provided in the lower half of the figure to show that no contaminating or background sequence is present, indicating that single cell clonal expansion from the corrected population was successful. Figure 27B illustrates the genotypic analyses of 17 eGFP-negative clonally expanded cells, selected at random from the sorted, uncorrected population. In eight of the seventeen clones, the TAG codon remains intact and no sequence variation is observed within the region examined. In the other nine clone's insertion immediately downstream from the targeted base are present. These results demonstrate that onsite mutagenesis occurs during RNP/ssODN gene editing reactions in cells that fail to achieve the desired phenotype. This led to the examination of the genotype of the two clones containing DNA insertions, neither of

which exhibit a change in phenotype as judged by the absence of green fluorescence. These two clones are instructive not only for what they tell us about the potential for on-site mutagenesis but also what they tell us about a mechanism for DNA insertion driven by single-stranded DNA. In Figure 27C, the DNA sequence of the uncorrected clone containing a 15 base pair insertion is presented. Here it is illustrated how this insertion created a new frameshift generating the corrected TAC tyrosine codon and concomitantly creating a new TAG, stop codon. Thus, this clone appears to have been corrected at the targeted base but that correction is not reflected in a phenotypic change. As such, the insertion expands the gene by five codons. In the same fashion, a second clone containing a 24 base pair segment, inserted at the identical position, is also displayed in Figure 27C. In this case, 8 new codons, preceding the newly created stop codon, have now become inserted. Thus, CRISPR/Cas9 single-stranded oligonucleotide gene editing can generate a novel stretch of amino acids that are not encoded by the targeted gene. The DNA insertion matches, in perfect register, a section of the single-stranded oligonucleotide when placed in this reading frame. Thus, these data suggest that sections of the oligonucleotide can be inserted into the target gene, resulting with the simultaneous correction of the examined on-site onsite mutagenesis consisting of both deletion and insertion mutations of varying lengths. This mutagenic activity ranged from a one base pair deletion surrounding the target site to a 19 pair deletion to a 24 base pair point mutation and the generation of a mutagenic footprint (see Figure 28). The degree of mutagenesis observed in the uncorrected population is broad, signaling the importance of analyzing a representative sample of the entire population of cells targeted for genetic alteration.

Discussion

Collateral mutagenesis, generated by the action the CRISPR/Cas9 gene editing tool, has been a central focus of both advocates and critics of this technology. Sophisticated molecular cloning approaches to refute, diminish or downplay the degree of off-site mutagenesis have been offered by many of the leading laboratories in the field[13,112,113]. But, in many cases, results rely in large part on proving a negative. In fact, it is debatable as to whether or not off-site mutagenesis can be completely eliminated as a potential side reaction in therapeutic gene editing. More recently, focus has been placed on the potential of onsite mutagenesis, an outcome of the normal activity of RGENs. The inherent response of a cell to repair the double strand break through the process of non-homologous end joining is at the core of the current genetic revolution, partially inspired by RGENs, that have made the generation of gene knockouts in many eukaryotic cell types a routine lab procedure. In contrast, onsite mutagenesis becomes more relevant when the objective of the gene editing protocol is not to disable but rather to repair a gene bearing a point mutation, and eventually, to direct seamless insertion of a fragment of donor DNA. Most of the studies focused on onsite genetic heterogeneity, examine genes for which there is complex readout, as in the case of human stem cells, often requiring drug selection to identify the targeted cells[114]. Primary cells and even transformed cells can respond negatively to selective pressures when challenged with either chemo-toxins or antibiotics. The Kmiec laboratory examined onsite mutagenesis as a function of single base gene editing, the repair of a point mutation in human cells[115]. While a native gene, human HBB, has been utilized here to validate these initial findings, it is essential to employ a reproducible, robust model system that has a long history of validated readout that can measure phenotype, protein function and genotype without

exogenous manipulation. In this study the model system in which a mutated eGFP gene, integrated as a single copy into HCT 116 cells[85], is targeted for repair by the combination of a CRISPR/Cas9 RNP and a specific single-stranded DNA oligonucleotide. Successful conversion of the point mutation transforms a stop codon to a tyrosine codon enabling translation and expression of functional eGFP. Because the cells can be cloned and examined as uniquely expanded populations, allelic analysis of gene editing activity in both corrected and uncorrected populations is simple. Here it is revealed how the combination of the RNP complex and a 72-mer directs gene repair of the point mutation in an efficient and reproducible fashion. Keeping the molecular ratio of the RNP and the single-strand DNA oligonucleotide constant but raising the total amount in the reaction induces a dose-dependent response which begins to plateau above 10%; at an 8 to 10-fold higher level than when the ssODN is used as the sole gene editing agent. All of the appropriate reaction components are required for successful point mutation repair and the separation of the corrected and uncorrected cells can be achieved in a straightforward fashion. The rationale for using the RNP is that the active components of the CRISPR/Cas9 system will be delivered to the nucleus at approximately the same time facilitating a more constant initialization of the gene editing reaction. Previously, a plasmid expression system was used in which Cas9 is expressed from the same plasmid as the sgRNA. In the presented study, point mutation repair driven by the combination of the RNP and the ssODN is confirmed. With regard to the examination of DNA heterogeneity in corrected and uncorrected population of cells, identified by FACS as being corrected, exhibit precise single base repair at the target site. No genetic alteration was observed in the cells from the corrected population for a proximal distance of 718 bases. In

addition, there weren't any observations of nucleotide changes that would result in a conservative change in amino acid sequence still enabling expression of alternative wild type eGFP.

The clonal expansion of a population of cells that did not exhibit phenotypic correction generated a panel of genetic alterations ranging from uncorrected, yet intact, to a cell line bearing a 19 base deletion surrounding the target site to the insertion of 24 bases surrounding the target site respectively. Of the 17 clones tested, eight had no change to the mutant DNA sequence, perhaps indicating that the RNP complex had not reached the target site in those cells or had induced a double strand break which was properly and efficiently repaired with or without the aid of the single-stranded oligonucleotide.

A wide range of DNA sequence deletions were observed starting with a single base deletion and ending with a 19 base deletion, heterogeneity that surrounds the nucleotide targeted for gene repair. No other sequence alterations outside of the target site were observed, again within the proximal 718 bases. The one clone harboring a 15 base insertion, appears to have arisen through a duplication of the adjacent 15 bases located 5' to the target site, as well as a 24 base insertion that appears to have come from the same DNA source. Our results indicate that onsite mutagenesis clearly occurs in the uncorrected population of cells, exhibiting a wide range of indel formation. Interestingly, a similar phenomenon has been observed in a separate series of studies wherein in the objective was to induce a single base change in the genome of K562 cells [115]. In that system, however, it was not possible to examine the impact on phenotypic changes and thus this model system expands and confirms those studies,

demonstrating on-site heterogeneity as a function of gene editing reactions that include single-stranded DNA oligonucleotides.

The data collected in this study aligns with associated studies that examine the insertion of a longer fragment of DNA at a precise site. Merkel et al[114] recently published an elegant study in which indel formation was observed at the target site catalyzed by intact CRISPR/Cas9, as well as associated single or dual Nickases. Taken together, these results expand upon data from earlier work in which the objective was to modify the target site excessively without introducing unwarranted changes[13,105,113]; all of these studies reported site alteration.

The specific interest herein is not to develop a strategy to insert a large fragment of DNA, but rather to use a short piece of donor DNA, a short single-stranded oligonucleotide, to perform genetic surgery as a way to repair single point mutations. Over the course of the last 15 years the Kmiec lab and others have established the mechanism and regulation of short oligo induced gene editing[4,5,38,40,86,116]. More recent studies have shown that oligonucleotides of length between 49 and 72 bases respectively can direct single base repair as the sole agent of the gene editing reaction or in combination with both TALENs and CRISPR/Cas9[13,37]. Thus, while the mechanism of action of single base repair, directed by these oligonucleotides of restricted length, may differ from the mechanism by which DNA fragment or gene insertion takes place, it is now apparent that the same type of allelic analysis should be performed on at least a sample of both corrected and uncorrected targeted cells generated from both approaches. It is therefore critical that studies are conducted in reliable, robust and validated testing systems to explore the degree of collateral damage directed by CRISPR/Cas9 to more fully understand the

remarkable power of this gene editing tool especially in light of its therapeutic potential.

Homologous recombination was not observed in any events from distal sites, wherein genetic information is provided by adjacent chromosomes to aid in the repair of the fragmented DNA. The DNA sequence of the insertion clones, however, enables a continuation of the reading frame through several codons until a stop codon is generated so additional genetic information was proved in some fashion. Both insertion clones contain the exact corrected point mutation but do not score as eGFP+ because the inserted DNA creates a stop codon 15 or 24 bases downstream from the targeted nucleotide respectively. This is an interesting example of how double strand DNA breakage can provide a site for DNA insertion of exogenous or repetitive segments as the cell responds to chromosomal damage. These data provide insight into the overall mechanism by which short oligonucleotides and the RNP execute the repair of a point mutation in a mammalian cell and enable the development of a model that explains our results. Figure 28 displays a model that is believed to explain the generation of cells bearing only a corrected genotype, as well as cells bearing both corrected genotype and phenotype.

A number of sophisticated models have been put forward to explain the insertion of exogenous DNA templates for the repair of single base mutations in gene editing reactions[99,115,117]. These models are based on a process known as Homology Directed Repair (HDR) and likely help explain the results in many of the studies wherein the objective was to insert longer pieces of DNA. In their important study, Paquet et al[118] developed a gene editing methodology, known as CORRECT, for introducing mono- and bi-allelic sequence changes. These workers were successful

in elevating the accuracy of HDR through the incorporation of blocking mutations that modify the interaction between CRISPR/Cas9 and the PAM sites resulting in scar-less genome editing. As a group, these approaches and the models that are generated by them are somewhat complex because they involve specific enzymatic activities and sophisticated reengineering of some of the reaction components.

The model presented here is much simpler as it has been based on the well-accepted and standard model of Double Strand Break Repair[118–120]. When a double strand break occurs in a mammalian chromosome (in the case of gene editing, induced by CRISPR/Cas9 activity), activated exonucleases recognize the break and resect the broken ends to varying degrees, a biochemical reaction that takes place regardless of whether the break is designated for repair through the process of homologous recombination or nonhomologous end joining. In the case of homologous recombination however, usually occurring during S-phase of the cell cycle[121,122], proteins involved in DNA recombinational repair load onto the broken ends. Subsequently, a sister chromatid provides the DNA template to enable the broken strand to once again be made whole through the process of gap filling by DNA replication. Since crossover of one strand of DNA from the sister chromatin provides the template, its original partner strand is displaced and becomes the template for gap filling through DNA replication, albeit in the opposite polarity. Thus, the gap created by the original double strand break is repaired through the utilization of an exogenous piece of DNA that serves as a source of genetic information and the template for replication activity and gap filling. This general concept can help explain the appearance of these two insertional mutants and may also explain previous data

including the overall reaction of how single-stranded oligonucleotides direct point mutation repair in mammalian cells[3,5,116,123].

The RNP particle, as illustrated in Figure 28, Panel A, interacts at the target site and catalyzes a double strand break leaving two 3' hydroxyl ends available for extension by the DNA replication machinery. Non-homologous end joining activity then resects the broken ends and the degree of this resection varies from clone to clone (Panel B). The clones expanded from the uncorrected population support the fact that varying degrees of resection take place (see Figure 28B) because DNA insertions of 15 and 24 baes were found. As illustrated in Panel C, the oligonucleotide (red) pairs stably with the target gene via sequence complementarity, bridging the gap in the top strand. The binding is more stable upstream since the ssODN aligns in homologous register using perfect complementarity. Downstream from the break site, the base pairing must be incomplete because the data reveals a duplication of adjacent sequences. The partial binding downstream from the resected site is, in fact, energetically favorable based on calculations of free energy (approximately ΔG of -2.6). In our system, the oligonucleotide has been designed to be complementary to the non-transcribed strand and thus we can depict the polarity of pairing partners with confidence. As illustrated in Panel D and as a result of resection, a free 3' hydroxyl end on the top strand is now available for extension by DNA replication. In this simple model, the oligonucleotide acts as a template for the replication machinery to fill in the gap in the upper strand. For these two clones, the single-stranded oligonucleotide used in the gene editing reaction contains a G residue at its center because it is designed to create a single base mismatch with the G residue in the gene and promote mismatch repair. This strategy is based on work on single agent gene editing wherein the

objective is not DNA insertion but rather nucleotide exchange through the process of mismatch repair or by incorporation of the oligonucleotide into a growing replication fork (see models in references [3–5]). In contrast to some other models of gene editing[99,115,117,118], the sequencing data acquired here indicate that the oligonucleotide itself does not insert directly because, if this had happened, then the base at that position, identified in the genomic sequence (Figure 28C), would have been a G, not a C. After serving as a template for replication, the oligonucleotide dissociates (Panel D) and DNA replication is initiated on the opposite strand and in the opposite direction by utilizing the free hydroxyl group for extension as illustrated in Panel E. This variant of gene editing was termed, EXACT, for EXcision And Corrective Therapy; it may also be the general mechanism by which point mutations are repaired in gene editing reactions as directed by short oligonucleotides and double strand DNA breaks at the target site.

In principle though, these results do align with the conclusions of Schumann et al[105] and others[99,117] in that the activity of the CRISPR/Cas9 system provides a framework for the repair of resected regions of genomic DNA. Importantly, however, no fragment insertions were observed for point mutation repair because none of the clones examined in this study contained the G nucleotide at the target site. Since double strand DNA breaks are widely recognized as being both dangerous to cell viability and highly recombinogenic, it is likely that multiple pathways are used to regenerate a contiguous chromosome. The mechanism of repair may be dictated by the type and structure of donor DNA available at the site of damage.

The analysis of corrected and uncorrected cells reveals both the precision of gene editing and the development of genetic lesions, when indels are created in

uncorrected cells in the DNA sequence surrounding the target site. Through the above presented work it was possible to outline and develop the specific methodology used to analyze this combinatorial approach to the gene editing of a point mutation, coupled with a detailed experimental strategy to measuring indel formation at the target site. This protocol outlines a foundational approach and workflow for investigations aimed at developing CRISPR/Cas9-based gene editing for human therapy. The conclusion of this work is that on-site mutagenesis takes place as a result of CRISPR/Cas9 activity during the process of point mutation repair. This work puts in place a standardized methodology (see Appendix A) to identify the degree of mutagenesis[124]. While an enormous amount of attention has been paid to the analysis and mapping of off-site mutagenesis, it is likely that a heterogeneous mutation created at the target site will have a greater effect on the success or failure of gene editing in the clinical arena. Additional technologies or modified Cas9 proteins may be required to improve the precision of the homology-directed repair of inborn errors in mammalian cells [81]. Some of these technologies include the use of auxiliary oligonucleotides to act as a bridge, holding the chromosomal ends together and avoiding the destructive action of NHEJ. Defining the degree of heterogeneity at the target site as a result of gene editing activity is and must be an important part of any protocol designed for therapeutic intervention.

Mutant eGFP gene sequence:

5′ TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA **G**GGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT ‘3 **NT**

3′ AGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGGGAGCACTGGTG **GGACTGGATCCCGCACGTCACGA**AGTCGGCGATGGGGCTGGTGTACTTCGTCGTGCTGAAGAA ‘5 **T**

2C

72mer(NT):

3′ C*A*C*GGGACCGGGTGGGAGCACTGGTGGGACTGG**ATG**CCGCACGTCACGAAGTCGGCGATGGGGCTGGTG*T*A*C 5′

Wild Type eGFP gene sequence:

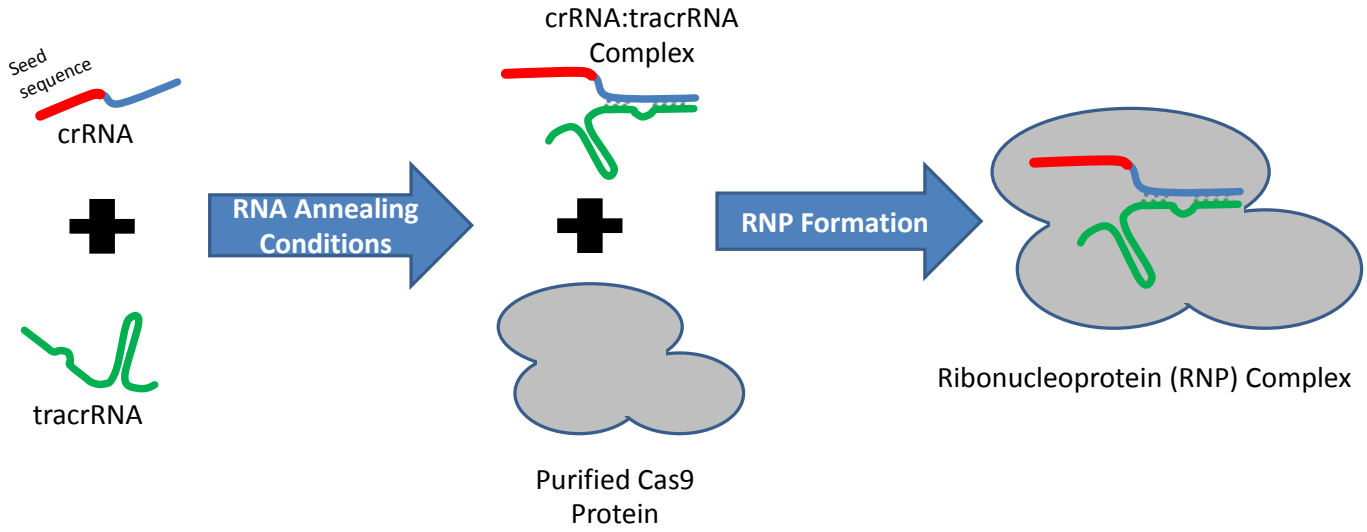
5′ TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA **C**GGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT ‘3 **NT**

3′ AGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGGGAGCACTGGTGGGACTGGAT **G**CCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTACTTCGTCGTGCTGAAGAA ‘5 **T**

CRISPR	Protospacer (5’-3’)
2C	**AGCACTGCACGCCCTAGGTC**

Figure 23. Model system for gene editing of the mutant eGFP gene.
The appropriate segments of the wild-type and mutated eGFP gene with the targeted codon, located in the center of the sequence, are displayed in green and red. The nucleotide targeted for exchange is bolded and underlined. The highlighted bases in blue represent the 2C CRISPR protospacer sequence and the orange bases highlight the PAM site. The oligonucleotide used in these experiments is 72 bases in length bearing phosphorothioate modified linkages at the three terminal bases; the 72-mer targets the non- transcribed (NT) strand (72NT).

A



CRISPR	crRNA	tracRNA
2C	5’-AGCACTGCACGCCCTAGGTC-3’	5'-GTTGGAACCATTCAAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTT-3'

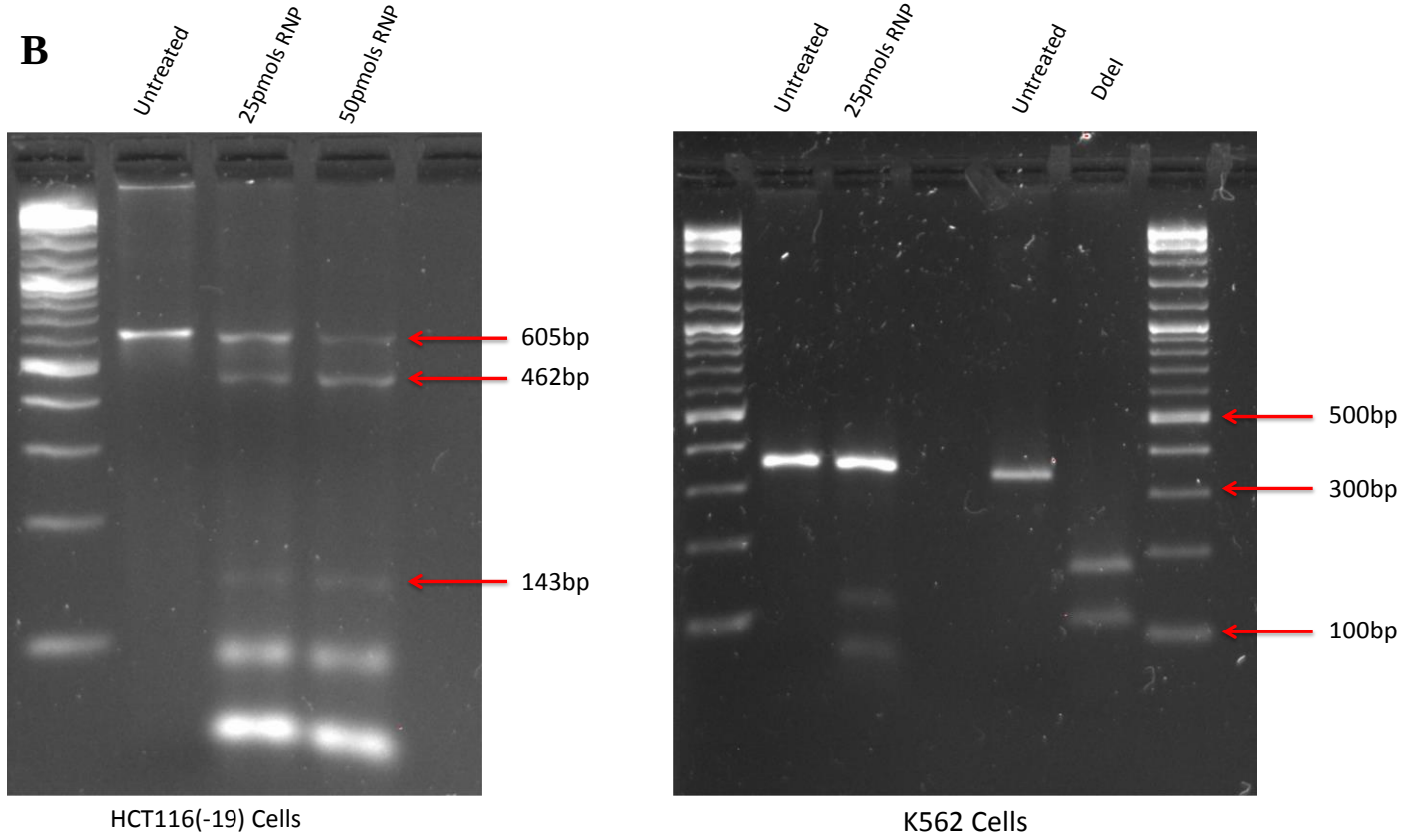
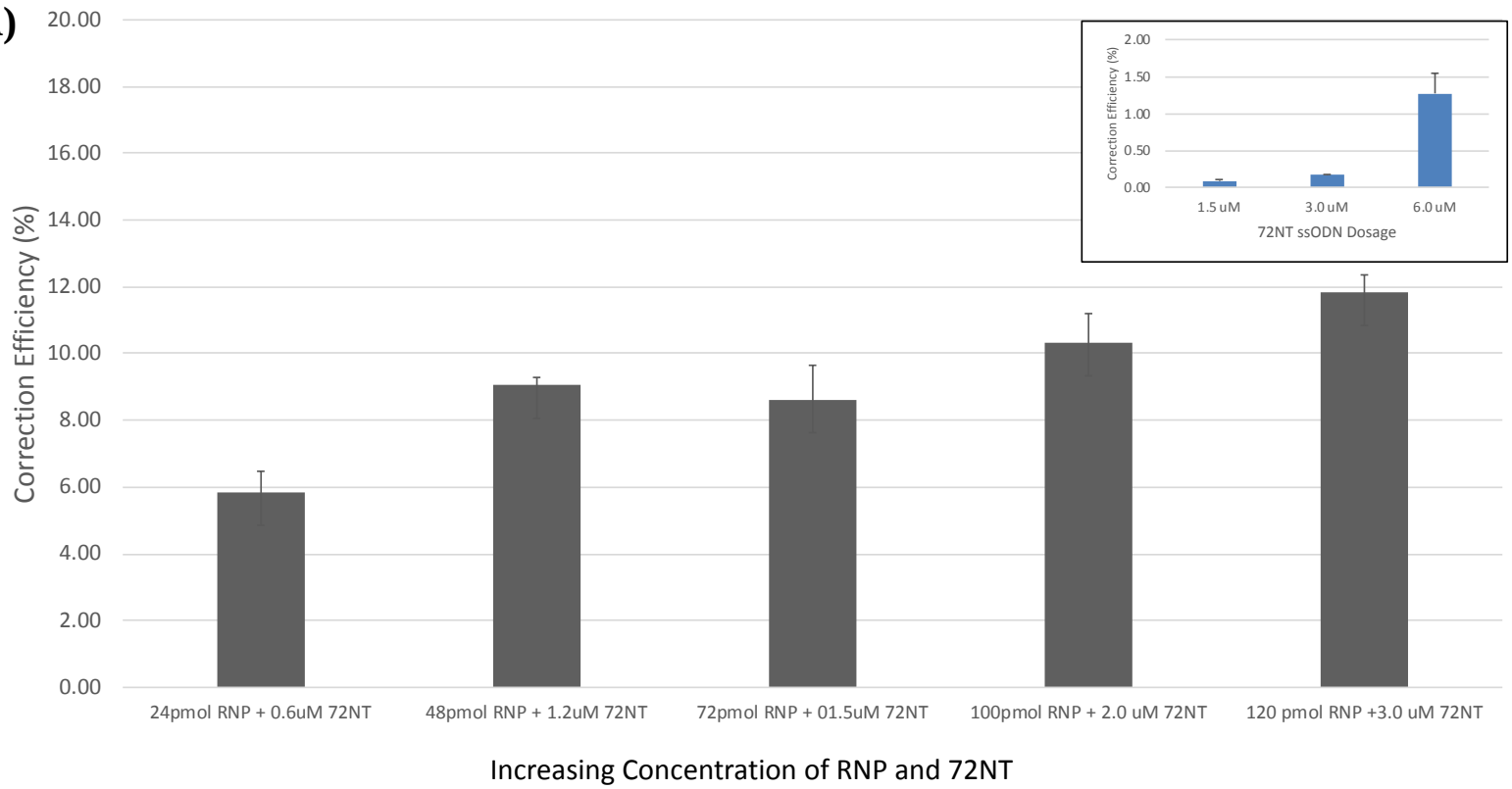


Figure 24. CRISPR/Cas9 Ribonucleoprotein Gene Editing of the mutant eGFP gene.
(A) CRISPR/Cas9 Ribonucleoprotein Assembly Reaction.
crRNA provides target specificity (20 bases, red section) corresponding to the 2C protospacer sequence and an interaction domain (blue) with the tracrRNA (green). crRNA and tracrRNA are annealed in equimolar concentrations. Cas9 protein (gray) is added to complete RNP assembly. Guide RNAs (gRNAs) direct and activate the Cas9 endonuclease which then

cleaves the target DNA. The lower section of the figure shows the 2C seed sequence and the tracrRNA sequence. **(B) In vitroRNP Digestion**. Genomic DNA was isolated from untreated HCT 116–19 cells and PCR used to generate an amplicon of size 605bp, which surrounds the sequence of the integrated mutant eGFP gene. The amplicon was combined with 25pmols and 50pmols of RNP complex respectively and incubated for 40 minutes at 37˚C. In the complete reaction, two products were generated with sizes consistent with fragments predicted from the specific cut site designed for the RNP complex. As a control, the RNP complex was incubated with an amplicon generated from the HBB gene 345 base pairs in length from cell line K562. A control digest was performed on the 345 base amplicon with the restriction enzyme DdeI.

(A)


20.00
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
Correction Efficiency (%)
24pmol RNP + 0.6uM 72NT
48pmol RNP + 1.2uM 72NT
72pmol RNP + 01.5uM 72NT
100pmol RNP + 2.0 uM 72NT
120 pmol RNP +3.0 uM 72NT
Increasing Concentration of RNP and 72NT
2.00
1.50
1.00
0.50
0.00
Correction Efficiency (%)
1.5 uM
3.0 uM
6.0 uM
72NT ssODN Dosage

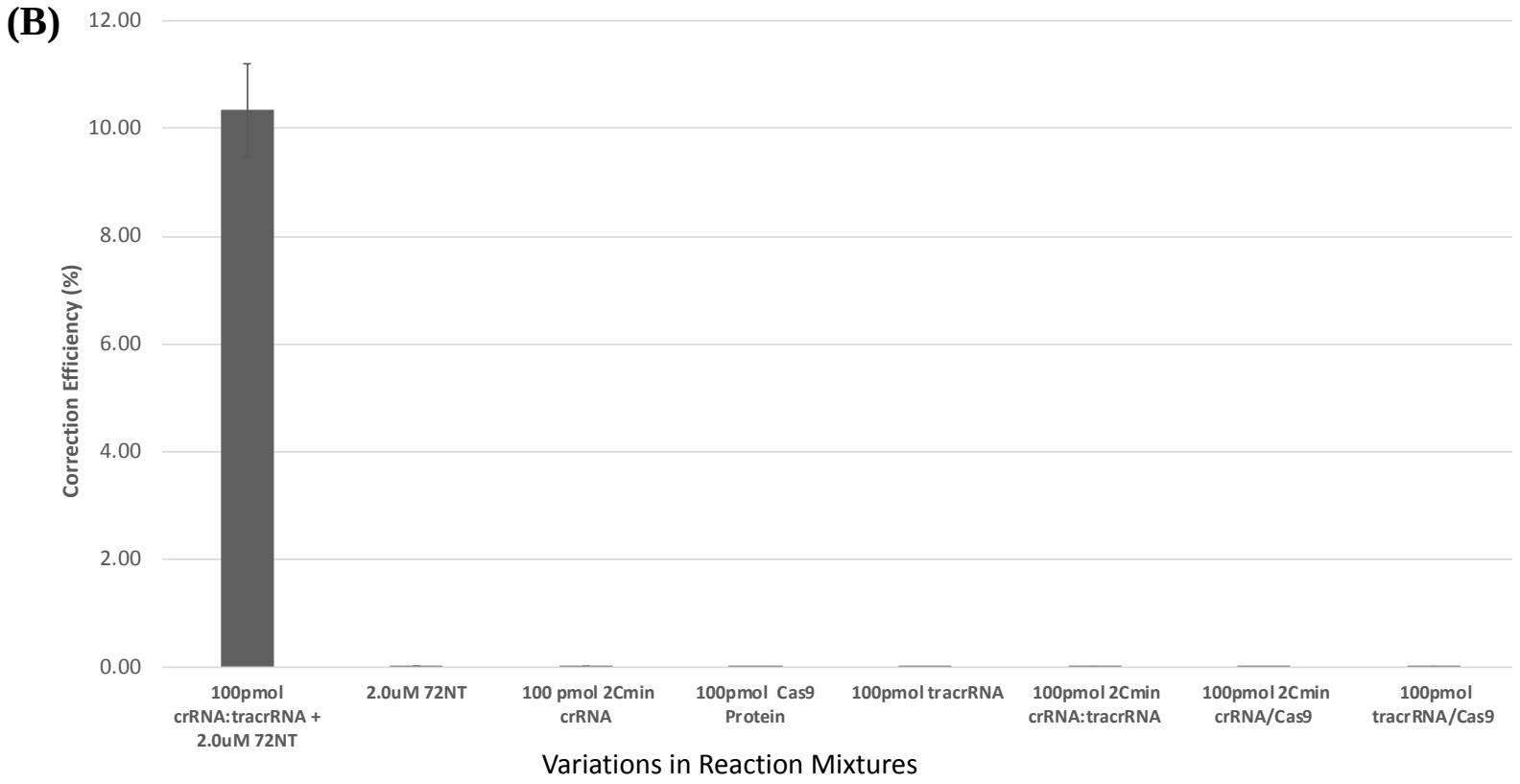
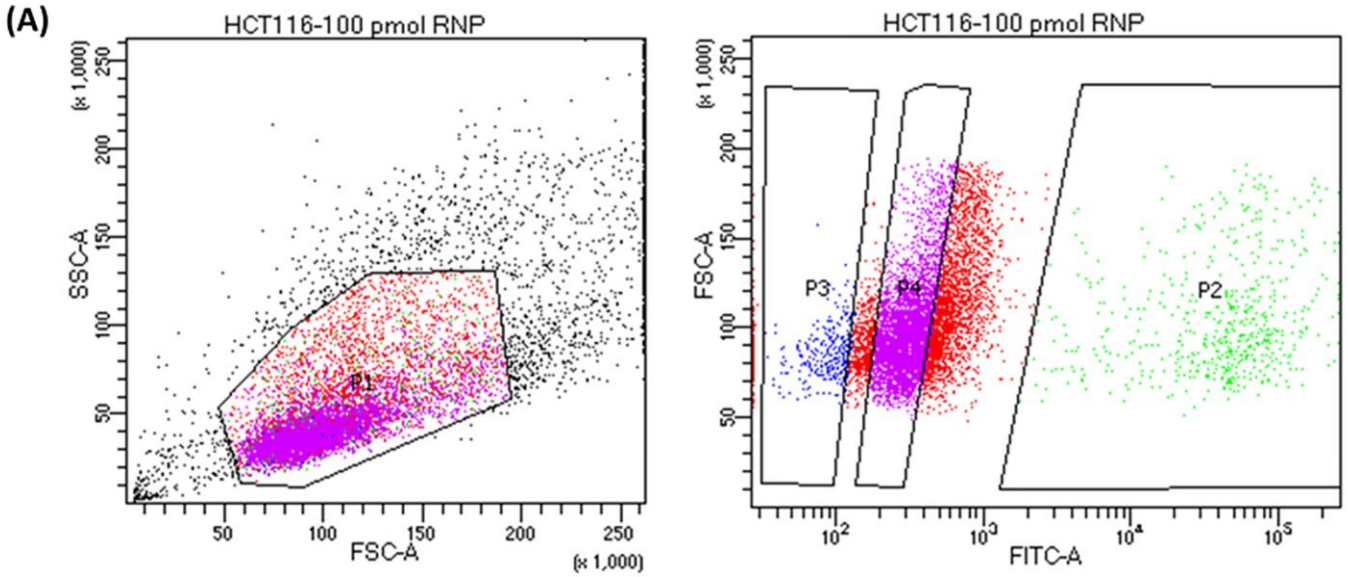


Figure 25. CRISPR/Cas9 Ribonucleoprotein Gene Editing of eGFP.
(A). Gene editing is dose dependent when directed by RNP and the ssODN.
Synchronized and released HCT 116–19 cells were electroporated with 24–120 pmol CRISPR/Cas9 RNP and 0.6–3.0 μM of 72mer. After a 72-hour recovery period, gene editing activity was measured using a FACSAria II flow cytometer. Gene editing is displayed as correction efficiency (%), determined by the number of viable eGFP positive cells divided by the total number of viable cells in the population. Each treatment was performed in triplicate and standard error is illustrated with accompanying bars. **Inset: Single agent gene editing.** Gene editing activity directed by the single-stranded

oligonucleotide (72NT) in the absence of the RNP complex under identical conditions is presented as a function of increasing concentration. **(B) Gene editing activity is dependent on all components being present in the reaction mixture.** Synchronized and released HCT 116–19 cells were electroporated with 100pmol of the crRNA, Cas9 Protein, tracrRNA and 2.0 μM of the 72NT, as a complete reaction. Identical mixtures, lacking the indicated reaction component, were carried out in parallel. In one specific reaction mixture, the RNP specific for the beta globin gene replaced the RNP specific for the eGFP gene (far right bar). After a 72 hour recovery period, gene editing activity was measured using a FACSAria II flow cytometer. Gene editing is displayed as correction efficiency (%), determined by the number of viable eGFP positive cells divided by the total number of viable cells in the population. Each treatment was performed in triplicate and standard error is illustrated with accompanying bars.



FACS Analyses

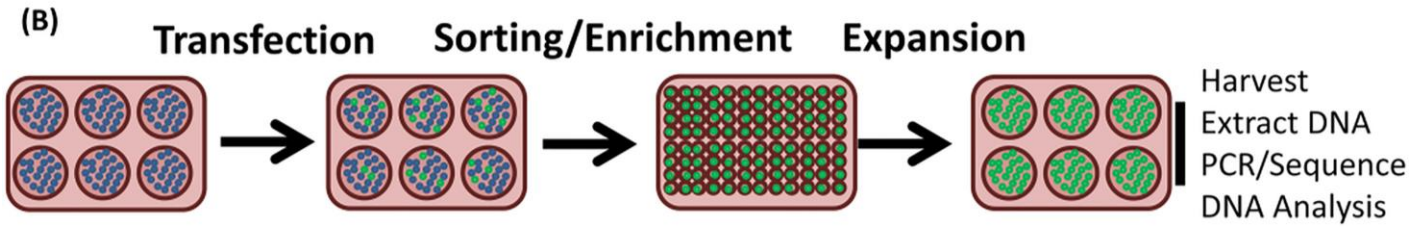
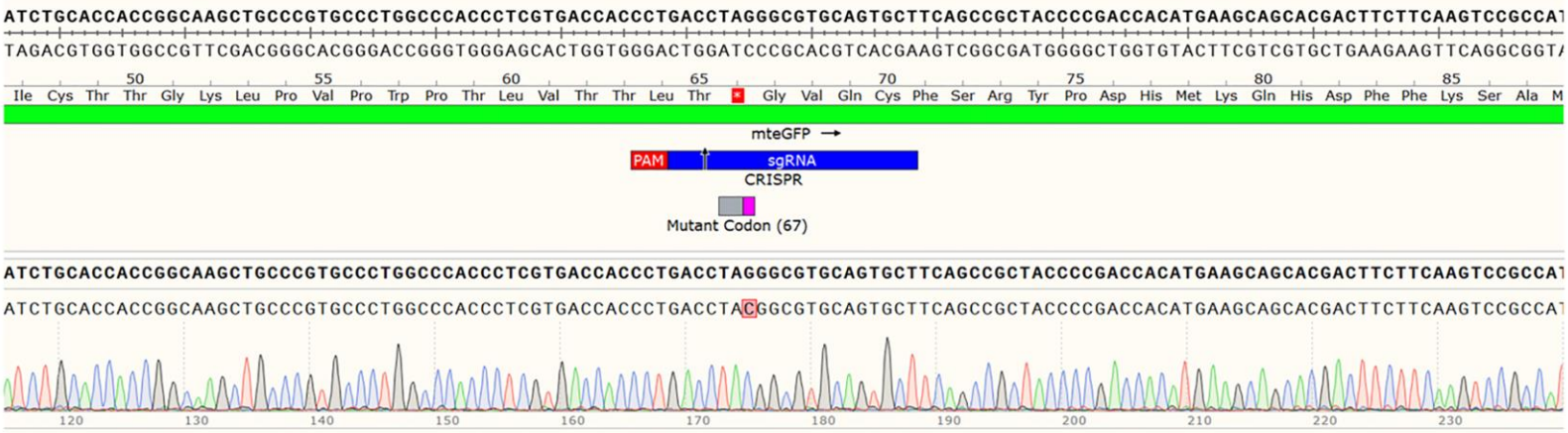
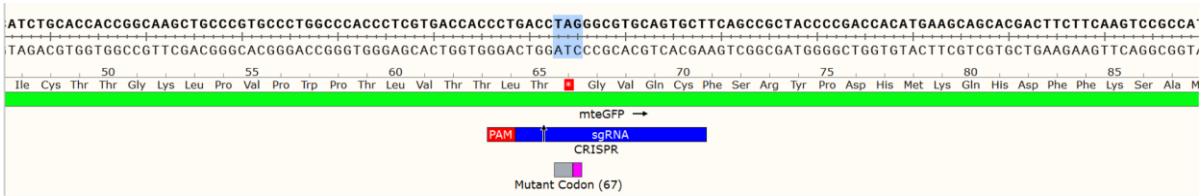


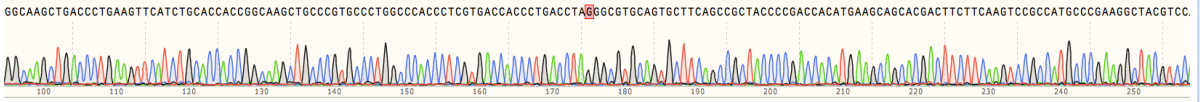
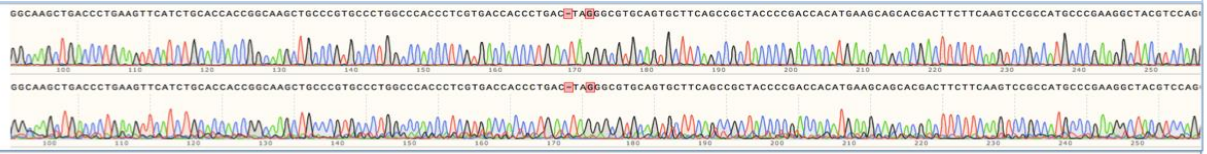
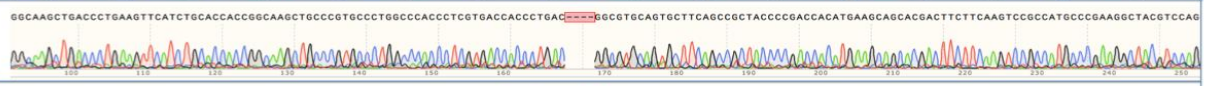
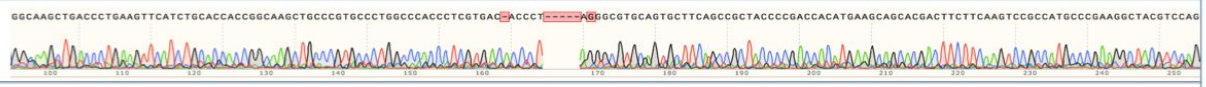
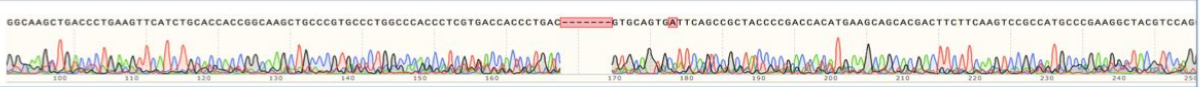
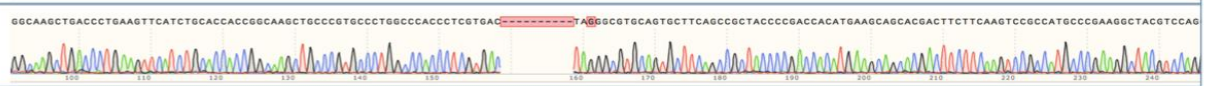
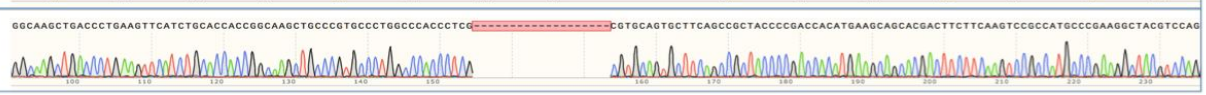
Figure 26. Experimental Design.
(A) FACSAria II plots of gene editing activity in HCT 116–19 cells.
HCT 116–19 cells synchronized for 24 hours at the G1/S border and released were electroporated with 100 pmol of RNP complex and 2.0 μM of the 72NT ssODN. After 72 hours, the cells were analyzed using FACS and single cells were sorted individually into 96-well plates. Two distinct populations were collected. The population of live, green cells (labeled as P2 on the FACS plot) as well as the population of live, non-green cells (labeled as P3) were segregated into separate clonal expansion plates. **(B) Experimental strategy isolation of single cell clones.** Cells exhibiting eGFP expression were scored positive and sorted using a FACSAria II flow cytometer as single cells into individual wells for clonal expansion. Cells lacking eGFP expression isolated and sorted in a similar fashion and expanded under the same conditions. DNA was then isolated and the eGFP gene was amplified and subjected to Sanger sequencing to analyze gene editing activity surrounding the target site.

(A)


ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTAGGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA
TAGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGGGAGCACTGGTGGGACTGGATCCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTACTTCGTCGTGCTGAAGAAGTTCAGGCGGTA
50 55 60 65 70 75 80 85
Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys Ser Ala M
mteGFP
PAM
sgRNA
CRISPR
Mutant Codon (67)
ATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA
120 130 140 150 160 170 180 190 200 210 220 230

(B)



No change **8 clones**	
1 bp deletion	
4 bp deletion	
5 bp deletion	
7 bp deletion	
10 bp deletion	
14 bp deletion	
19 bp deletion	

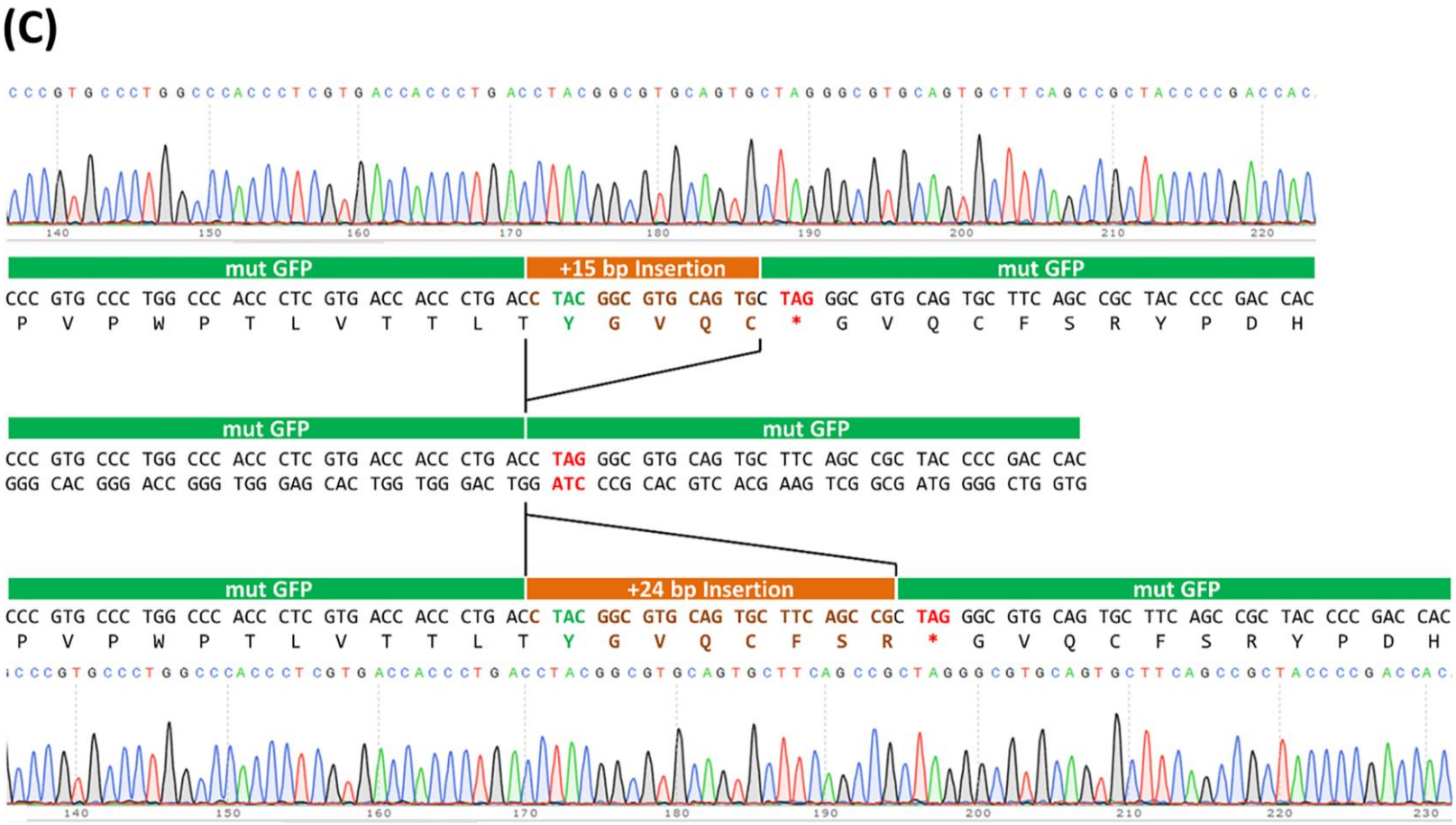


Figure 27. CRISPR/Cas9 Ribonucleoprotein Gene Editing Allelic Analysis.
(A) Allelic analysis of eGFP positive cells expanded as a clonal population. Clonally isolated and expanded eGFP positive samples (sixteen clones) were analyzed at the site surrounding the targeted base and DNA from each, harvested, purified, amplified and sequenced. Allelic analysis was carried out using Sanger sequencing, assembled using SnapGene

and compared to the sequence of a wild-type allele which is illustrated at the top of the figure; the cut site of the RNP complex is indicated as a small black arrow located on the green bar (2C crRNA). **(B) Allelic analysis of eGFP negative cells expanded as a clonal population.** Fifteen individual samples, expanded from cloned originating from the uncorrected population were randomly selected and analyzed for indel formation at the site surrounding the target nucleotide. As above, allelic analysis was carried out using Sanger sequencing and assembled SnapGene. Once again, the sequence of a wild-type allele at the top of the figure along with the cut site of the RNP is presented. **(C) Allelic analysis of eGFP negative cells presenting insertions.** Two individual clones from the uncorrected population displayed insertions of 15bp (top panel) and 24bp (bottom panel), respectively. The center panel represents the mutant eGFP gene sequence with the mutant codon in red. The inserted bases are highlighted in orange with the corrected tyrosine codon depicted in green and the mutant stop codon represented by a red asterisk. The boundaries of the insertions are denoted by black bars.

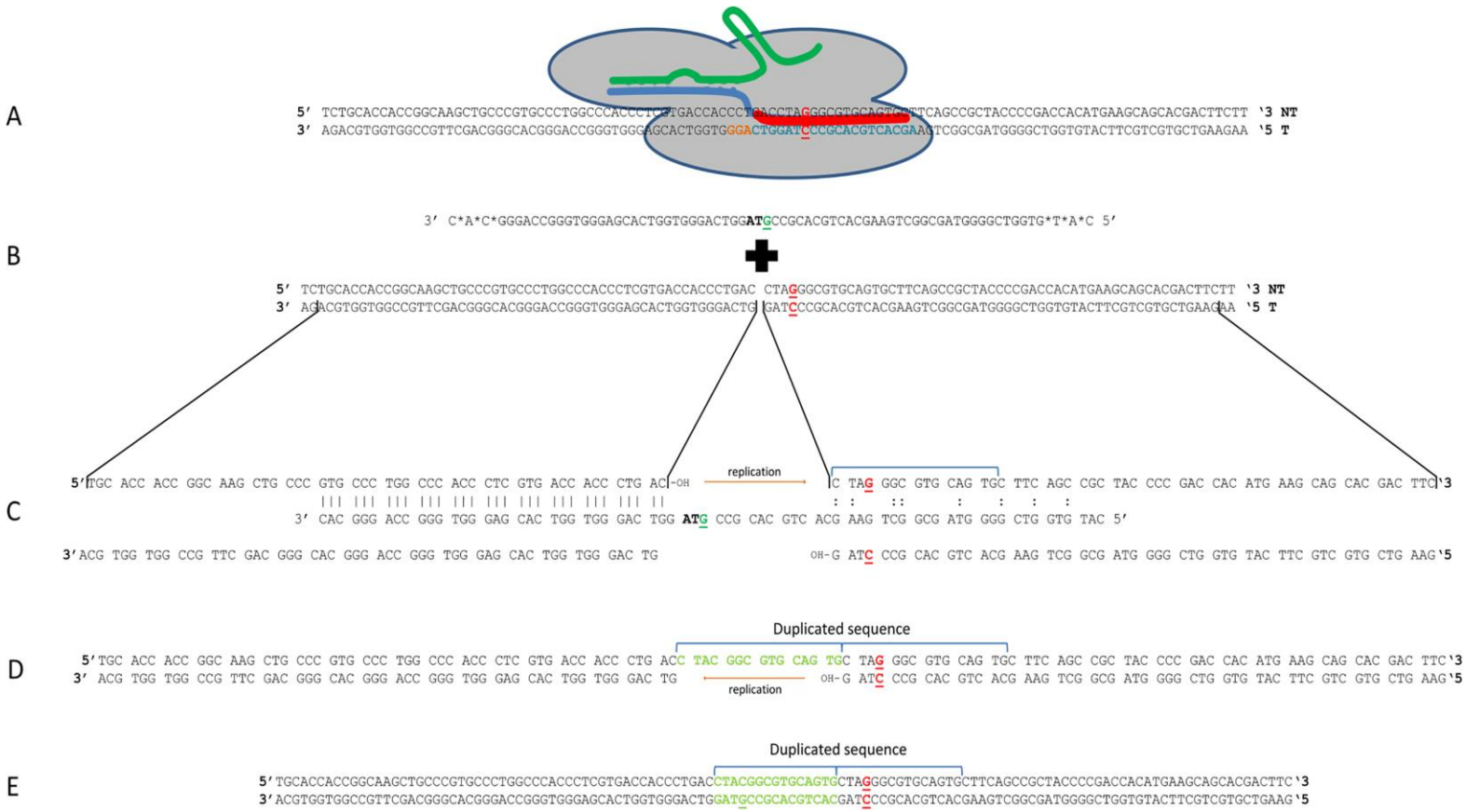


Figure 28. A model for point mutation repair directed by an RNP complex and a short single-stranded DNA oligonucleotide.

(Panels A and B); the RNP particle induces a double strand break at the target site generating two free 3' hydroxyl ends on each strand of the broken DNA. (Panel C); the oligonucleotide aligns in imperfect homologous register with the non-transcribed strand of the chromosome. The DNA replication machinery fills the gap starting from the 3' hydroxyl end and

completing by ligation to the 5’ phosphate at the opposite side of the gap. The single-stranded oligonucleotide serves as a template for the replication process. (Panels D/E), dissociation of the single-stranded oligonucleotide allows for the newly synthesized DNA to act as a template for DNA replication in the opposite direction on the bottom strand followed by ligation.

Chapter 5

HIGH LEVELS OF MAMMMALIAN CELL TRENSFORMATION IS NOT A VALID INDICATOR OF SUCCESSFUL GENE REPAIR DIRECTED BY SINGLE -STRANDED OLIGONUCLEOTIDES AND A CRISPR/CAS9 RIBONUCLEOPROTEIN PARTICLE

Introduction

Combinatorial targeting with ssODNs and a properly designed CRISPR/Cas9 RNP produces correction frequencies between 6 and 12% routinely. A key reaction parameter is the placement of the cleavage site within 50 bases upstream or downstream from the point mutation [37,101]. Yet, it is still unclear if the activity takes place as a function of the efficient RNP uptake into the target cells. Here, focus on testing and better understanding the relationship between efficient delivery of the CRISPR/Cas9 RNP complex into the nucleus and its correlation to gene editing activity. These experiments were carried out to determine if the level of transformation of a transformed cell line with a ribonucleoprotein (RNP) particle, through electroporation, can predict the degree of single base repair of a mutated eukaryotic gene. This will serve as a guide to design gene editing experiments in complex cell lines and most importantly, primary cells.

Methods and materials

Cell Culture Conditions

HCT116 cells were acquired from ATCC (American Type Cell Culture, Manassas, VA). The HCT116-19 was created by integrating a pEGFP-N3 vector

(Clontech, Palo Alto, CA) containing a mutated eGFP gene. The mutated eGFP gene has a nonsense mutation at position +67 resulting in a nonfunctional eGFP protein[85]. For these experiments, HCT116 (-19) cells were cultured in McCoy's 5A Modified medium (Thermo Scientific, Pittsburgh, PA) supplemented with 10% fetal bovine serum, 2mM L-Glutamine, and 1% Penicillin/Streptomycin. Cells were maintained at 37°C and 5% CO_2. The eGFP targeting custom designed 72-mer oligonucleotide was synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

Assembly of Cas9 RNP Complex

The recombinant Cas9 protein, ATTO647-labeled tracrRNA, ATTO550-labeled tracrRNA, HBB gene targeting custom made crRNAs (G5 and G10), ATTO550-labeled crRNA, and eGFP crRNA were a gift from IDT (Integrated DNA Technologies, Coralville, IA). The individual components were reconstituted to make the RNP complex at various molar ratios of protein and RNA, 1:1 crRNA and tracrRNA: RNP. The RNP complex was built and utilized according to the ALT-R CRISPR/Cas9 System transfection protocol provided on IDT's website.

Confocal Microscopy

HCT116-19 cells post-electroporation were grown in Nunc Lab-Tek II Chambered Coverglass plates for 72 hours, rinsed with PBS and fixed with BD Cytofix/Cytoperm Fixation and Permeabilization Solution at 4˚C for 20 minutes. The fixation solution was aspirated, cells were rinsed with PBS twice, and a drop of ProLong Diamond Antifade Mount with DAPI (Life Technologies) was applied onto the cells. The treated cell samples were investigated under a ZEISS LSM880 Laser

Scanning Confocal Microscope at DBI. A 63x oil objective lens and four channel light sources (bright field transmitted light, 405nm, 488nm, and 561nm lasers) were used. In order to construct the 3D images of the cells, Z-stacking and 3D building functions of the confocal microscope system were used. The Z-stacking slice intervals were set as 280nm, and the pixel size was selected as1024 x 1024. An average of 4 or 8 scanning data per light channel were acquired for each sample. The image data were processed to construct stationery images of 3D movies using Image J (NIH).

Gene editing reactions

HCT116-19 cells were targeted and analyzed as previously described [45]. The BioRad Gene Pulser was used to deliver the RNP and the ssODN. The eGFP targeting custom designed 180-mer oligonucleotide was synthesized by IDT (Integrated DNA Technologies, Coralville, IA).

Results

What influences the accuracy and efficiency of gene editing in Primary cells?

Figure 29 displays the standard targeting model system with the G residue at the third position of the TAG stop codon, highlighted in red. In addition, the 72-base single-stranded oligonucleotide is also presented. For this study, an RNP constructed following manufacturer's (IDT) annealing suggestions was utilized. The pairing of the two RNA components precedes the addition of the purified Cas9 protein which after complexation generates the complete RNP particle. These two components are traditionally are electroporated into HCT 116 –19 cells which bear the mutant gene accounting for our standard gene editing reaction. The position of the protospacer binding which aligns the RNP at the target site is illustrated in blue. For this particular

experiment, the transformation of the RNP into HCT 116 -19 cells was analyzed by utilizing a tracrRNA labeled with an ATTO 647 at the 5' end, which is detected with a 633 laser. As part of the assembled and active RNP, the uptake of the RNP into the cells was measured after 24 hours using various dosages ranging from 10 picomoles to 75 picomoles respectively along with an electroporation enhancer at equimolar amounts. As seen in Figure 30, the RNP particle is introduced into HCT 116 -19 cells in a sigmoidal fashion, with a large increase seen between 20 pmol and 50 pmol respectively. At 75 pmol, approximately 64.2% of the cells are visibly containing the RNP. Analyses of RNP uptake were carried out using the BD FACS AriaII Flow Cytometer and the FACS DIVA software. Figure 31A is essentially a repeat of the previous experiment except that the length of time allowed for RNP incorporation into the cells has been extended to 72 hours and the electroporation enhancer was exchanged for a targeting 180mer ssODN to direct point mutation repair. With this extended timeline, a remarkably high percentage of cells, approaching 94%, appear to contain the RNP particle. In this case, the 10 and 20 pmol levels of transformation exhibit more similarity to the higher levels of RNP particles electroporated seen in this figure and previously in Figure 30. For this particular time point, the frequency of gene editing activity was coordinately measured by analyzing the restoration of productive and visible eGFP production. Those numbers are given in the upper right-hand quadrant of Figure 31A and reveal a more typical sigmoidal dose curve with a significant increase in gene editing activity observable between 20 and 50 pmol. The 50 and 75 pmol dosages appear to plateau suggesting that the system for gene editing activity has become saturated. In the gene editing system used here, significant success utilizing a 72-base single-stranded oligonucleotide as an effector donor DNA

molecule has been achieved to direct point mutation repair in these cells, while others have reported an increase in gene editing frequency as a function of length of the single-stranded oligonucleotide. This notion was tested in a series of experiments by exchanging the 72-mer with a 180-mer of identical, but extended equally in both directions, sequence. As shown in Figure 31B, and after 72 hours of incubation, we observed that the 180-base single-stranded oligonucleotide did not increase gene editing activity at the 75 pmol level. In fact, it could be argued that there is a discernible lowering of gene editing activity when the longer DNA oligonucleotide is utilized in this particular reaction mixture. Figure 32 is composite representation that directly compares in a side-by-side fashion, the degree of transformation and transfection of HCT 116 – 19 cells with a 5' ATTO 647 fluorescent dye labeled tracrRNA component of the RNP particle with the gene editing activity in the form of a correction of a single base point mutation. As can be seen in this figure, while extensive levels of transformation are achieved with as little as 10 pmol of RNP particle using a standard electroporation protocol, significant levels of correction do not appear until 50 pmol of RNP are transfected into the cells. In addition, the data also demonstrate that an increase in the dosage of the RNP particle increases the degree of intensity of the labeled molecule in the cells. In other words, while a rather insignificant jump in transfection efficiency appears between 20 and 50 pmol, a significant increase in the intensity of cells bearing the RNP is observed. That point lies between the 20 pmol and 50 pmol where a corresponding jump in gene editing activity is also seen.In addition, the data also demonstrate that an increase in the dosage of the RNP particle increases the degree of intensity of the labeled molecule in the cells. Figure 33 shows that the transfection results of non-primary cell, HCT116-

19 colon cancer cells that harbor mutated eGFP gene, targeted using CRISPR/Cas9 RNP and ssODNs. The progression of the cell depth is revealed as the slice number of the Z-stacking increases. RNPs were distributed throughout the HCT116-19 cells, including the nucleus, with noticeable gene correction of eGFP visible in one of the cells in the view. The eGFP corrected cell exhibits green while the labeled RNPs are distributed throughout the DAPI stained nucleus, exhibiting pinkish color as the slice number increases, and then disappear again as the sections reach the cell bottom.

Discussion

The acclaimed genetic engineering tool, CRISPR/Cas9, will realize its full potential when robust and reproducible methods for delivering the package into primary cells are fully developed. It is widely accepted that improving the transport efficiency into the cell, and moreover into the nucleus, is the benchmark of success for generating a genetically altered genome. To establish a baseline relationship between delivery and successful gene editing, first a well-established gene editing system was utilized, in which all of the reaction components have been validated and clearly defined[45]. Through this study it can be concluded that while significant RNP delivery is achieved at low concentrations, a significant level of gene editing activity is realized only when the highest levels of RNP are electroporated into the cell. It was also shown that nuclear delivery is easily achieved in the HCT116-19 cell model systems. These observations lead to the hypothesis that there is no direct correlation between efficient cellular uptake and genome modification directed by an RNP. Taken together, these data suggest that gene editing activity is facilitated when a significant number of RNP particles have entered an individual cell and that high levels of

transfection frequency in the absence of highly intense individual cells bearing the RNP is not a sufficient nor predictable measure of the outcome of a gene editing reaction. The data herein reported may begin to establish the basis for carving out guidelines as to how best to evaluate practicality and efficiency of gene editing activity in cells for clinical application.

A

Mutant eGFP gene sequence:

5′ TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA GGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT ‘3 NT

3′ AGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGGGAGCACTGGTG GGACTGGATCCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTACTTCGTCGTGCTGAAGAA ‘5 T

2C

72mer(NT):

3′ C*A*C*GGGACCGGGTGGGAGCACTGGTGGGACTGG**ATG**CCGCACGTCACGAAGTCGGCGATGGGGCTGGTG*T*A*C 5′

Wild Type eGFP gene sequence:

5′ TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTA CGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT ‘3 NT

3′ AGACGTGGTGGCCGTTCGACGGGCACGGGACCGGGTGGGAGCACTGGTGGGACTGGAT GCCGCACGTCACGAAGTCGGCGATGGGGCTGGTGTACTTCGTCGTGCTGAAGAA ‘5 T

CRISPR	Protospacer (5’-3’)
2C	AGCACTGCACGCCCTAGGTC

B

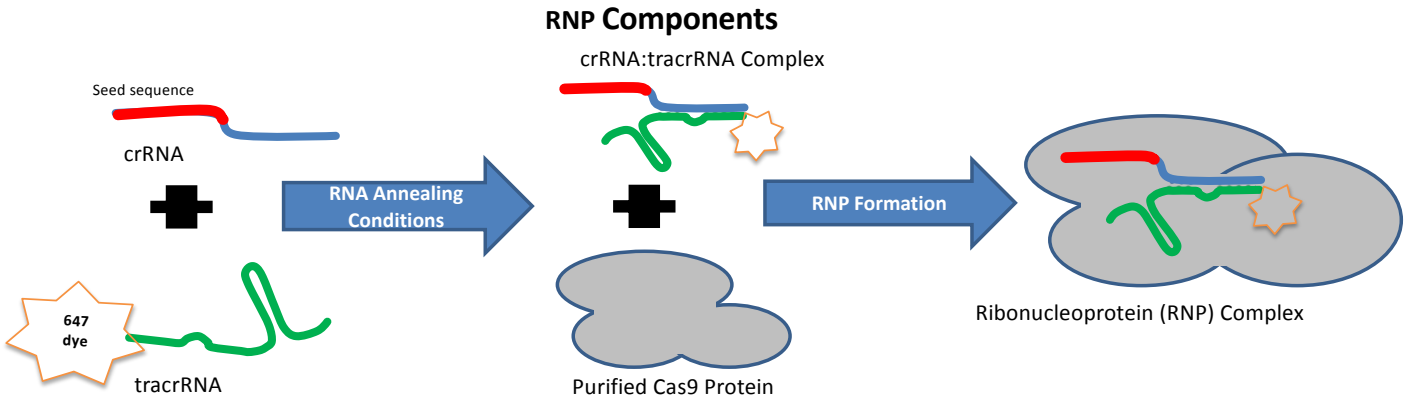


Figure 29. Model system for gene editing of the mutant eGFP gene with CRISPR/Cas9 Ribonucleoprotein.
(A.) Model system for gene editing of the mutant eGFP gene. The appropriate segments of the wild-type and mutated eGFP gene with the targeted codon, located in the center of the sequence, are displayed in green and red. The nucleotide targeted for exchange is bolded and underlined. The highlighted bases in blue represent the 2C CRISPR protospacer sequence and the orange bases highlight the PAM site. The oligonucleotide used in these experiments is 72 bases in length

bearing phosphorothioate modified linkages at the three terminal bases; the 72-mer targets the non- transcribed (NT) strand (72NT). **(B) CRISPR/Cas9 Ribonucleoprotein Assembly Reaction**. crRNA provides target specificity (20 bases, red section) corresponding to the 2C protospacer sequence and an interaction domain (blue) with the tracrRNA (green) which has an ATTO 647 fluorescent dye attached to the 5' end. crRNA and tracrRNA are annealed in equimolar concentrations. Cas9 protein (gray) is added to complete RNP assembly. Guide RNAs (gRNAs) direct and activate the Cas9 endonuclease which then cleaves the target DNA. The lower section of the figure shows the 2C seed sequence and the tracrRNA sequence.

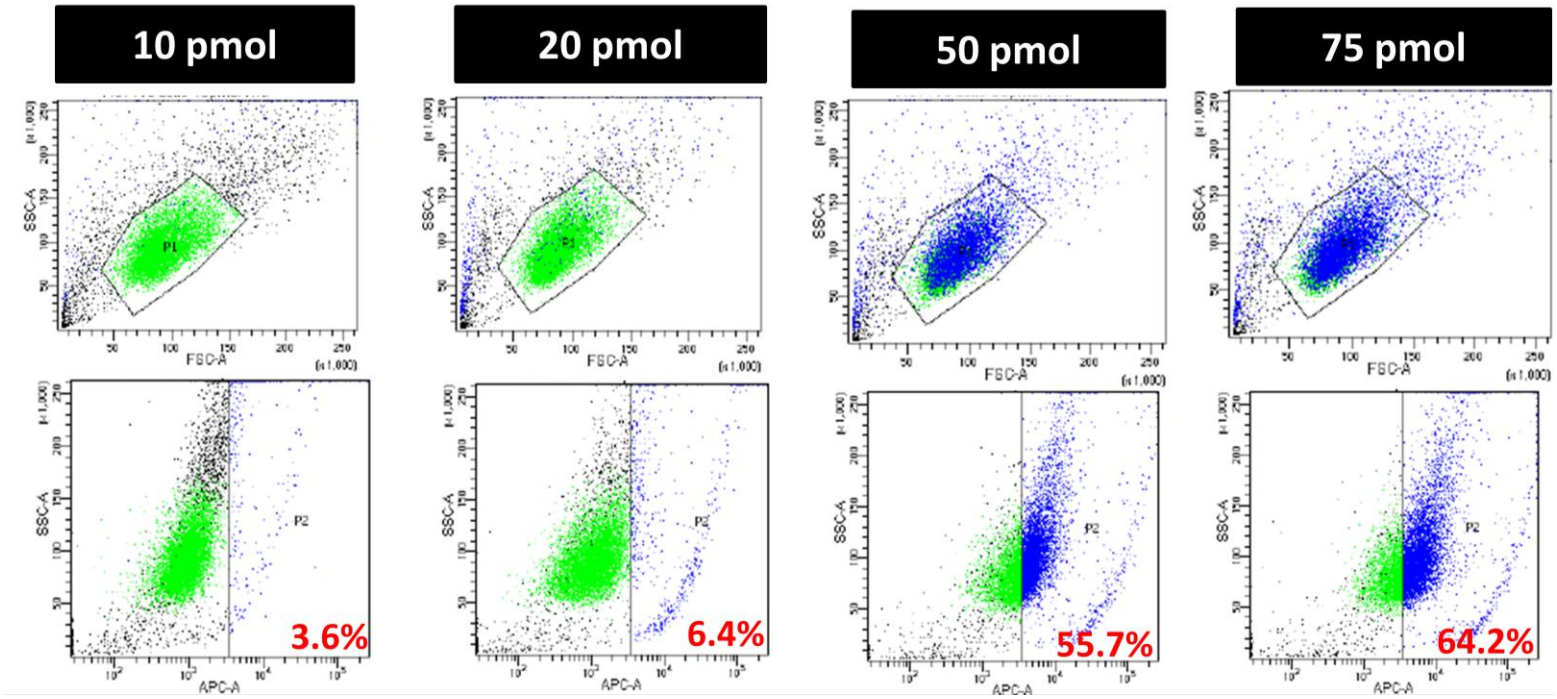
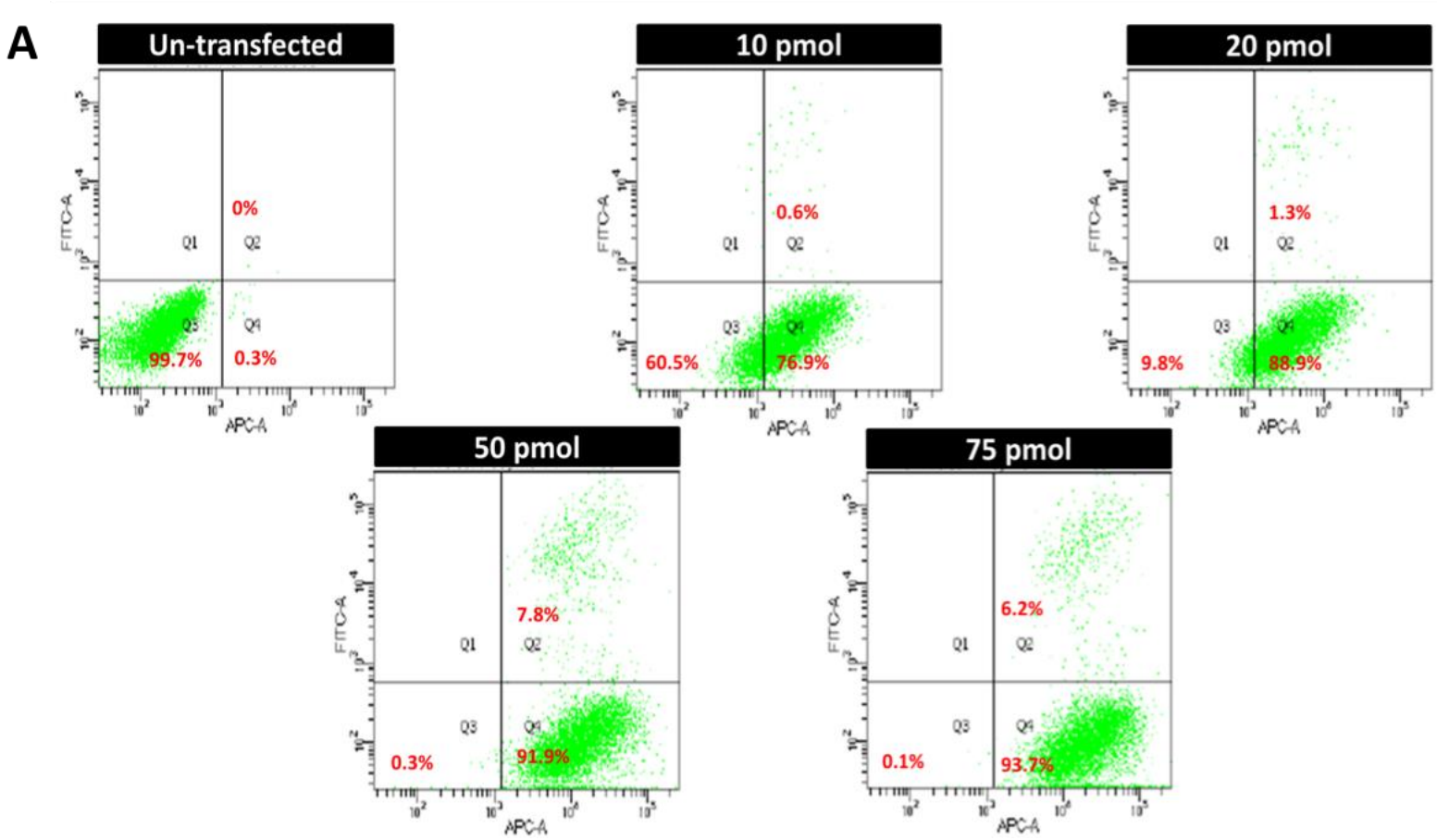


Figure 30. Transfection of CRISPR/Cas9 RNP into HCT116 is dose dependent.
HCT 116-19 cells were electroporated with10-75 pmol at equimolar amounts of CRISPR/Cas9 RNP and IDT enhancer. The tracrRNA used for these experiments has ATTO 647 fluorescent dye attached to the 5' end. Transfection was measured at 24 hours using a FACSAria II flow cytometer. The percent of transfection is showed in red.


A
Un-transfected
0%
99.7%
0.3%
10 pmol
0.6%
60.5%
76.9%
20 pmol
1.3%
9.8%
88.9%
50 pmol
7.8%
0.3%
91.9%
75 pmol
6.2%
0.1%
93.7%
FITC-A
APC-A
Q1
Q2
Q3
Q4

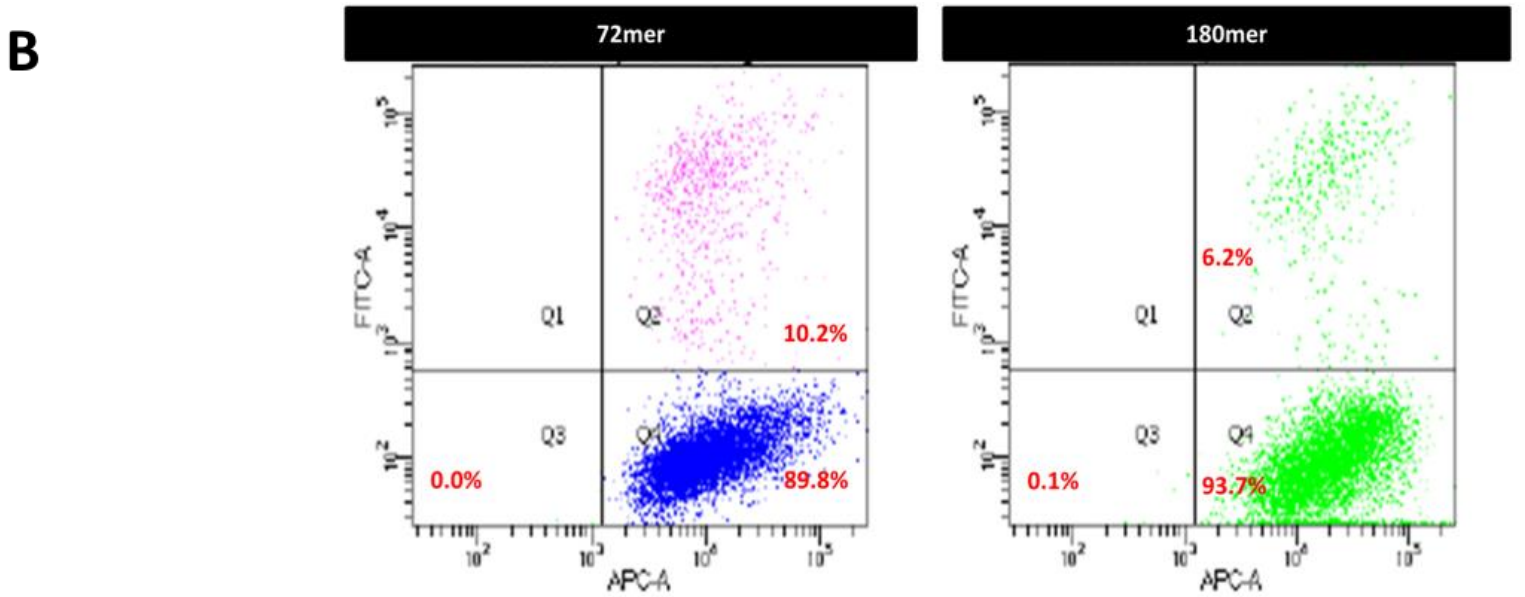


Figure 31. Relationship between transfection and gene correction.
(A.) Synchronized and released HCT 116-19 cells were electroporated with10-75 pmol at equimolar amounts of CRISPR/Cas9 RNP and 180 NT. The tracrRNA used for these experiments has ATTO 647 fluorescent dye attached to the 5' end which permitted the measurement of RNP transfection. After a 72-hour recovery period, transfection and gene editing activity was measured using a FACSAria II flow cytometer. Quadrant 2 shows the cells that were positive for both correction (+eGFP) and transfection (ATTO 647 dye). Quadrant 4 shows the cells that were positive for CRISPR/ Cas9 RNP alone. **(B) Increased length of the ssODN does not increase gene correction.** Synchronized and released HCT 116-19 cells were electroporated with 75 pmol at equimolar amounts of CRISPR/Cas9 RNP and 72mer or 180mer. The tracrRNA used for these experiments has ATTO 647 fluorescent dye attached to the 5' end which permitted the measurement of RNP transfection. After a 72-hour recovery period, transfection and gene editing activity was measured using a FACSAria II flow cytometer. Quadrant 2 shows the cells that were positive for both correction (+eGFP) and transfection (ATTO 647 dye). Quadrant 4 shows the cells that were positive for CRISPR/ Cas9 RNP alone.

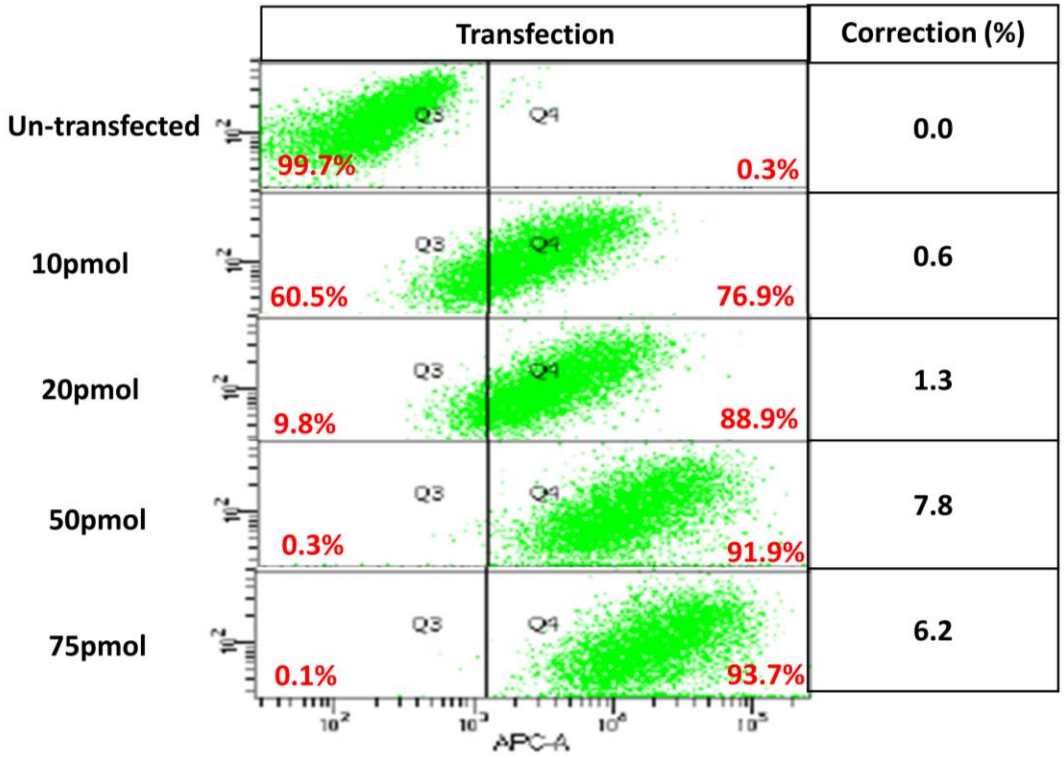


Figure 32. Degree of transformation and transfection of HCT 116 – 19 cells with a 5' ATTO 647 fluorescent dye labeled tracrRNA.

Synchronized and released HCT 116-19 cells were electroporated with10-75 pmol at equimolar amounts of CRISPR/Cas9 RNP and 180 NT. The tracrRNA used for these experiments has ATTO 647 fluorescent dye attached to the 5' end which permitted the measurement of RNP transfection. After a 72-hour recovery period, transfection was measured using a FACSAria II flow cytometer. Quadrant 3 shows the cells that were not transfected. Quadrant 4 shows the cells that were positive for CRISPR/ Cas9 RNP (647 fluorescent dye positive).

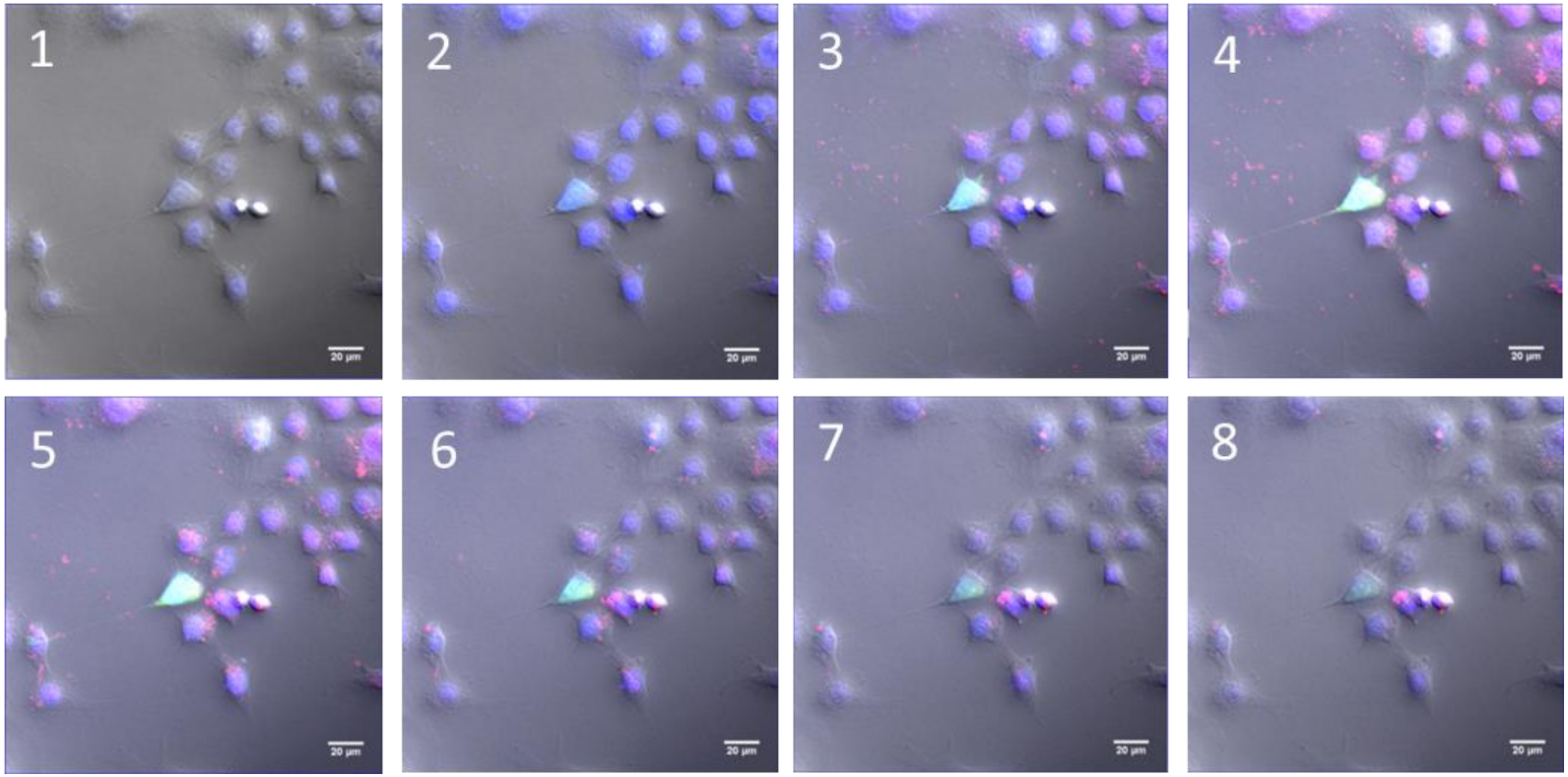


Figure 33. Representative z stack images of HCT116-19 cells transfected with ATTO550-labeled RNP complex 16 hr post-transfection.
The z stack top view images show a group of cells with gradual increment of the confocal slices. The green cell in the field of view exhibited gene editing due to a corrected EGFP gene. Blue represents DAPI-stained nuclei, and red represents ATTO550-labeled RNP.

Chapter 6

ENGINEERING LEUKEMIC HUMAN TUMOR-ASSOCIATED CHROMOSOMAL TRANSLOCATION t(4;11)(q21;q23) WITH CRISPR/CAS9 SYSTEM

Introduction

Rearrangements of the MLL gene are responsible for leukemogenesis. Molecular events result in the fusion of an array of partner genes that lead to infant acute leukemia. Importantly, the types of genetic fusion through genetic pairing influences prognosis. For example, a t(4;11)(q21;q23) translocation involving the genes MLL and AF4 (alias for AFF1) detected in 50-70% of infant leukemia is the most aggressive with a poor outcome (0% 3yr EFS). Genome editing tools have now been developed that can enable a recapitulation of the initial steps of leukemogenesis and in this study we present a strategy to induce the t(4;11)(q21;q23) translocation utilizing an RNA-guided CRISPR Cas9 system in a human cell line. A specifically designed CRISPR/Cas9 system is utilized for the generation of chromosomal translocation via double stranded breaks at specific sites. This approach will generate primary human cells as well as cell lines that will enable a molecular analysis of the events accompanying chromosomal rearrangements and a delineation of the variable prognosis. This new cell system will also form the basis for identification of new therapeutics for childhood leukemia. Genetic engineering tools are being used to help in the discovery of new anti-leukemic drugs by enabling the creation of modified cell lines that recapitulate the molecular events that precede the onset of leukemogenesis.

The goal of this study was to develop a genetically engineered cell line bearing a (4:11)(q21:q23) translocation that appears in a series of pediatric leukemias and bears a poor outcome. Here CRISPR-directed gene editing approach was successfully used to catalyze the exchange of the appropriate segments of chromosome 4 and chromosome 11 respectively. This cell line can now be used for the screening of appropriate compounds designed to inhibit leukemogenesis in children bearing this chromosomal translocation. We will provide the cell line to any group or individual in the scientific community that desires to use it for such drug discovery purposes.

Methods and materials

Cell Line and Culture Conditions

HEK 293T (ATCC CRL-3216) and MV-4-11 cells (ATCC CRL-9591) were acquired from ATCC (American Type Cell Culture, Manassas, VA). HEK 293T cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (ATCC 30-2002). To make the complete growth medium supplemented with: 10% Fetal Bovine Serum (heat inactivated) (ATCC 30-2020), 2mM L-glutamine (ATCC 30-2214), and 1% Penicillin/Streptomycin. MV-4-11 cells were cultured in Iscove's Modified Dulbecco's Medium, (ATCC 30-2005) supplemented with: 10% Fetal Bovine Serum (heat inactivated). Cells were maintained at 37°C and 5% CO_2.

CRISPR Design and Construction

The KMT2a and AFF1 gene sequences were entered into the Zhang Lab's online generator (http://crispr.mit.edu/) and the appropriate CRISPR guide sequence which binds in close proximity to the break point junction site of the t(4;11)(q21;q23) was chosen. The gRNA was cloned into the pX458 backbone vector (Addgene

plasmid 48138), a human codon optimized pSpCas9 and chimeric guide RNA expression plasmid with a 2A-eGFP. pX458 was purchased through Addgene (https://www.addgene.org). Following construction, clones were verified by DNA sequencing by Eton Bio Incorporated (Union, NJ).

Transfection of HEK293 Cells and Experimental Approach

HEK 293T Cells were transfected by CalPhos Mammalian Transfection Kit (Clontech Laboratories, Inc.) with endotoxin-free DNA (Qiagen, Spain) carried out in six-well plates (BD Biosciences, USA). The day before transfection 250,000 HEK293T cells were seeded in each well, and under a confluence of ~70–80% the transfection was carried out the following day. A total amount of 5µg of plasmid DNA (2.5µg KMT2a and 2.5µg AFF1) was prepared, adding 12.4ml of $CaCl_2$ (2M) in a final amount of 100µl of milliQ water; in parallel, 100µl of a 2X HBS solution was prepared and both tubes were mixed. Finally, the transfection mix was added in a drop-wise manner to the cells. Plates were incubated at 37°C overnight in a CO_2 incubator. Next day the medium was changed followed by 48hrs of incubation. After incubation, individual cells were sorted into each well of a 96-well plate with a FACSAria II flow cytometer (BD Biosciences, San Jose, CA). Clones were expanded into larger plates as the individual clones reached confluence, with DNA isolation occurring when cells reached confluence in a 6-well plate (~1 x 10^6 cells/mL).

Translocation PCR and Sequencing

Cellular gDNA was isolated from pellets of 1-2 x 10^6 HEK293 cells using the Qiagen DNEasy Blood and Tissue Kit (Cat. ID 69506, Valencia, CA). PCR was performed using AmpliTaq Gold Fast PCR Master Mix, UP (2X) (Applied

Biosystems, Foster City, CA) on isolated gDNA. Amplification parameters were optimized for both translocation outcomes. MLL-AF4 amplicon size of 944bp (FWD Primer: 5'- CTAAAGTAGTCGTTGCCAGCATCTGACTG -3', REV Primer: 5'-GACGAGTAATGTAAGTGATGCCGGCTTTC-3') and AF4-MLL amplicon size 764bp (FWD Primer: 5'- CAGGGAACCACTTAGGGTTTGAAAG -3', REV Primer: 5'- TTTGGGATGGAGTCTAGCTCTGTTGTCC-3'). Primers were obtained from Integrated DNA Technologies (Coralville, IA). Amplicon size was verified on 1% agarose gel, and PCR products were verified by DNA sequencing by Eton Bio Incorporated (Union, NJ).

RT-PCR

For RT-PCR, total cellular RNA was isolated using TRIzol Reagent from Thermo-Fisher Scientific (Cat. ID103 15596026, Waltham, MA). cDNA was synthesized from total RNA using High Capacity RNA-to-cDNA Kit from Applied Biosystems (Cat. ID, Foster City, CA) following manufacturer's protocol. Purified cDNA was then amplified using Amplitaq Gold Fast PCR Master Mix (Thermo-Fisher, Waltham, MA), optimized for MLL-AF4 cDNA (FWD RT Primer: 5'-GAGTGAAGAAGGGAATGTCTCG -3', REV RT Primer: 5'-GACTGTGGAGCACTTGGAGGT -3', obtained from Integrated DNA Technologies, Coralville, IA). Amplicon size was verified on 1% agarose gel, and PCR products were verified by DNA sequencing by Eton Bio Incorporated (Union, NJ).

Results

Can CRISPR/Cas9 system be implemented for the generation of chromosomal translocations?

Previous studies have identified hot spots within the MLL and AF4 genes, known as breakpoint cluster regions, that are frequently the sites of chromosomal translocations in patients with leukemia [125]. To design CRISPR sgRNAs for the MLL and AF4 genes, the (4:11)(q21:q23) translocation present in the MV411 human cell line was chosen as a model [126,127]. MV411 are often used in human acute monocytic leukemia studies since this cell line was established from a 10-year-old boy with acute monocytic leukemia (AML FAB M5) at diagnosis[128]. The translocation in this cell line can be representative of the most common site of translocations in patients with infant leukemia [129]. SgRNAs were designed targeting intron 8 of MLL (KMT2A) gene and intron 4 of AF4 (AFF1) gene following published guidelines and protocols[54] (see Figure 34). The sgRNAs were co-expressed with Cas9 in HEK 293T cells by Calcium Phosphate transfection. The CRISPR/Cas9 targeting system was harbored in an expression vector which also contained a wild-type eGFP gene so that HEK 293T cells successfully transfected could be discernible by FACS. The cells exhibiting eGFP expression were sorted using a FACS Aria II and placed into individual wells for single cell clonal expansion. Cells were allowed to expand for approximately 3–5 weeks at which time DNA was extracted from the clonal expansions (see Figure 35). Resultant PCR products were isolated, and genomic rearrangements confirmed by Sanger sequencing using primers spanning the expected breakpoint junction. A translocation occurring in cells expressing both MLL and AF4 sgRNAs was detected. Sequencing of breakpoints confirmed the formation of the MLL-AF4 translocation event and junction types resulting from both containing short

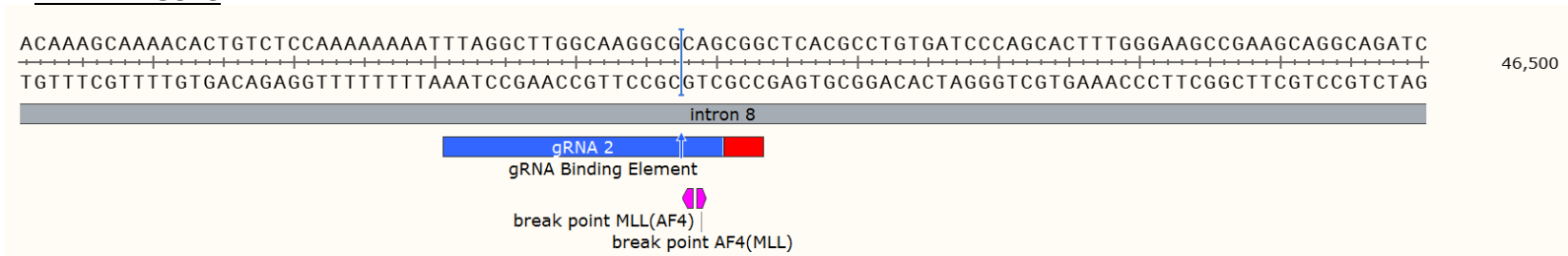
deletions and insertion that likely resulted from nucleolytic processing of DNA ends during DSB repair were observed (see Figure 36 A-C). In addition, expression of the predicted MLL-AF4 fusion transcript was detected from the cDNA samples using primers spanning the junction between MLL exon 8 and AF4 exon 5 (see Figure 36 D).

Discussion

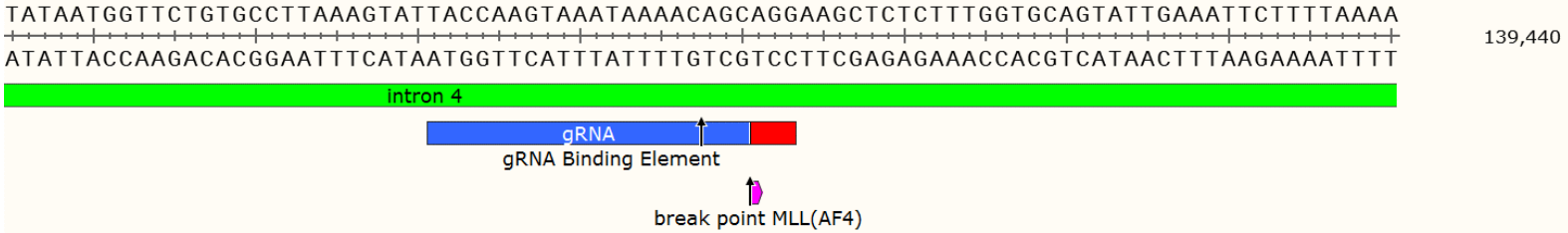
Many cancers carry recurrent chromosome translocations, which often result in the formation of fusion genes that are directly involved in the tumorigenic process [130,131]. For the precise modeling of human recurrent chromosome translocations and their impact on disease development the reenactment of the actual translocation in cells or in mice would be the closest possible recapitulation of the sequence of events in humans. Until now such reenactment was a daunting task as the translocation would require introduction of LoxP or Frt recombination sites into both translocation partners via homologous recombination, followed by expression of Cre or Flp recombinase to create DSBs that would mediate the translocation [132–134]. As shown by others [135–138] and here, the availability of the CRISPR-Cas9 system has paved the way to implementing this approach without such major technical or time investment. For translocation to occur both genes need to break and the disparate ends need to fuse via a process called non-homologous end joining [139]. The nature of an induced fusion site is also likely to depend on a variety of other factors, such as features specific to the genomic regions that are targeted, the unique properties of the cell lines that are used, and the levels of Cas9/sgRNAs that are expressed [138]. The ease with which sgRNAs can be designed and cloned makes the CRISPR system well-suited for high-throughput screening experiments. Also, the ability to multiplex Cas9 targeting opens

up the possibility of studying more complex combinations of rearrangements observed in cancer genomes. This study proposed the use of CRISPR-Cas9 technology to provide a convenient platform for rapid modeling of cancer-related genetic mutations in vitro. The creation of this cell line will facilitate cancer research in both mechanism investigations and therapy testing. Although our system currently only produced the unbalanced translocation product, which is present this way in some patient samples, this system is believed to have substantial extent to recapitulate in vitro the genetic and biological complexity of cancer. This will allow us to test the current and new drug therapies on patient mutation to further understand the effect of these specific mutations in the drugs mechanism of action and help us develop better drug screening methods.

A **KMT2A Gene**



AFF1 Gene



B

Chromosome	Gene	Strand Target	CRISPR Binding Element(5′-3)	GUIDE OLIGOs
4	AFF1	Top	**TACCAAGTAAATAAAACAGCAGG** ATGGTTCATTTATTTTGTCGTCC	**5′CACC G TACCAAGTAAATAAAACAGC** **3′ C ATGGTTCATTTATTTTGTCG CAAA**
11	KMT2A	Top	**TTAGGCTTGGCAAGGCGCAGCGG** AATCCGAACCGTTCCGCGTCGCC	**5′ CACC G TTAGGCTTGGCAAGGCGCAG** **3′ C AATCCGAACCGTTCCGCGTC CAAA**

C

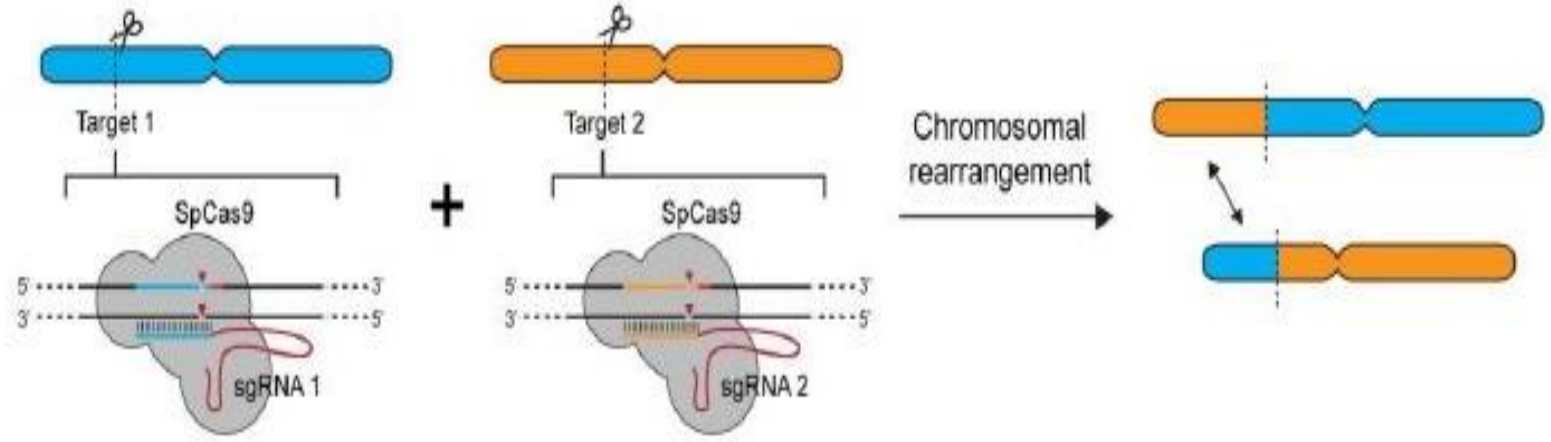


Figure 34. CRISPR/Cas9 Mediated Translocation.
(A-B,D.) The translocation strategy. CRISPR/Cas9 constructs were designed and built following published guidelines and protocols DSBs are introduced by the sgRNAs (arrowheads) mapping to introns in KMT2A (gray) and AFF1 (Green). Double-strand breaks in the targeted loci were generated by transfection of HEK293A cells with a plasmid expressing the Cas9 gene and sgRNA. (C) The custom CRISPR/Cas9 RNA guide sequence is depicted in the box.

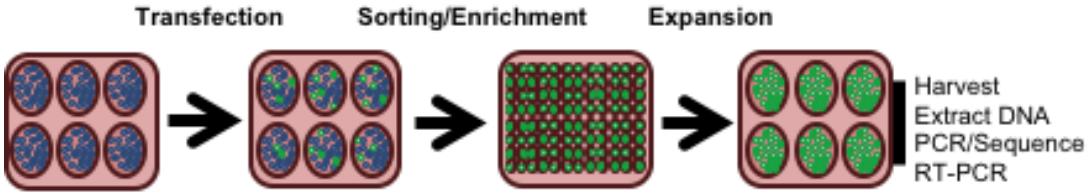
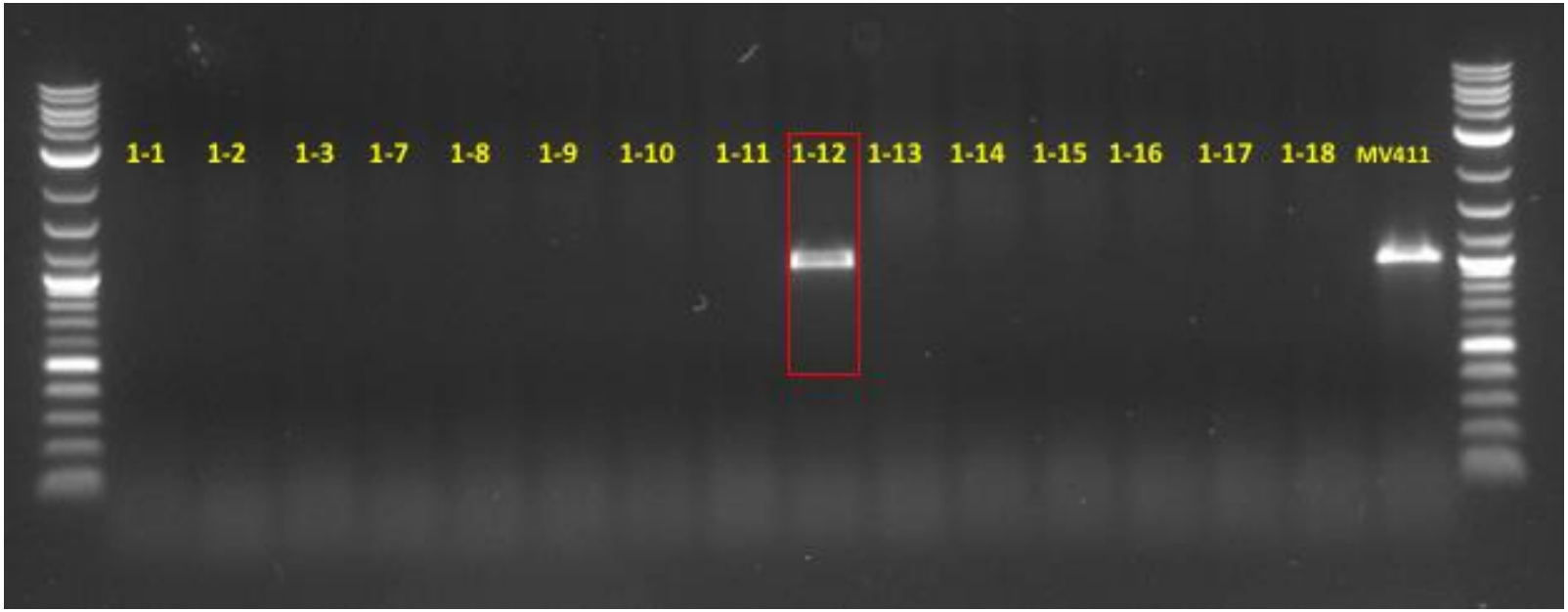
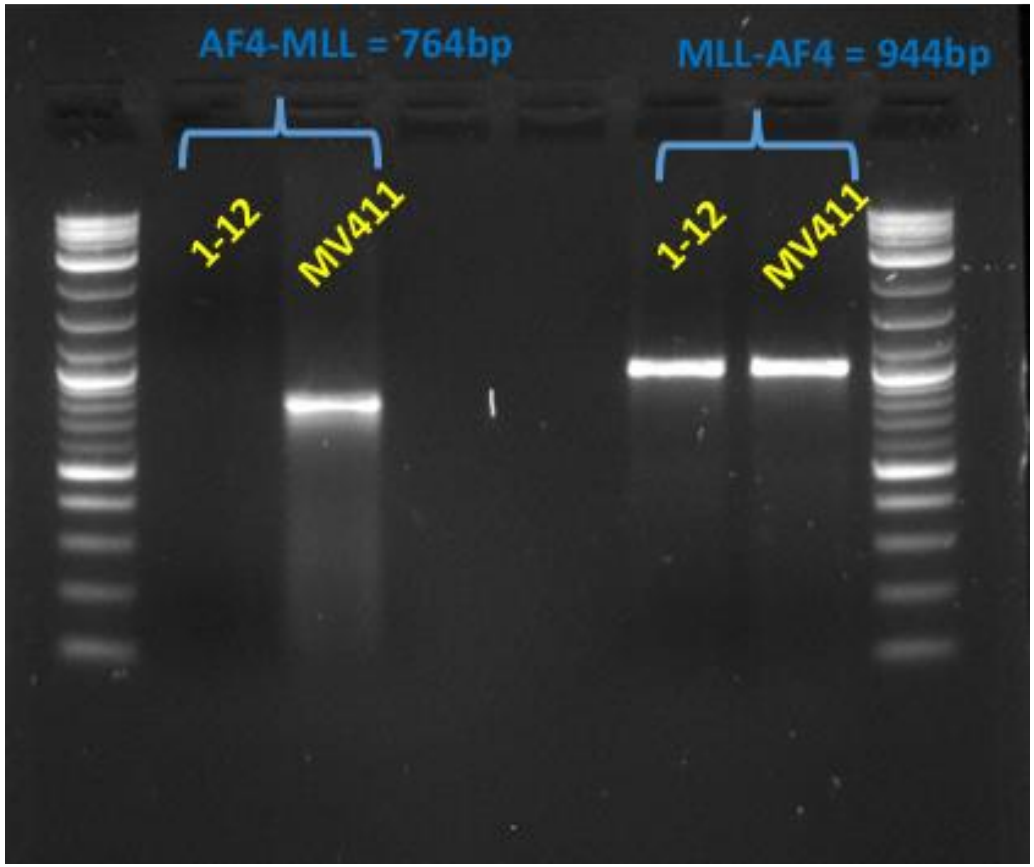


Figure 35. Experimental Overview.
Hek293A cells were transfected using Calcium Phosphate at a concentration of $5x10^5$ cells/100ul with 2.5ug each of the indicated CRISPR/Cas9 KMT2A and AFF1. Following transfection, cells were allowed to incubate for 48 hours followed by single sorting and clonal expansion.

A



B



C

D

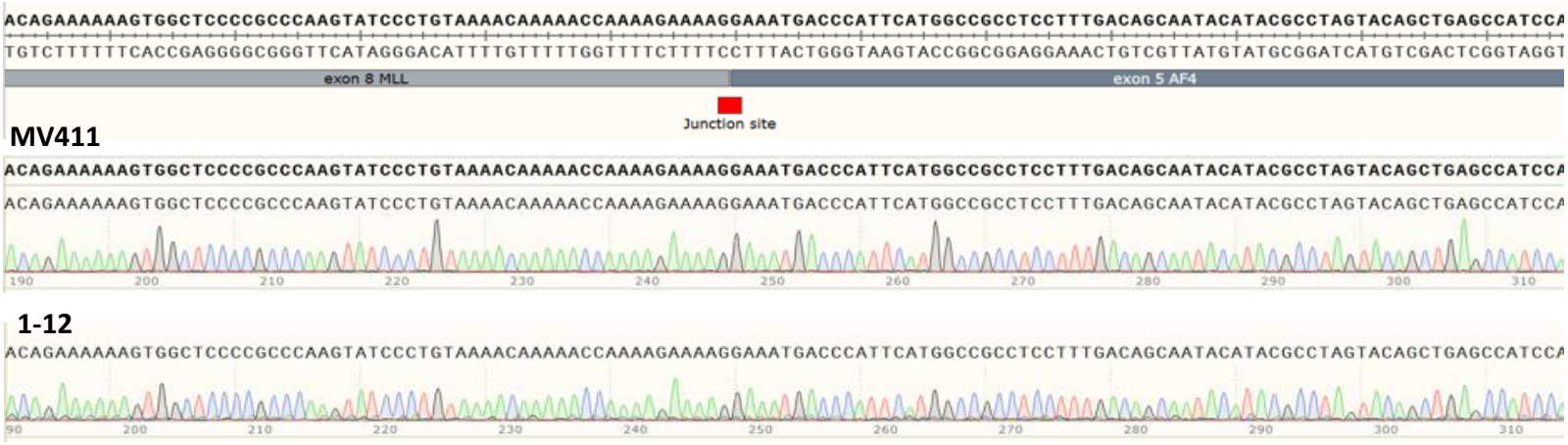


Figure 36. Identifying translocation breakpoint junctions.
(A) PCR analysis of single cell expanded clones to identify translocation breakpoint junction. (B) PCR analysis of clone 1-12 for the secondary translocation AF4-MLL. (C) Sequence chromatogram of the detected MLL-AF4 breakpoint junction from cells in which Cas9 and both KMT2A and AFF1 sgRNAs were expressed. (D) RT-PCR sequence chromatogram of the detected MLL-AF4 breakpoint junction transcript.

Chapter 7

MODELING PEDIATRIC AML FLT3 MUTATIONS WITH CRISPR-Cas12a MEDIATED GENE EDITING

Introduction

AML is a disorder of the hematopoietic progenitor cell wherein cells stall during differentiation followed by uncontrolled growth[140], through mutations found in FLT3, NPM1, CEBPA, RAS, c-KIT, and WT. These mutations affect many cellular activities, specifically cell proliferation and cell survival[141,142]. Pediatric AML is a rare disease with only ~500 children diagnosed each year (stjude.org). While prognosis has improved over the last few decades[143], relapse is still a major concern and it accounts for more than half of the deaths from pediatric leukemia [140,143]. The primary treatment for AML is chemotherapy, sometimes coupled to targeted therapy or stem cell transplant[143]. One of the key genes involved in AML pathogenesis, FLT3, is located on chromosome 13q12 and encodes a class III receptor tyrosine kinase that regulates hematopoiesis. This receptor is activated by binding the fms-related tyrosine kinase 3 ligand (FL) to its extracellular domain, which induces homodimer formation in the plasma membrane and to auto-phosphorylation[144]. The activated receptor kinase subsequently phosphorylates and activates multiple cytoplasmic effector molecules in pathways involved in apoptosis, proliferation, and differentiation of hematopoietic cells[144]. Mutations that lead to the constitutive activation of this receptor result in acute myeloid leukemia and acute lymphoblastic leukemia [144–147]. FLT3 is the most common gene mutated AML [148] and the two

main FLT3 mutations are internal tandem duplications (ITD) and point mutations. FLT3 ITDs range from ten to several hundred bases and are repeated within exon 14, which encodes the juxta-membrane domain (JDM)[149,150]. All duplications are "in frame," with the number of nucleotides added in multiples of three; the overall reading frame of the protein remains unchanged. Point mutations have been found in the tyrosine kinase encoding domain (TKD) of the FLT3 protein in 5-7% of AML patients[151]. FLT3 ITD and FLT3 TKD mutations induce aberrant activation of FLT3 signaling, leading to proliferation [146]. AML bearing an FLT3-ITD usually presents with high blood blast counts and a normal karyotype; it has poor treatment outcomes. Libura et al.[152] found that 54% of AML, exhibiting either a MLL-PTD or a breakage of the MLL gene at the Topo II (Topoisomerase II) site, carry either an internal tandem (FLT3-ITD) or a variety of point mutations. Protein kinase activation can be induced by somatic mutation, a common mechanism of tumorigenesis. There is an initial response to treatment, but high relapse rate, short relapse-free survival (RFS) and diminished overall survival often follow [143,149].

Tyrosine kinase inhibitors (TKI) are effective in the targeted treatment of various malignancies[153] through competitive ATP inhibition at the catalytic binding site of tyrosine kinase[154]. This inhibition helps regulate the growth of some types of cancer cells by blocking signaling for cell growth and division[147]. TKIs represent anticancer drugs that are designed to interfere with a specific molecular target, usually a protein with a critical role in tumor growth or progression (i.e.FLT3)[155]. Several FLT3 kinase inhibitors are now in development and although some have shown promise in clinical arenas, responses tend to be transient. FLT3 inhibitors are classified into first- and second-generation based on their specificity for FLT3, and

into type I and type II categories based on their interaction with FLT3. First-generation inhibitors block multiple RTKs (receptor tyrosine kinases). A lack of specificity for a single RTK may enhance anti-leukemia efficacy by inhibiting targets downstream of FLT3 and/or in parallel signaling pathways, or other targets in AML cells. Second-generation FLT3 inhibitors often target FLT3, and do not have efficacy against targets downstream or on parallel signaling pathways. Type I inhibitors bind to the ATP-binding site when the receptor is active, while type II inhibitors interact with a hydrophobic region immediately adjacent to the ATP-binding, only accessible when the receptor is in the inactive conformation thereby preventing receptor activation. Sensitivity of the different FLT3 mutations to the diverse FLT3 inhibitors has been reported to be highly variable[144,156].

Advancement in the treatment of leukemia over the last 40 years has impacted and transformed uniformly fatal disease into a one that is somewhat manageable. There are, however, a number of subtypes of pediatric and adult leukemia that evade treatment and continue to present poor prognosis, many of these involve FLT3 mutations. For example, the FLT3 ITD associated with a single point mutation in the tyrosine kinase domain is known to induce resistance to TKI treatment. Some patients bear polyclonal cells that, when exposed to TKI, are selected and as such, relapse in both adult and pediatric leukemia cases. The major difference between the pediatric and adult AML is the higher proportion of the children with activating point mutations within the TKDs of the FLT3 gene [157]. Thus, more relevant cell-based models are needed to enable the discovery of new therapeutic targets through the screening of anti-leukemic drugs designed to reverse the negative prognosis encountered in infants and adults with leukemia.

In here I describe newly developed and innovative usage of the Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/(with the associated nuclease, Cas12a) gene editing tool functioning as fully assembled ribonucleotide protein (RNP) complex to engineer FLT3 ITD mutations. CRISPR/Cas9 has emerged as the preferred agent to catalyze site specific double stranded DNA cleavage, generating a template for non-homologous end joining [12,13,54–57,138,158] A major advantage in using CRISPR systems for genome editing is the simplicity with which the vectors expressing the CRISPR components can be created and utilized. The innovation in this project centers on the use of a fully synthetic ribonucleoprotein (RNP) with an alternative CRISPR/ Cas system [159]. Cas12a associates with crRNA and is able to fully engage the DNA target site without the requirement of a second tracrRNA[160]. Here, I use Cas12a to carry out gene editing reactions on plasmid DNA templates bearing the FLT3 gene *in vitro*. This *in vitro* system affords us the opportunity to carry out biomedical engineering and *in vitro* site-directed mutagenesis in a controlled environment that does not require PCR and is easily scalable. By implementing this Cas12a directed *in vitro* cell-free gene editing system, we re-create the patient specific FLT3 mutations in expression plasmids that can be further characterized for involvement in AML leukemogenesis, specifically in the context of differential sensitivity to FLT3 TKIs. In summary, we aim to (1) construct the expression plasmids that will express the FLT3 ITD along with a point mutation by implementing in vitro gene editing; (2) characterize the effect of both FLT3 ITD and point mutations (singly or in combination) at a cellular level and how these play a role in AML progression, and last (3) validate our system by showing by showing differential sensitivity of FLT3 mutants to FLT3 TKI inhibitors. Characterizing the effect of the FLT3 TKI on

proliferation and survival of the FLT3 mutant clones will give us a better understanding of how drug resistance evolves in some patient.

Methods and materials

CRISPR-Directed In Vitro Gene Editing

Cell-free extracts were prepared following the technique outlined by Sansbury et al.[161]. RNP complexes used in in vitro reactions consisted of a purified AsCas12a nuclease (Integrated DNA Technologies, Coralville, IA) and a target-specific crRNA (Integrated DNA Technologies). In vitro DNA cleavage reaction mixtures contained 250 ng (0.007566 lM) of FLT3 plasmid DNA (Gift from Tarlock Lab) and 10pmol of RNP mixed in a reaction buffer (100mM of NaCl, 20mM of Tris-HCl, 10mM of MgCl2, and 100µg/mL of bovine serum albumin), which was brought to a final volume of 20µL. Each reaction was incubated for 15 min at 37° C after which DNA was recovered from reaction mixtures and purified using QIAprep Spin Miniprep silica columns (Qiagen, Hilden, Germany). Secondary in vitro recircularization reactions contained DNA recovered from the initial cleavage reaction, 20µg of cell-free extract supplemented with Quick Ligase (New England Biolabs, Ipswich, MA), and a reaction buffer (20mM of Tris, 15mM of MgCl2, 0.4mM of DTT, and 1.0mM of adenosine triphosphate), which was brought to a final volume of 35µL. Each secondary reaction was incubated for 15 min at 37° C. Double-stranded donor DNA templates (Integrated DNA Technologies), 4.464µg was added into the secondary reaction mixture. DNA from the secondary in vitro recircularization reactions was recovered from reaction mixtures and purified using silica spin columns. Plasmid

DNA recovered from in vitro reactions was transformed into 50µL of DH5a competent E. coli (Invitrogen, Carlsbad, CA) via heat shock transformation. Competent cells were incubated on ice for 30min after plasmid introduction, heat shocked for 20 s at 42°C, placed on ice for 2min, brought to a final volume of 1mL in SOC media and incubated for 1 hr at 37°C, with shaking (225 rpm). Undiluted competent cells were plated on media containing ampicillin antibiotics and incubated overnight at 37°C. Single ampicillin-resistant colonies were selected, and plasmid DNA was isolated via a QIAprep Spin Mini- prep Kit (Qiagen). Modifications made to the plasmid DNA selected from bacterial colonies were evaluated via DNA sequencing (GeneWiz, South Plainfield, NJ).

Cell lines and transfection

The murine IL-3-dependent pro-B cell line, Ba/F3, was maintained in RPMI 1640 medium (GIBCO, Rockville, MD, USA) supplemented with 10% fetal bovine serum (FBS, Gemini Bio- Products, Calabasas, CA, USA) and 1 ng/ml IL-3. Ba/F3 cells were transfected with plasmid DNA by Nucleofection using the SG Cell Line 4D-Nucleofector X Kit. BA/F3 cells were transfected with program CM-147 in 20 µl cuvette Strips. Transfected cells were cultured in IL-3-containing medium for 72hr and then selected in 1 mg/ml Puromycin (GIBCO) for a period of 7 days. The ability of the cells to survive in the absence of IL- 3 was determined by trypan blue exclusion.

Cell proliferation of Ba/F3 cells

Cells were seeded at a density of $4x10^4$/mL in the presence or absence of IL-3 as indicated. Viable cells were counted at indicated time periods in a standard hemacytometer after staining with trypan blue.

Cytotoxicity

Ba/F3 or Ba/F3-derived cells were plated in quadruplicate at 1×10^5 cells per well in 96-well plates with or without IL-3 supplement and in the presence of increasing concentrations of Sorafenib or Gilteritinib for 72 hours. To evaluate the cytotoxic effects of Sorafenib and Gilteritinib, an MTS-based assay system (Boehringer Mannheim, Indianapolis, IN, USA) was used. Sorafenib (S7397) and Gilteritinib (ASP2215) were purchased from sellecckchem.com.

RT-PCR

For RT-PCR, total cellular RNA was isolated using TRIzol Reagent from Thermo-Fisher Scientific (Cat. ID103 15596026, Waltham, MA). cDNA was synthesized from total RNA using High Capacity RNA-to-cDNA Kit from Applied Biosystems (Cat. ID, Foster City, CA) following manufacturer's protocol. Purified cDNA was then amplified using Amplitaq Gold Fast PCR Master Mix (Thermo-Fisher, Waltham, MA), optimized for FLT3 cDNA (FWD RT Primer: 5'-GAGTGAAGAAGGGAATGTCTCG -3', REV RT Primer: 5'-GACTGTGGAGCACTTGGAGGT -3', obtained from Integrated DNA Technologies, Coralville, IA). Amplicon size was verified on 1% agarose gel, and PCR products were verified by DNA sequencing by Eton Bio Incorporated (Union, NJ).

Immunofluorescent Staining and Confocal Microscopy Imaging

Baf/3 FLT3 expressing cells and un-transfected cells were fixed in 4% formaldehyde for 15 min at room temperature followed by rinsing three times in PBS. For plasma membrane staining 2.0 $x10^6$ cells/ mL were incubated with 50µg/mL of WGA-Alexa 647 (pink) for 10 min at room temperature followed by rinsing two times in PBS. Cells were then permeabilized with 0.3% Triton 100-X for 10 min at room

temperature followed by rinsing two times in PBS. Blocking of non-specific protein-protein interaction was performed in 5% Goat Serum and 0.1% Triton 100-X for 2hr at room temperature followed by rinsing two times in PBS. To satin the ER cells were then stained in 125ug/mL of Concanavalin A- Alexa 594 (Red) for 30min at room temperature followed by rinsing two times in PBS. FLT3 protein was stained by 10ug/mL of Rabbit anti-FLT3 (ab37847) overnight at 4ºC followed by washing three times in 1% PBS. The secondary antibody was Goat anti-Rabbit IgG(H+L) Alexa Fluor 488 (green) conjugate was used at a 1/1000 dilution for 1 hr. Fixed and stained cells were mounted on to slides for imaging with ProLong™ Diamond Antifade Mountant with DAPI. The treated cell samples were investigated under a ZEISS LSM880 Laser Scanning Confocal Microscope at DBI. A 63x oil objective lens and four channel light sources (bright field transmitted light, 405nm, 488nm, and 561nm lasers) were used.

Protein Structural Analysis

Protein multiple sequence alignment was performed by utilizing the online webtool Clustal Omega: https://www.ebi.ac.uk/Tools/msa/clustalo/ [162,163]. The Phyre2 web portal was used for the modeling, prediction and analysis of all FLT3 proteins and mutant variants (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) [164].

Flow Cytometry

Baf/3 FLT3 expressing cells and un-transfected cells were fixed in 4% formaldehyde for 15 min at room temperature followed by rinsing three times in PBS. Cells were then permeabilized with 0.3% Triton 100-X for 10 min at room

temperature followed by rinsing two times in PBS. Blocking of non-specific protein-protein interaction was performed in 5% Goat Serum and 0.1% Triton 100-X for 2hr at room temperature followed by rinsing two times in PBS. FLT3 protein was stained by 10ug/mL of Rabbit anti-FLT3 (ab37847) overnight at 4ºC followed by washing three times in 1% PBS. The secondary antibody was Goat anti-Rabbit IgG(H+L) Alexa Fluor 488 (green) conjugate was used at a 1/1000 dilution for 1 hr. Samples were run at the Center for Translational Cancer Research (CTCR) Flow Cytometry Core Facility at the Helen F. Graham Cancer Center & Research Institute (HFGCC&RI), using the BD FACSAria II flow cytometer to measure the Mean Fluorescence Intensity (MFI).

Results

Can Cas12a mediated gene editing be implemented to develop a model system to study pediatric AML FLT3 mutations to evaluate the progression of oncogenesis and efficacy of novel AML drugs?

By implementing the Cas12a directed *in vitro* cell-free gene editing system, we intend to re-create patient specific FLT3 mutations in expression plasmids that can be further characterized for involvement in AML leukemogenesis, specifically in the context of differential sensitivity to FLT3 TKIs.

The construction of the plasmids expressing the FLT3-ITD gene bearing different point mutations was carried out using a novel *in vitro* cell-free gene editing system as site-directed mutagenesis approach (Figure 37) [161]. For these experiments FLT3-ITD sequences, designated Patient 4 37bp ITD (P4) and the Patient 576bp ITD (P5) respectively, identified in pediatric AML patients, will be examined (Figure 39) [165]. These point mutations identified D835Y and F691L, were found in patients

after FLT3 inhibitor therapy [166,167] were combined with the P4 and P5 FLT3 ITD elements to create expressions plasmids that would not only express the FLT3 ITD of interest but also one of the selected point mutations (Figure 39). Given their proximity in Exon 14 of the FLT3 gene, the P4 ITD (37bp) and P5 ITD (76bp), replacement reactions were initiated by a double cut reaction of Cas12a enzyme. These cut sites produced a 108bp fragment that is then replaced by a 145bp fragment to produce P4 ITD (37bp) and a 184 fragment to produce the P5 ITD (76bp) (Figure 38). *In vitro* DNA cleavage reaction mixtures contained the pMX-Puro plasmid bearing the FLT3wt cDNA (Figure 40) and the Cas12a RNP (see Methods and Materials). The DNA recovered from the initial cleavage reaction were mixed with cell-free extract and duplexed DNA replacement fragment containing the ITD or PM sequence. Modified plasmid DNA was transformed into DH5α competent *E. coli* via heat shock transformation and plated on media containing ampicillin antibiotics and incubated overnight. Single ampicillin resistant colonies were selected and plasmid DNA was evaluated for the specific modifications to the FLT3wt cDNA (Figure 41). Six different FLT3 mutation expression plasmids (P5, P4, D835Y, P5/D835Y, P4/D835Y, and P4/F691L) were successfully created with this method. These FLT3 mutations were then characterized for their transforming leukemic capacity.

The length of FLT3/ITD may influence FLT3 activation by altering its structure and impacting the response to therapy [168,169]. To identify oncogenes involved in AML, specific FLT3, functional genetic screens in myeloid cells have been established[170]. The murine pro-B Ba/F3 cell line and 32D myeloid progenitor cells depend on IL-3 for growth and viability and in the absence of IL-3, these cells undergo cell cycle arrest and apoptosis[170]. When transformation is achieved, the

cells lose their IL-3 dependence[171], making IL-3 growth independence an appropriate measure of transformation. The methodology of plasmid transfection containing FLT3 ITD or TK point mutations into cytokine-dependent hematopoietic cell lines (Ba/F3 and 32D) have already been established[148,171–175].

For this study, the murine IL-3-dependent pro-B cell line, Ba/F3, was used. Plasmids containing FLT3 ITD with or without the point mutations (P5, P4, D835Y, P5/D835Y, P4/D835Y, and P4/F691L) were transfected into Ba/F3 cells by nucleofection. To determine the efficiency of transformation, the mutated FLT3 cells were grown in the presence of IL-3 for three days and assayed for the following 27 days in the absence of IL-3 (determined with the Trypan Blue Exclusion test (Figure 42)). Throughout this study we use a "Seattle patient" FLT3 ITD (24bp) expression plasmid (gift from the Katherine G. Tarlock Lab, Department of Hematology-Oncology at Seattle Children's Hospital), as a positive transformation control. As seen in Figure 42, P5, P4, P5/D835Y, P4/D835Y, and P4/F691L took approximately 25 days to exhibit transformation. The D835Y mutation exhibited sustained transformation by 21 days. Un-transfected (NT) BaF/3 cells showed growth in the presence of IL-3 which was not sustained after 3 days without IL-3. These data helped confirm that transfected cells, surviving IL-3 depletion, have undergone transformation. The transformed Ba/F3 cells were selected in puromycin for 6 days to ensure the transformation was due to integration of the FLT3 cDNA expression cassette (Figure 40). Cell growth and viability were again determined by the Trypan blue exclusion (Figure 43). The transformed cells showed growth in the presence of Puromycin. Un-transfected BaF/3 cells cultured in the presence or absence of IL-3 were also put under puromycin selection and did not survive.

The Human MV4-11 cell line [128] is commonly used for AML FLT3 assays [176–178] because it constitutively expresses FLT3. We confirmed mRNA expression by RT-PCR followed by sequencing the cDNA of the transformed cell lines with the P5, P4, D835Y, P5/D835Y, P4/D835Y, and P4/F691L plasmids (Figure 44). The RT-PCR products were visualized by gel electrophoresis and he difference in band size due to the different sizes of the ITD were distinguished among the samples. The P5 band is the highest (817bp) since it contains the longest ITD, 76 bp ITD. The Wt FLT3 and the D835Y have the same band molecular weight (739bp) since the difference between the two is a single base point mutation. The Seattle patient (24 ITD) and P4 (37 bp ITD) only have a difference in size of 13 bp giving them very similar molecular weight bands at 765bp and 778bp respectively. All RT-PCR products were sequenced to confirm the presence of the different FLT3 mutations of interest. The sequence results were aligned to the FLT3 WT gene sequence with the triangle depicting ITDs (Figure 44).

Sequence-dependent protein conformational changes FLT3-ITD lead to autophosphorylation and display a constitutive intracellular localization by aberrant trafficking of the receptor the endoplasmic reticulum (ER) along with receptors in the plasma membrane [176,179–181]. To characterize our assay AML associated patient FLT3 mutations, the subcellular localization of the mutated proteins was established (Figure 45). Two main subcellular localization compartments of FLT3 (anti FLT3-488, green) the plasma membrane were stained with WGA-Alexa 647(pink-like color stain) and the ER stained with ConA-Alexa 594 (Red color stain). Un-transfected Baf/3 cells were used as control to demonstrate that no FLT3 is expressed in these cells. In accordance with published data, all FLT3 ITD mutations (24 ITD, P4, P5 and

MV4-11(30 ITD)) and ITD + Point mutations (P4/D835Y, P4/F691L and P5/D835Y) characterized in this study show co-localization; yellow pigment of FLT3 with the ER *and* plasma membrane. This was further tested by flow cytometry providing the MFI (Mean Fluorescent Intensity) ratio which quantifies the expression of the different FLT3 variants (Figure 46).

To determine the effect of the FLT3 mutations on their sensitivity to TKI cells expressing these FLT3 mutations (P5, P4, D835Y, P5/D835Y, P4/D835Y, and P4/F691L) were evaluated against first generation and second generation FLT3 inhibitors, Sorafenib and Gilterintinib respectively for their TKI sensitivity. Cell growth was calculated by the quantification of viable cells in proliferation through the MTS assay. Based on Trypan Blue exclusion viable cells were seeded in quadruplicate in a 96-well microtiter plate in increasing concentrations (0-20 nM) of inhibitor for 72 h. Following treatment, cells were incubated for 4 h in MTS/PMS solution followed by recording the absorbance (Figure 47). The concentration of drug that inhibited absorbance by 50% (IC_{50}) was calculated and is displayed in Table 7. For both TKIs, a variability in the IC_{50} can be observed where the difference in mutations resulted in a variable range of sensitivity. The P5, P4, D835Y and P4/F691L mutations displayed higher sensitivity to the Type 1 inhibitor, Gilterintinib because of lower IC_{50} concentrations as compared to the Type 2 inhibitor, Sorafenib. Interestingly Wt FLT3, MV4-11(30bp ITD), P4/D835Y and P5/D835Y reveal the same IC_{50} concentrations for both inhibitors. ITD length has been a studied parameter to determine clinical prognostic outcome of AML patients, in which the longer the ITD the lower the favorable prognostic [169,182]. This same parameter proved true in this study system; the P4 (37bp ITD) and P5 (76bp ITD) showed low sensitivity to the

presence of the TKI having IC_{50} concentrations above 10nM. To measure the effect of the TKI on the expression and localization of the FLT3 protein the relative expression was calculated using flow cytometry analysis (Figure 49) and fluorescence imaging (Figure 48). Interestingly FLT3 expression was increased in the presence of both TKIs when compared to the relative expression of the un-treated cells. The D835Y point mutation and the MV4-11 bearing a 30pb ITD have the most dramatic increase in FLT3 expression. FLT3 protein in the presence of TKI shows to shift the location of the protein to the plasma membrane (external localization) from the ER (internal localization).

Discussion

Inhibition of mutated FLT3 kinase activity by pharmacologic agents is an important therapeutic strategy for AML[148]. Most patients with AML and FLT3-internal tandem duplications (ITD) initially show favorable responses to FLT3 inhibitors but the development of resistance emerges with time. One of the most common mechanisms of resistance is the acquisition of mutations in the secondary FLT3 tyrosine kinase domain (TKD)[166]. For example, twelve somatic mutations have been identified at distinct positions within the juxta membrane domain (JMD), the region involved in apoptosis, proliferation and differentiation regulation. Among the JMD mutations, 9 are unique in pediatric cases all with significant activating potential (E573D/G, L576R, T582N, D586Y, Y589H, E596K/G, E598D, Y599C, D600G). These lead to auto phosphorylation of the receptor without the presence of FLT3 ligand (FL) binding[183]. Another parameter influencing FLT3 activation is the

length of FLT3/ITD, which causes structural alterations and altered responses to therapy[182].

Somatic mutations that lead to constitutive activation of FLT3 are frequent in AML patients. These mutations fall into two classes, the most common being in-frame internal tandem duplications of variable length in the juxta-membrane region that disrupt the normal regulation of the kinase activity. The juxta-membrane autoregulatory region is important for normal regulation of the kinase activity and for maintaining the kinase in an inactive state in the absence of bound ligand. Upon tyrosine phosphorylation, it mediates interaction with the SH2 domains of numerous signaling partners. Amino acid region 591-943 are important for normal regulation of the kinase activity and for maintaining the kinase in an inactive state in the absence of bound ligand. In-frame internal tandem duplications (ITDs) in the juxta-membrane domain result in constitutive activation of the kinase. The activity of the mutant kinase can be stimulated further by FLT3LG binding. Likewise, point mutations in the activation loop of the protein kinase domain, amino acid 610-943, can also result in a constitutively activated kinase. To analyze how these genomic mutations, disrupt the protein domain we have aligned the Wt FLT3 protein quaternary structure to the one of all the patient mutations discussed in this study, Figure 51. The presence of the ITD causes highly noticeable structural changes to the FLT3 protein, which account for the loss of its autoinhibitory function of the juxta-membrane domain rendering the receptor to its active conformation without the need of the FLT3LG binding.

Several FLT3-TKIs have been tested as monotherapy or combined with other chemotherapeutic agents. Sorafenib, first-generation inhibitor, is a type II FLT3 inhibitor that also inhibits RAF, VEGF receptors (VEGFR)-1,2,3, platelet-derived

growth factor receptor (PDGFR)-b, KIT, and RET. It is active against FLT3 ITD, but not against most TKD mutations, including D835 mutations [166,184,185]. This inhibitor has been FDA-approved for the treatment of renal (2005), hepatocellular (2007), and thyroid (2013) carcinomas. Gilteritinib, a second-generation inhibitor, is a small molecule type I inhibitors that is currently in clinical trials. Gilteritinib has demonstrated inhibitory activity against FLT3 ITD as well as TKD mutations. On July 2017, U.S. Food and Drug Administration (FDA) granted orphan-drug designation to Gilteritinib in patients with acute myeloid leukemia (AML)[186]. The role and timing of administration of these inhibitors in the sequence of AML therapy (induction, consolidation, maintenance) or in combination with conventional chemotherapy are still indecisive and unclear[187]. FTL3 inhibitors induce responses in AML patients with FLT3 mutations but responses are not durable, as the AML progresses in virtually all patients. FLT3 inhibitors may be administered with chemotherapy or after chemotherapy or in combinatorial regimens. Thus, the application of sensitive methods to detect and monitor leukemic clones with drug-resistant FLT3 mutations during therapy may allow individualized treatment with the currently available FLT3 inhibitors. The outcome of this study is the development of a model system to study and monitor mutant FLT3 ITD clones, more specifically, the identification of TKD-point mutations, which may be more susceptible to FLT3 inhibitors (Figure 52).

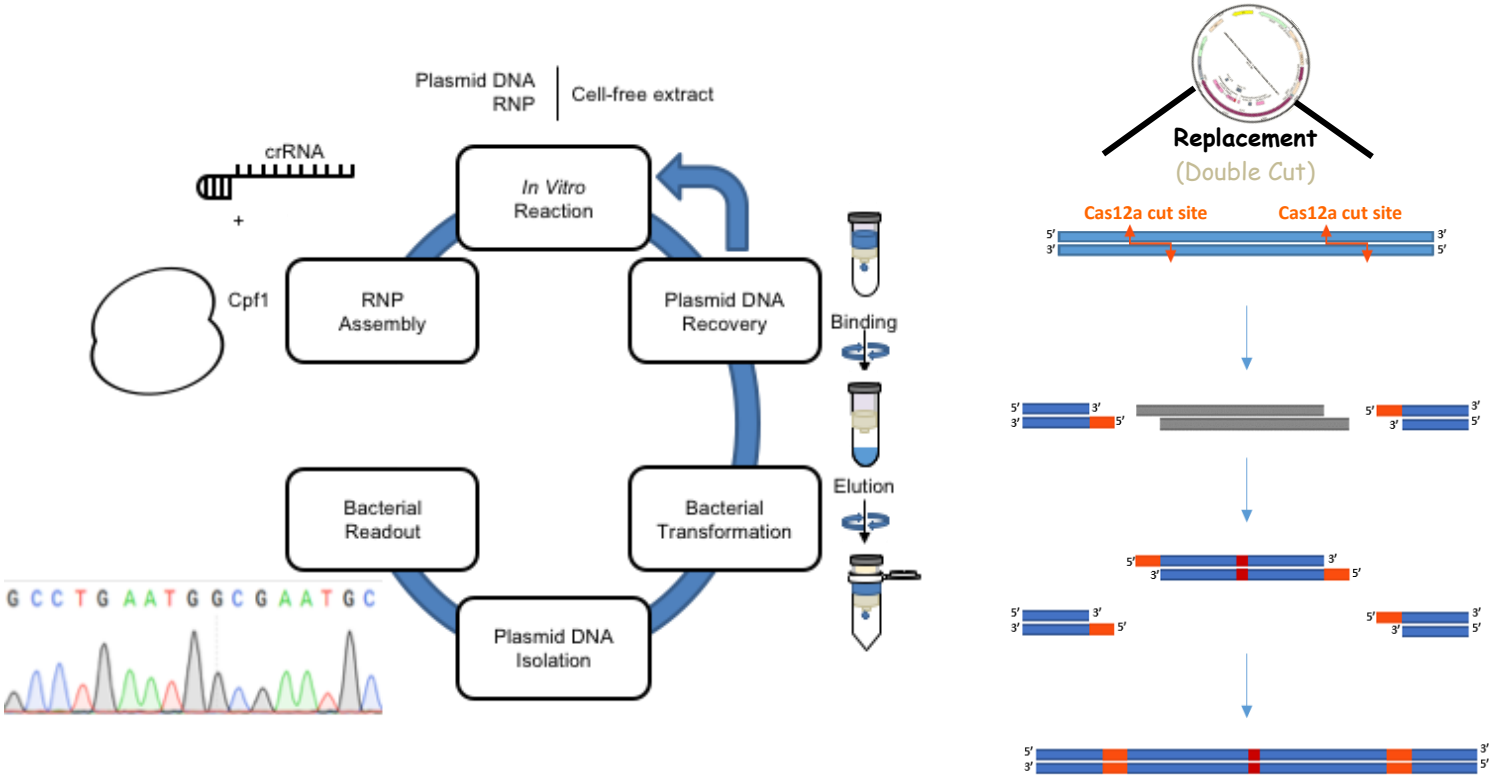
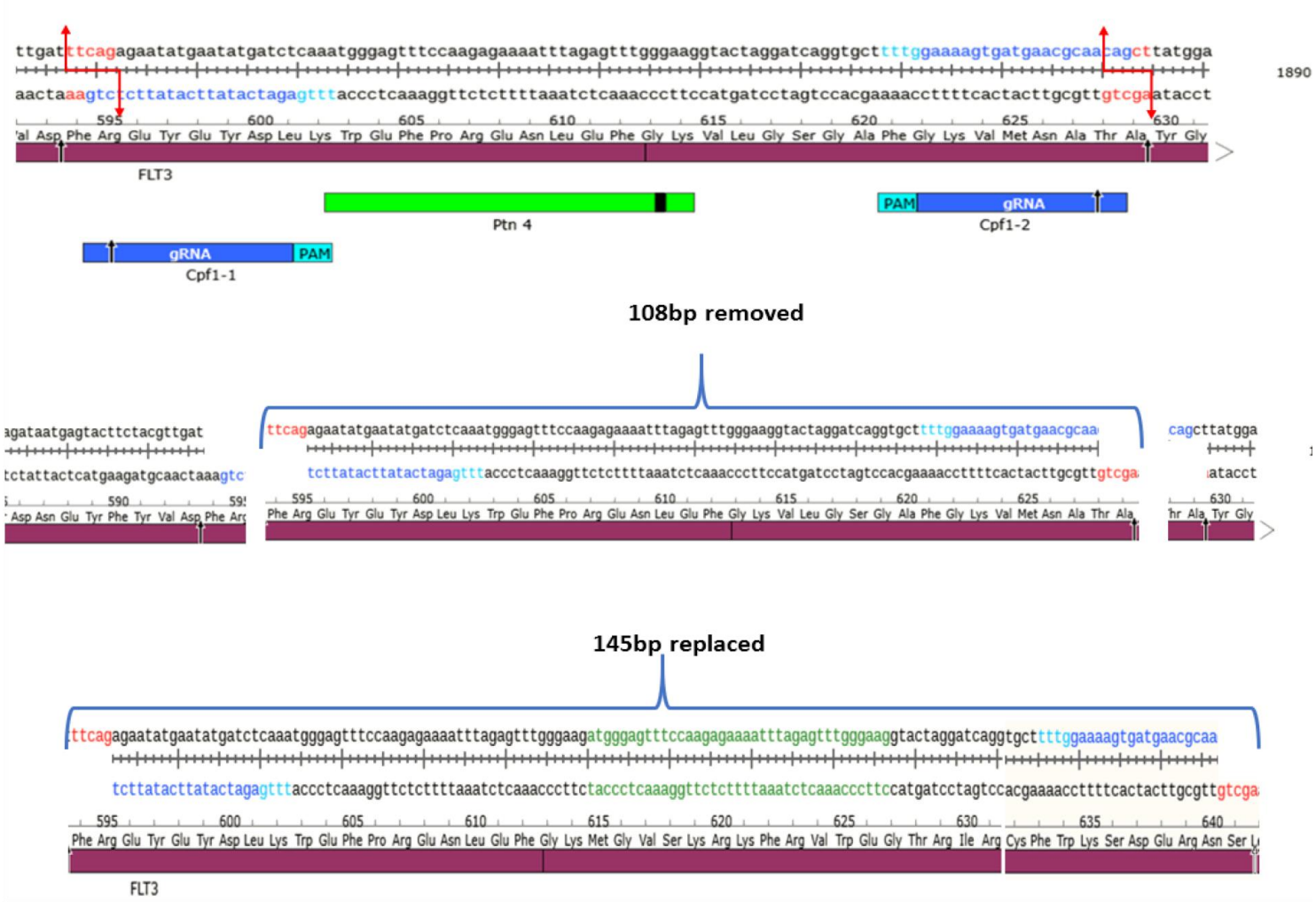
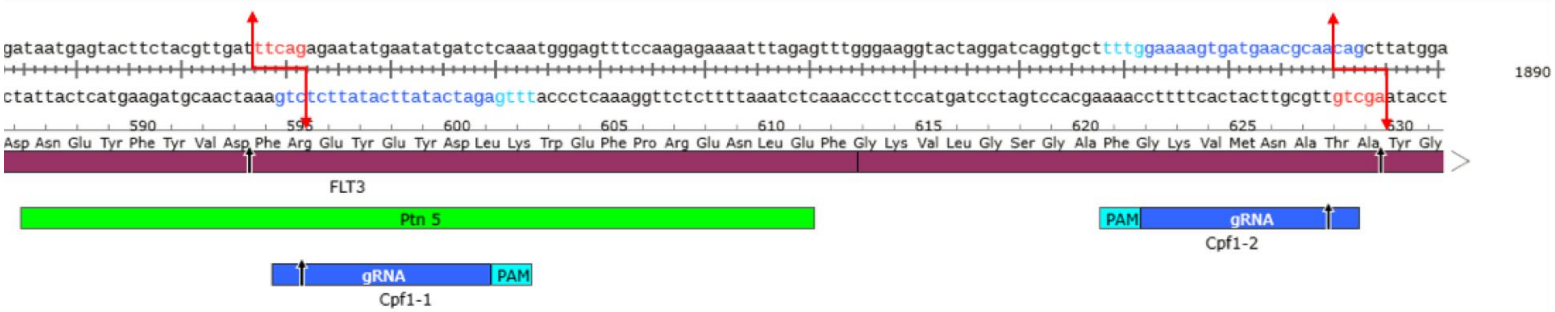
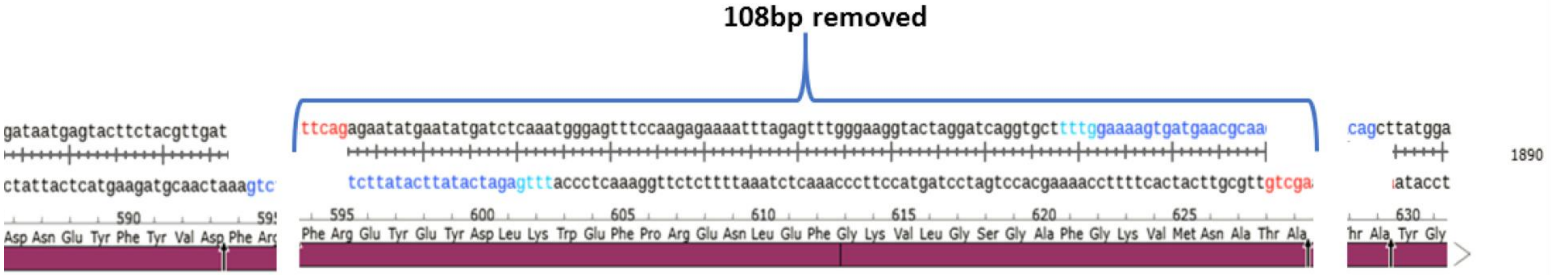


Figure 37. *In vitro* gene editing experimental protocol and tools.
Cpf1 RNP is complexed and added to the first *in vitro* cleavage reaction mixture with plasmid DNA. Plasmid DNA is recovered and added to a second *in* vitro re-circularizing reaction mixture with cell-free extract. After the reaction is complete, plasmid DNA is recovered from the reaction and transformed into competent *E. coli*. DNA is then isolated from transformed cells and sequenced to identify modifications made *in vitro*. Modified from Sansbury et al.[188]


1890
FLT3
Ptn 4
PAM
gRNA
Cpf1-2
gRNA
PAM
Cpf1-1
108bp removed
145bp replaced
FLT3


FLT3
Ptn 5
gRNA
PAM
Cpf1-1
PAM
gRNA
Cpf1-2


108bp removed

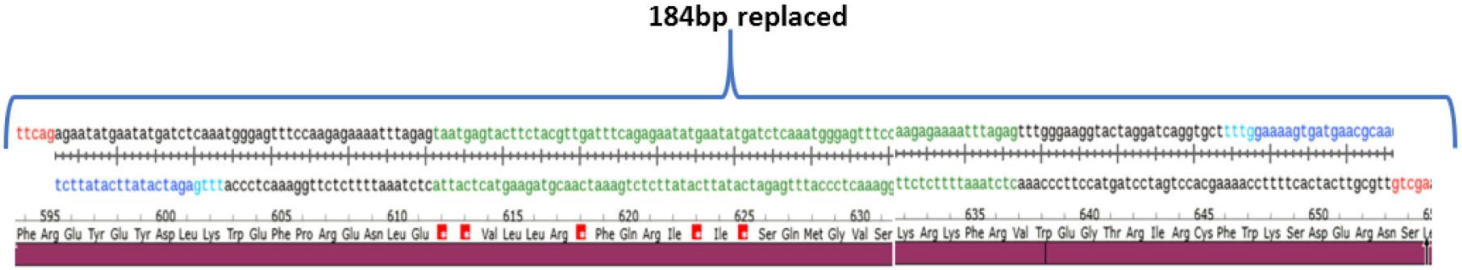

184bp replaced

Figure 38. FLT3 In-vitro Gene Editing Replacement Design.
Replacement reactions were initiated by two cut site designs for the Cas12a enzyme. These cut sites produced fragment that would then be replaced by a double stranded DNA with the mutations of interest.

Mutation	Therapy	Reference	Validated in vitro
F 691 L	Sorafenib	Baker et al. [42]	Ba/F3
D 835 Y	Sorafenib	Baker et al. [42]	Ba/F3

Annesley, CE and Brown, P (2014). The Biology and Targeting of FLT3 in Pediatric Leukemia. *Front. Oncol.* **4**: 1–18.

ID	Sample Type	Sequence	Length
Patient 4	Diagnosis/Relapse	CTTACCAAACTCTAAATTTTCTCTTGGAAACTCCCAT	37
Patient 5	Diagnosis/Relapse /Remission	CTCTAAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATATTCTCTGAAATCAACGTAGAAGTACTCATTA	76

Crowgey, EL, Kolb, A and Wu, CH. Development of Bioinformatics Pipeline for Analyzing Clinical Pediatric NGS Dataat

37

Figure 39. FLT3 Mutant Plasmid Construct.
FLT3 expression plasmid template was used in the in vitro gene editing reaction to produce FLT3 mutant plasmids that express FLT3 ITD mutations, ITDs, point mutations or different combination of both.

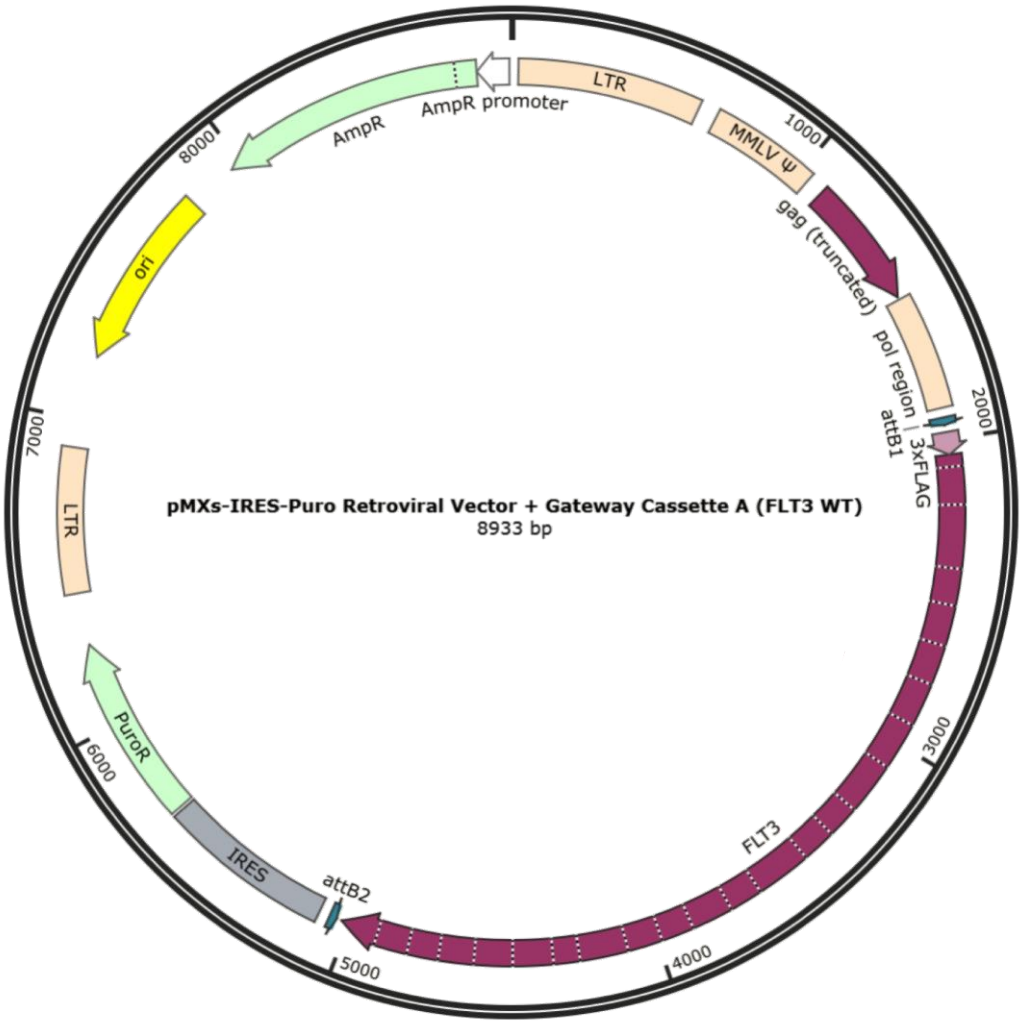


Figure 40. FLT3 Expression Plasmid.
pMxs-IRES-Puro Retroviral vector containing the FLT3 WT cassette was used as the template plasmid for the in vitro gene editing.

FLT3
Ptn 4
FLT3
145bp fragment replacement (ptn 4)
gRNA
PAM
Cpf1-1
Ptn 4
PAM
gRNA
Cpf1-2
(in frame with FLT3)
Pt4
P4/D835Y

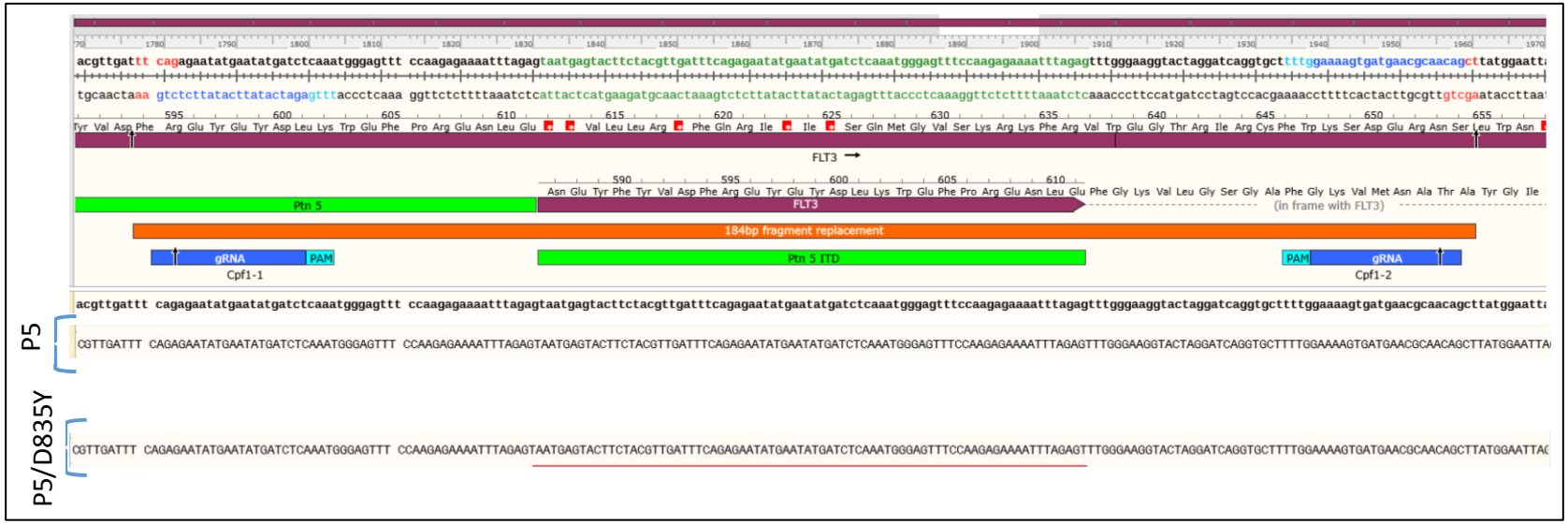


Figure 41. In-vitro Gene Editing Sequencing Confirmation.
Individual ampicillin resistant colonies were picked and sequenced to corroborate that the intended FLT3 mutant plasmids had been created by in vitro gene editing. The sequencing results were aligned to the designed template to confirm the presence of the mutation (ITD) of interest.

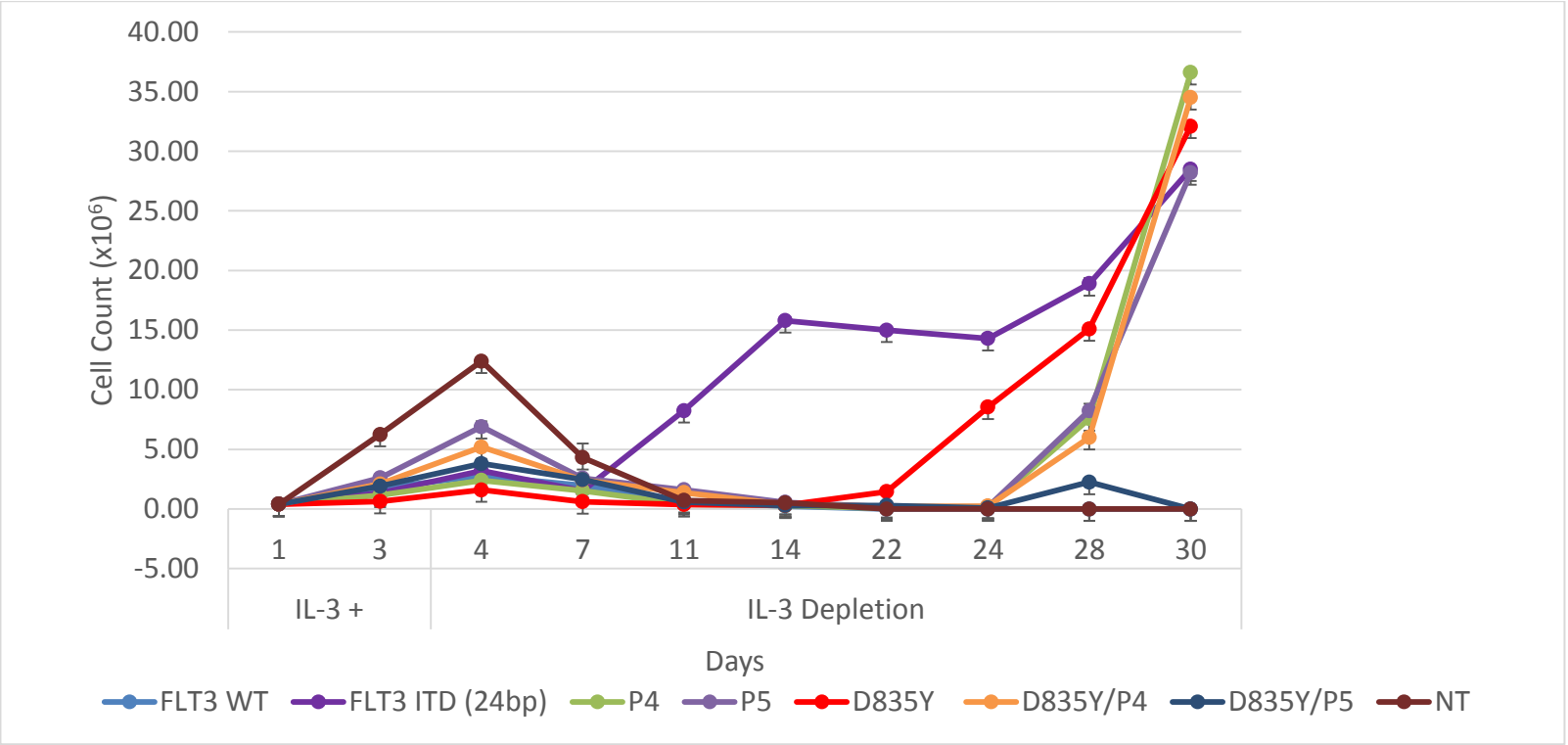


Figure 42. Baf/3 cell transformation by FLT3 Expression.
4 x10^5 Baf/3 cells were transfected with 2ug of plasmid DNA by Nucleofection using the 4D-Nucleofector SG Cell Line X Kit and CM-147 program. Cells were incubated for 72hrs post transfection in IL-3 supplemented media. At this time 2 x10^5 cells were plated into a 6-well plate containing IL-3 supplemented media. At 4 days cells were counted and plated into media lacking IL-3 to select for transformed cells. Cell count was performed by trypan blue exclusion.

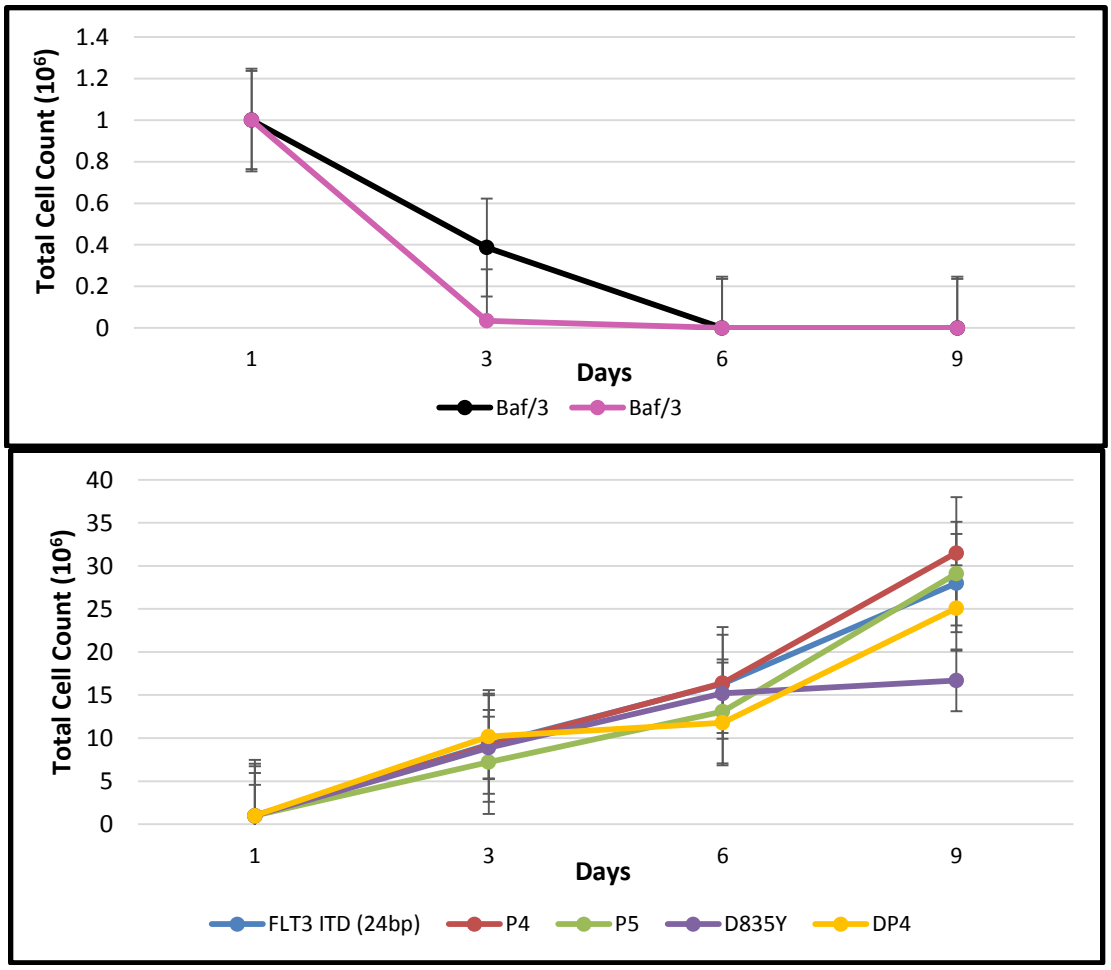


Figure 43. Puromycin selection growth curve.
$1x10^6$ cells that showed IL-3 depletion survival were cultured for 9 days in the presence of puromycin to further confirm the FLT3 cassette stable integration. Top graph was used as control to show the sensitivity of un-transfected Baf/3 cells to puromycin in the presence or absence of IL-3.

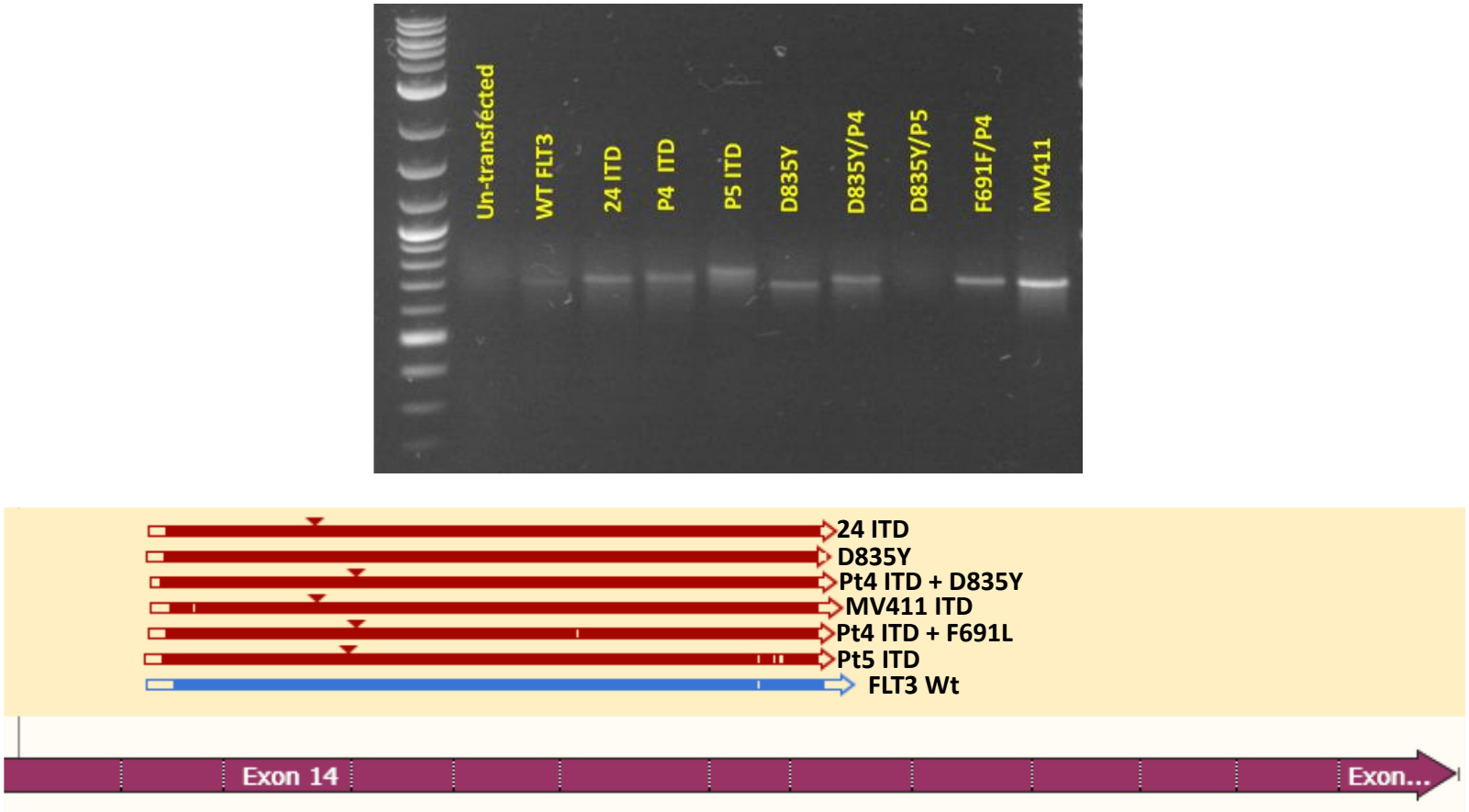
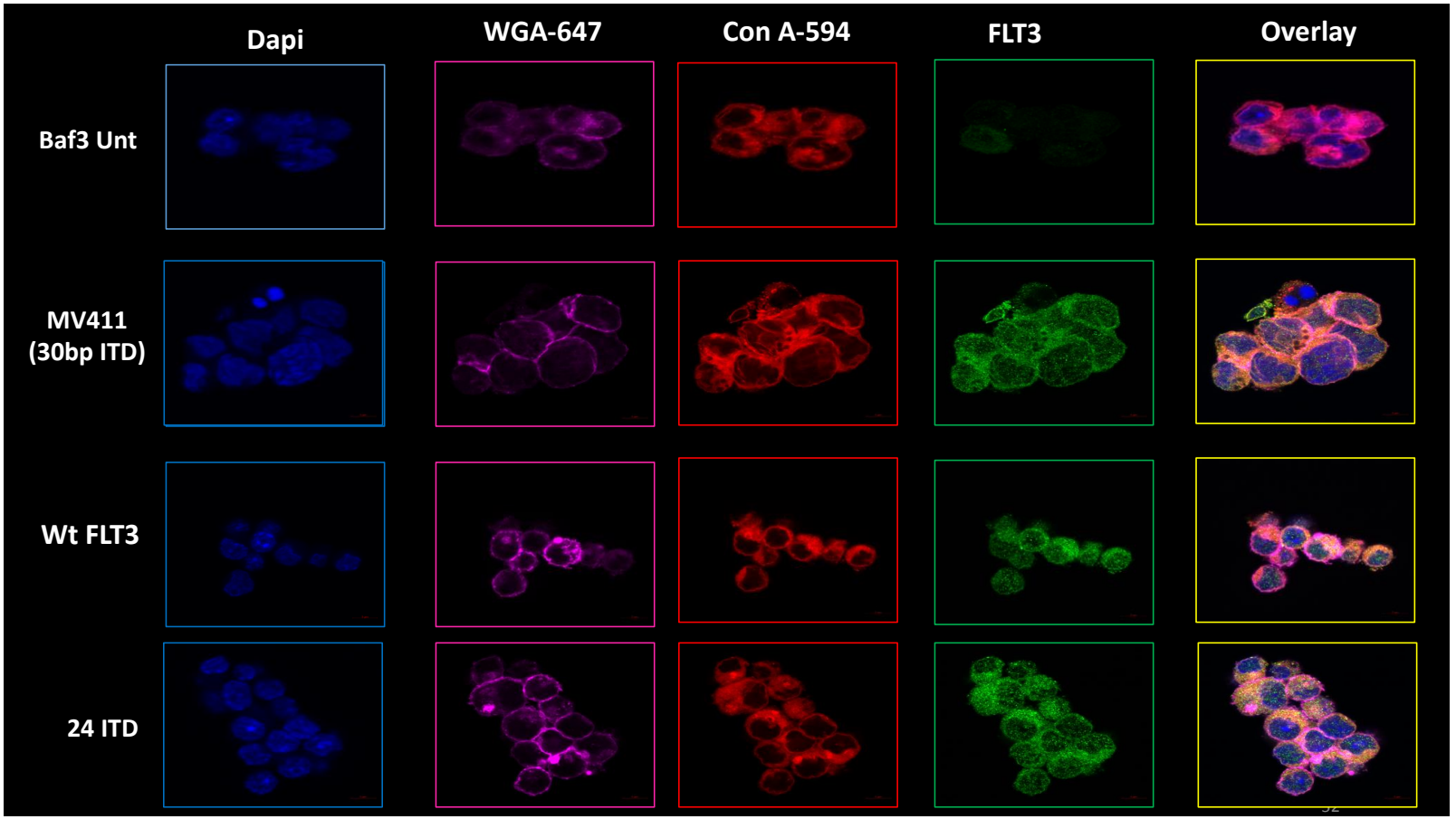
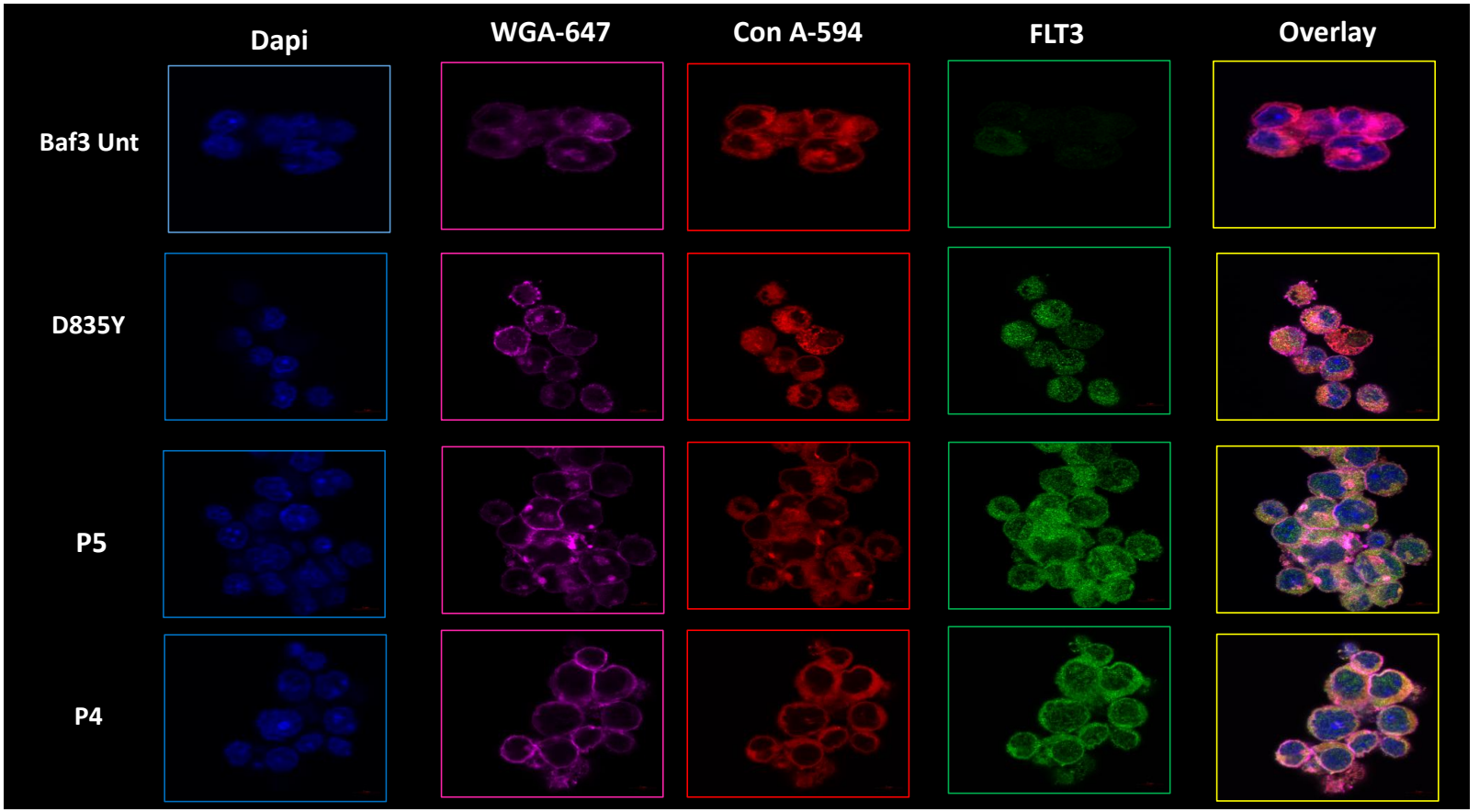


Figure 44. FLT3 Expression Analysis of Transformed cells.
RT-PCR products were analyzed through gel electrophoresis. All RT-PCR products were sequenced to confirm the presence of the different FLT3 mutations of interest. The sequence results were aligned to the FLT3 WT gene sequence with the triangle depicting were the ITDs are present.


Dapi
WGA-647
Con A-594
FLT3
Overlay
Baf3 Unt
MV411
(30bp ITD)
Wt FLT3
24 ITD


Dapi
WGA-647
Con A-594
FLT3
Overlay
Baf3 Unt
D835Y
P5
P4

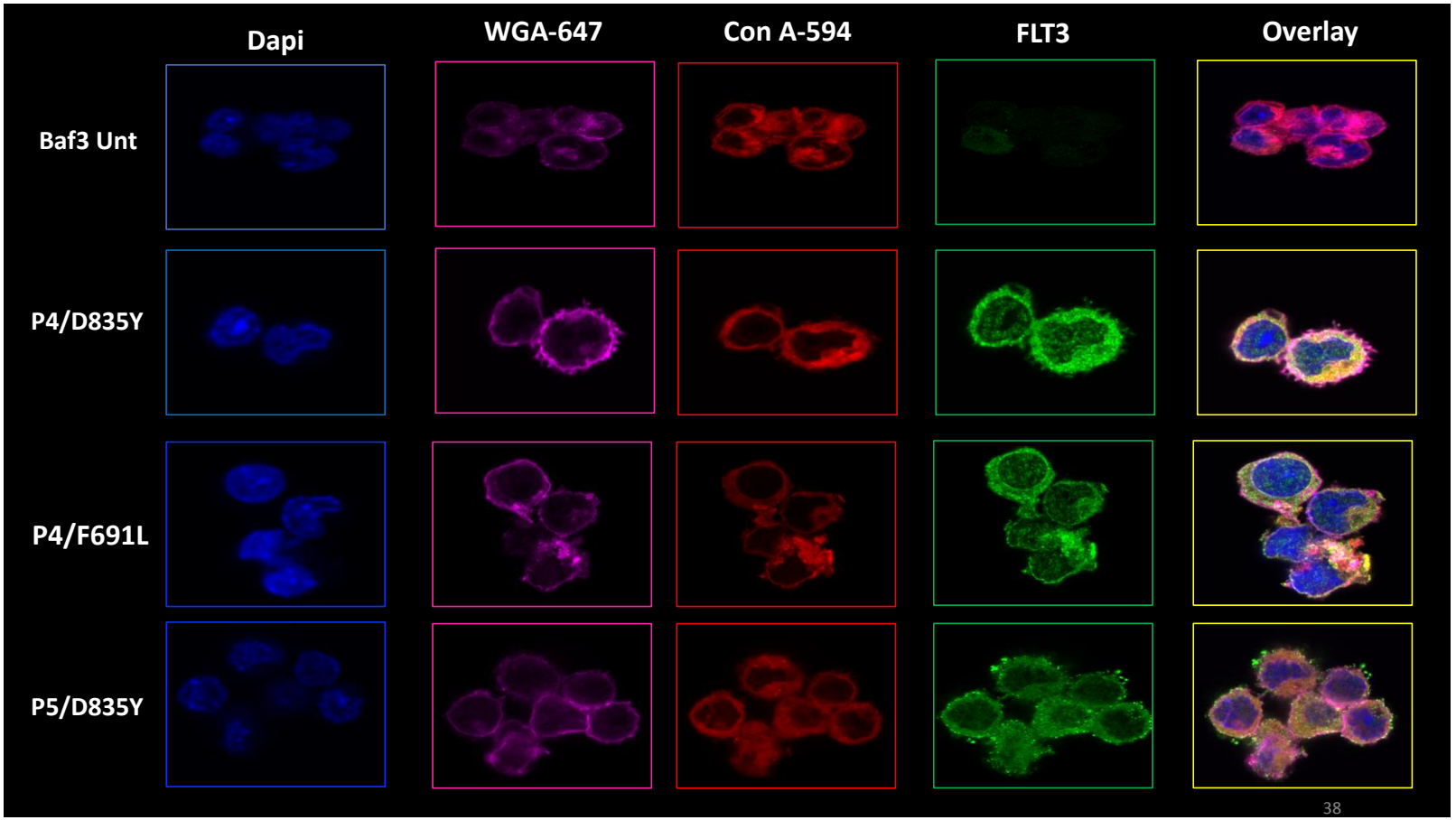


Figure 45. Localization of FLT3.
Immunofluorescence staining of Baf/3 FLT3 expressing cells and un-transfected cells.

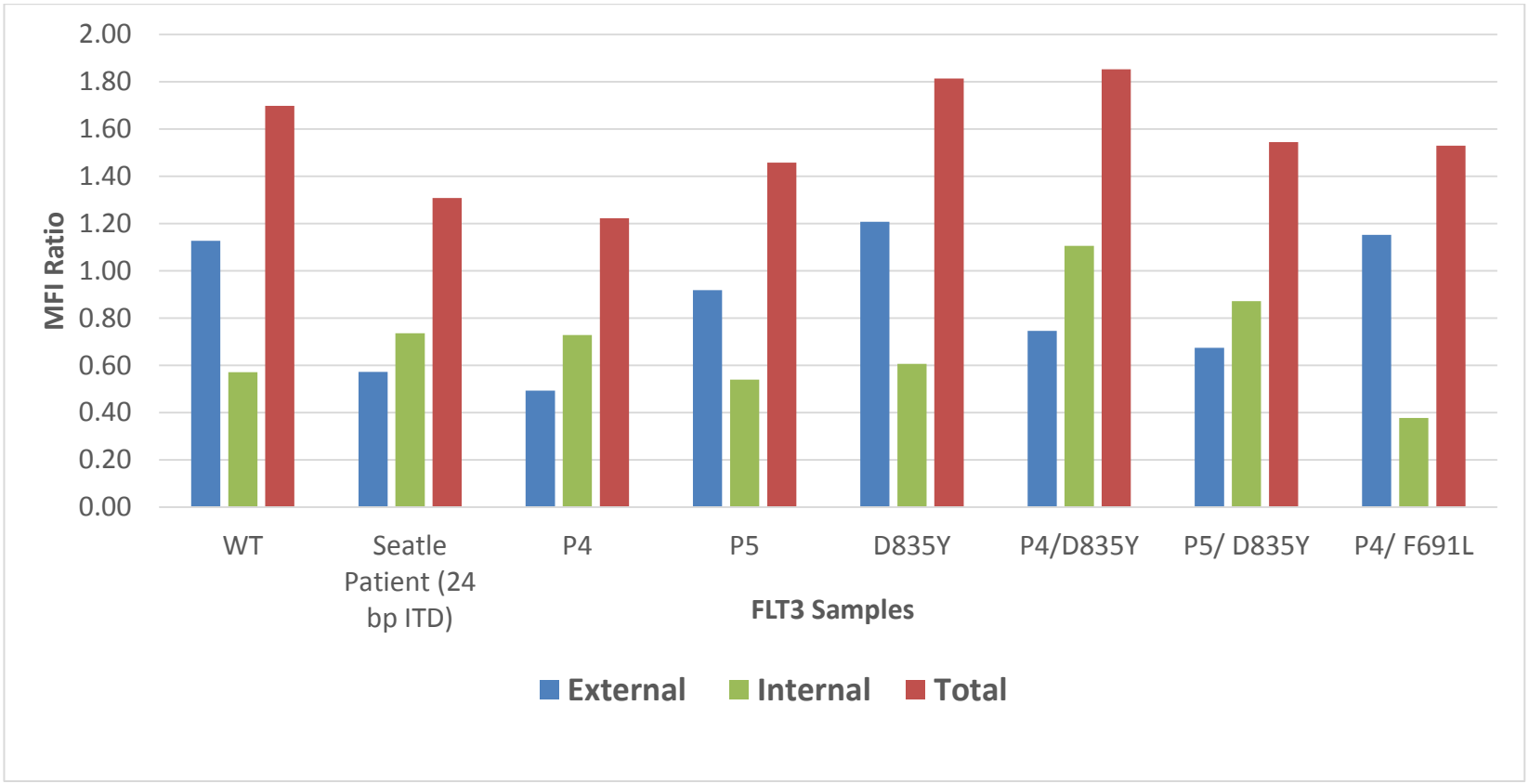
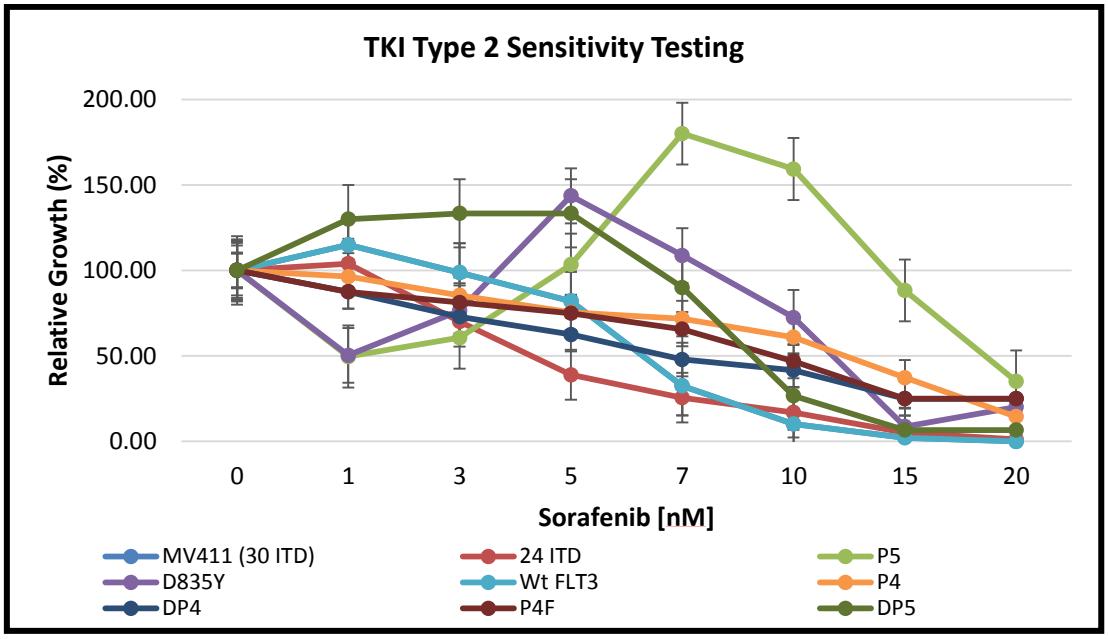


Figure 46. FLT3 Expression Quantification.
Bar graph showing the FLT3 surface expression in Baf/3 cells, harboring different FLT3 genotypes.



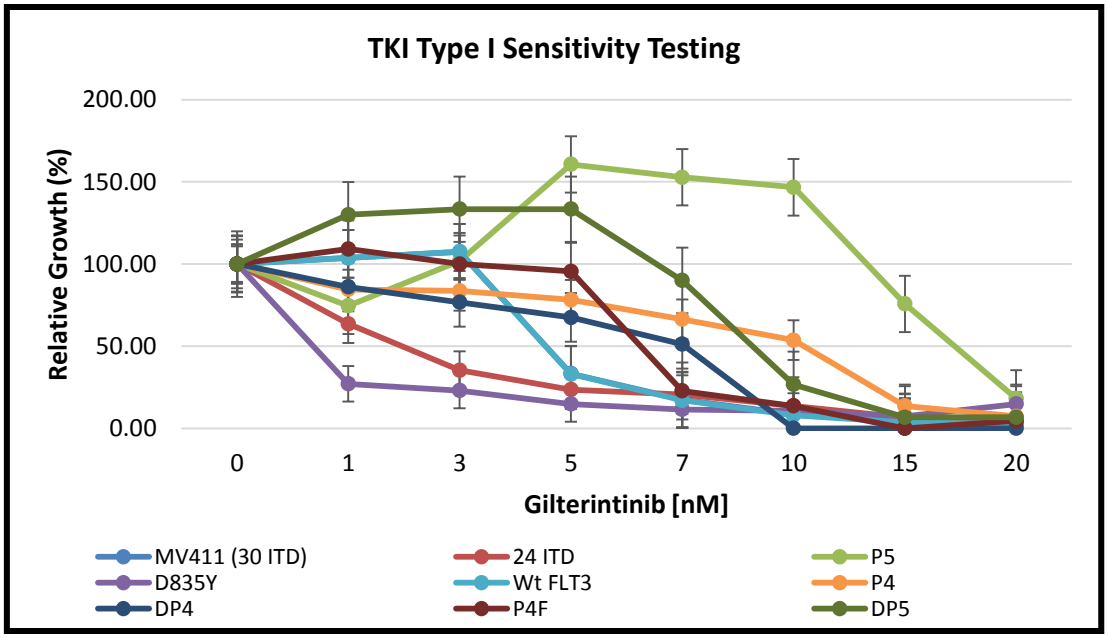
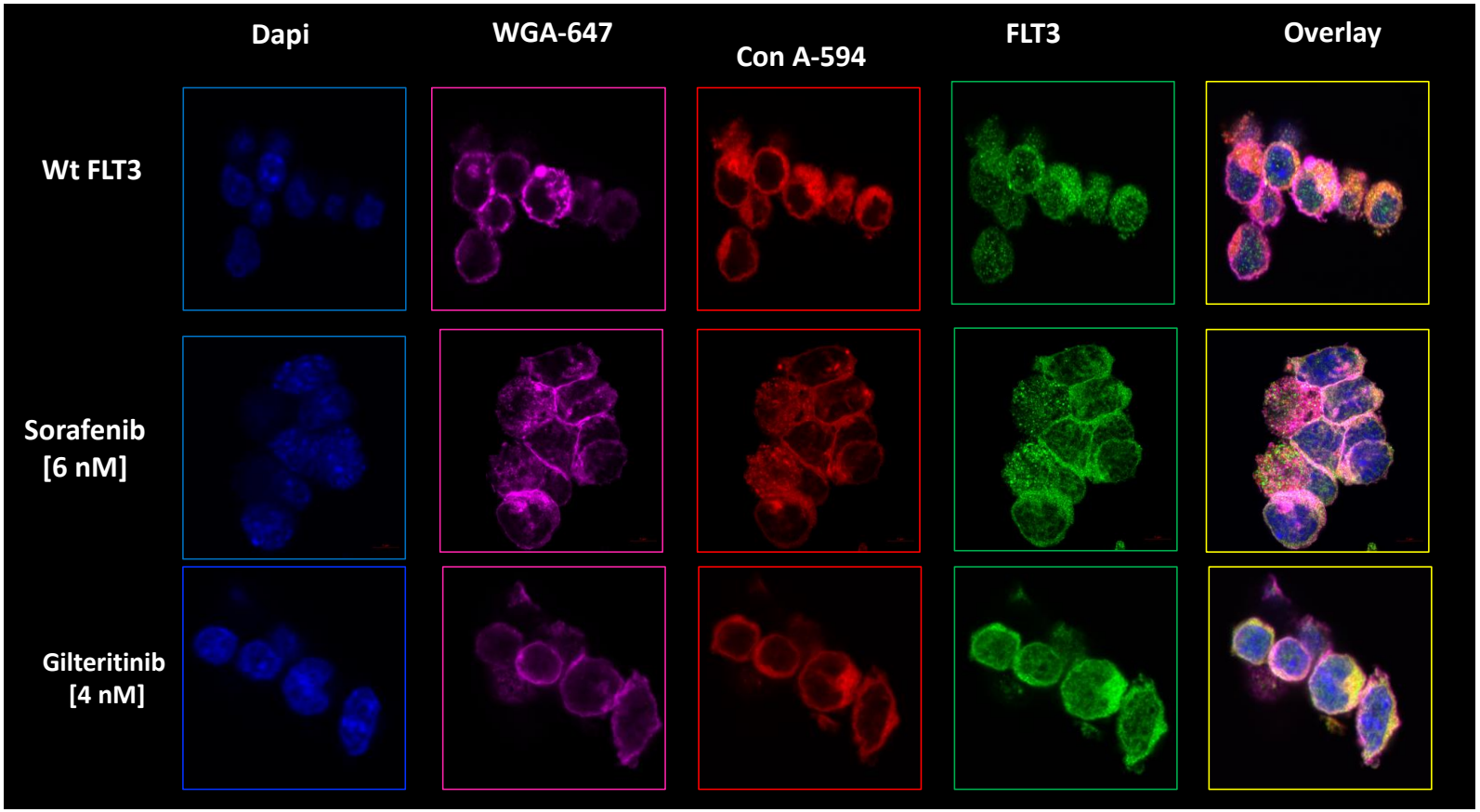
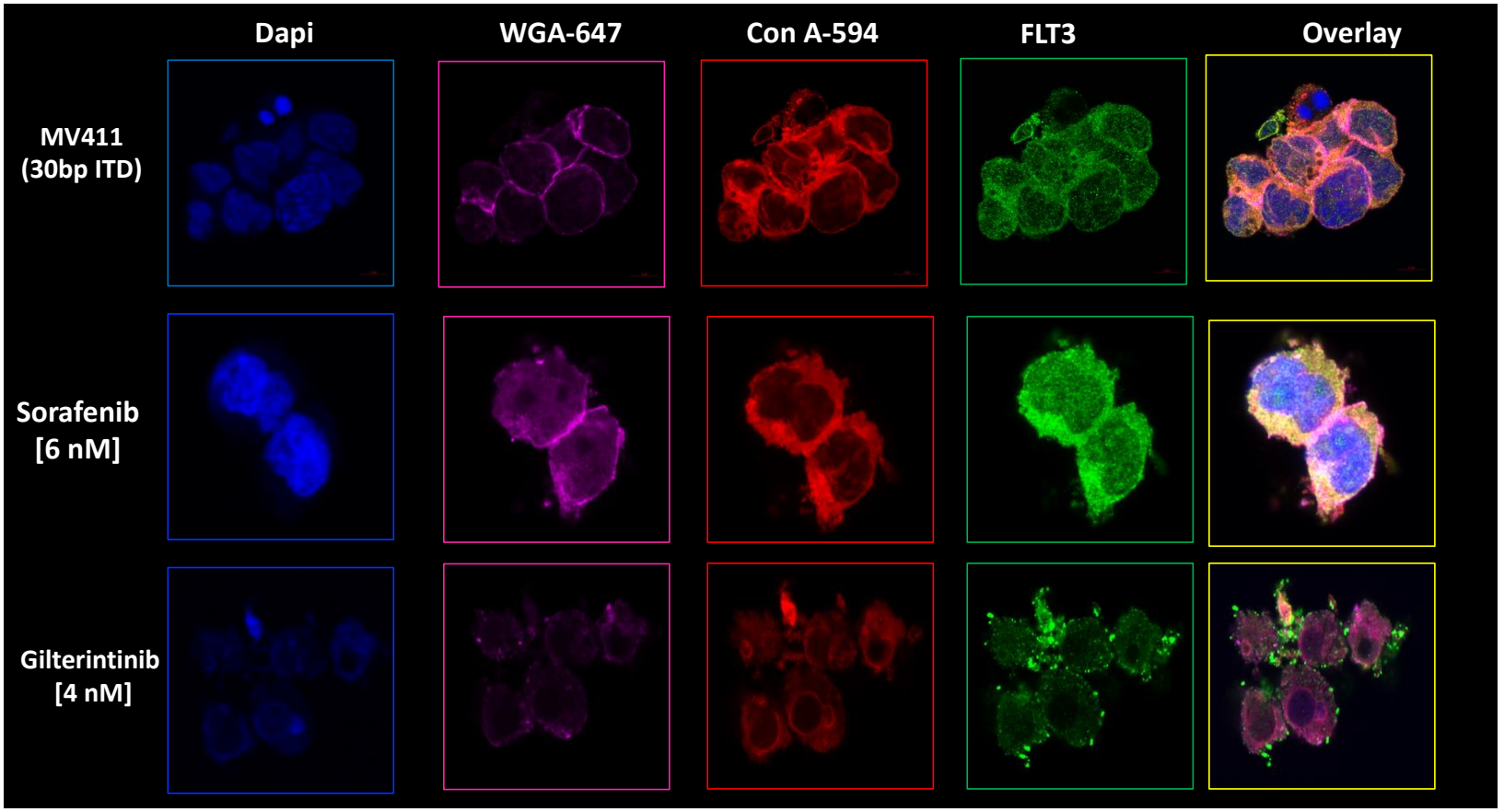


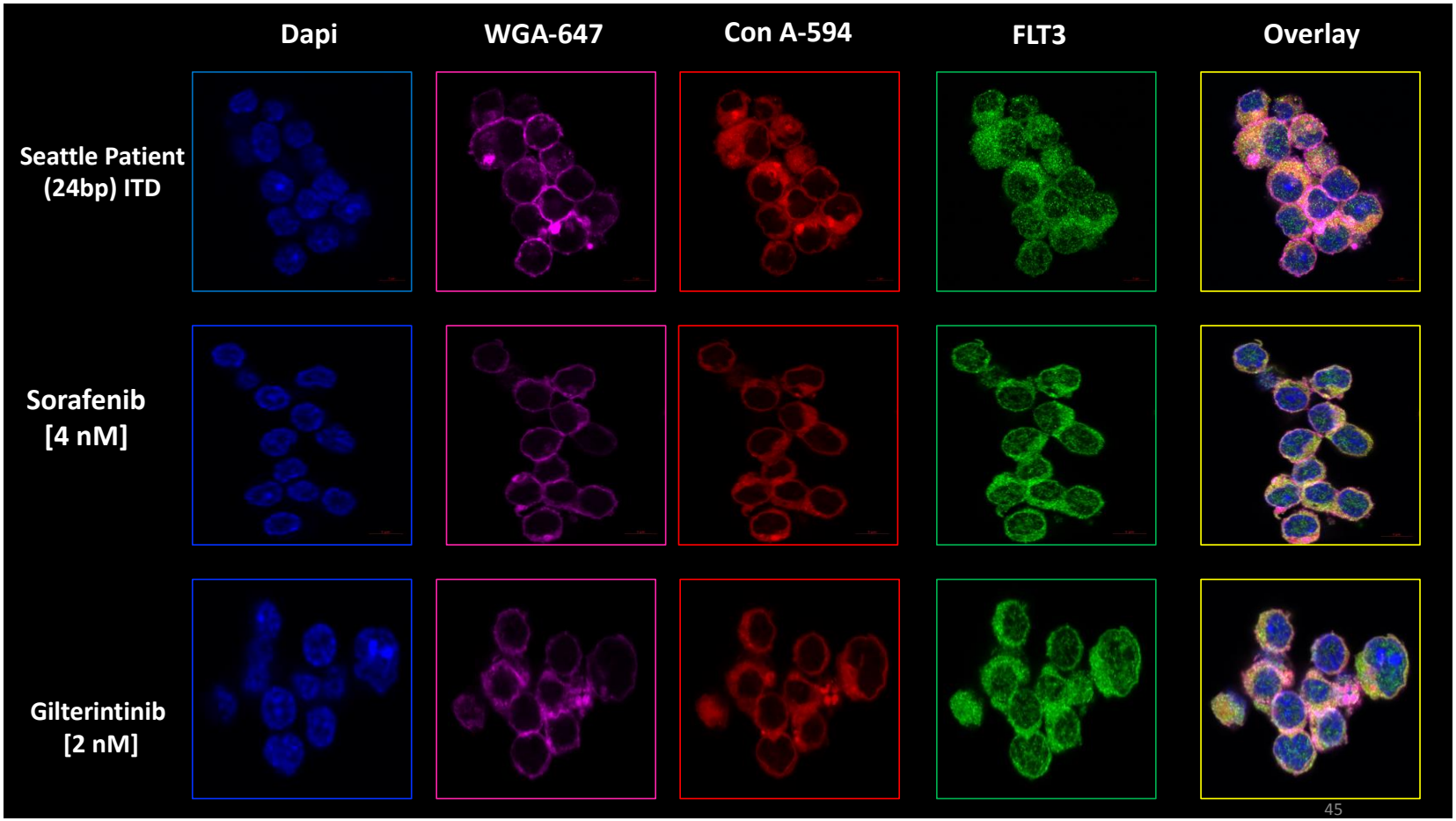
Figure 47. FLT3 Mutation Drug Sensitivity Assay.
Relative cell growth of Ba/f3 cells expressing the indicated FLT3 mutants after 72hr treatment with increasing concentrations of Sorafenib or Gilteritinib. Errors bars represent the Standard Error (n=4)

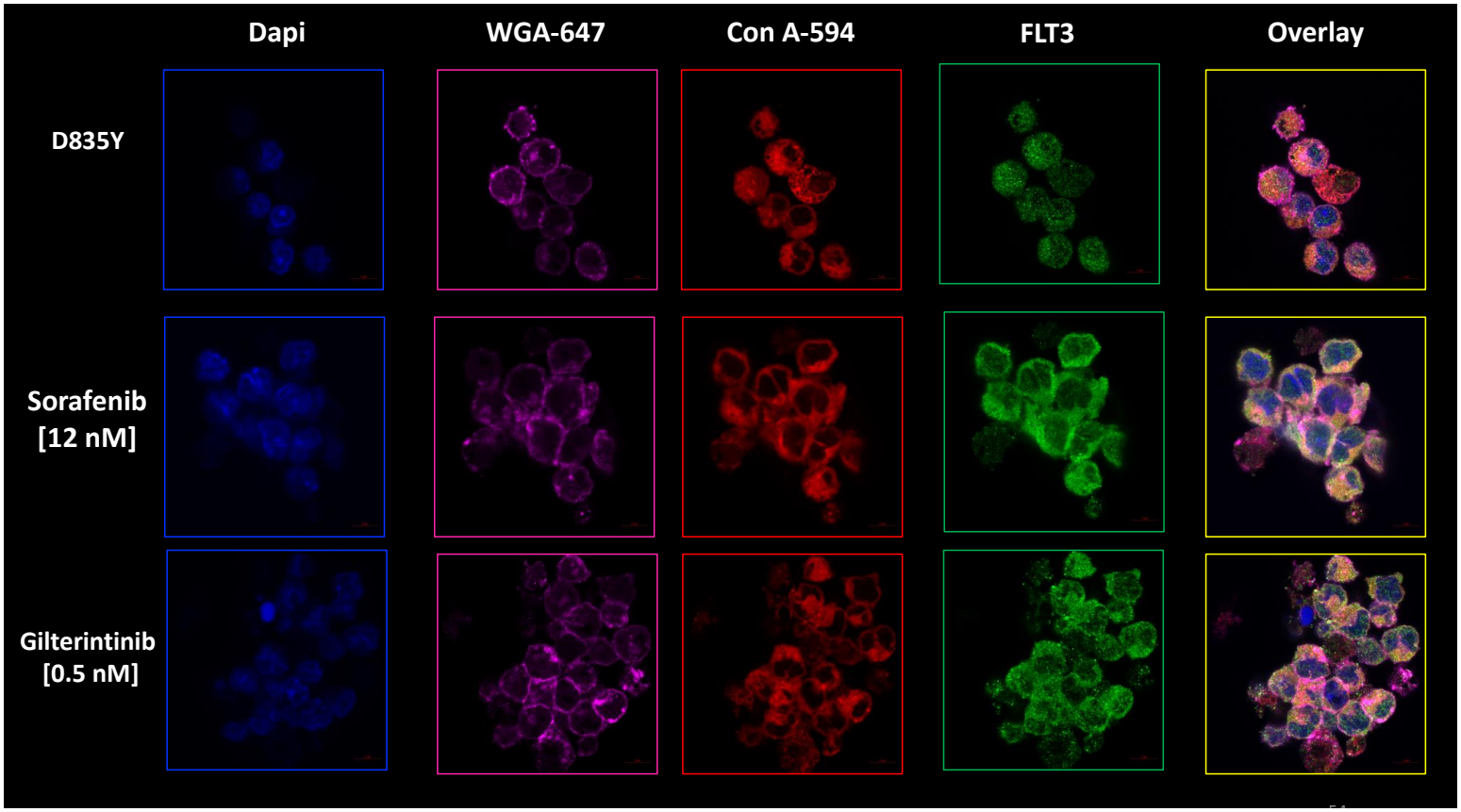
Table 7. FLT3 TKI IC_{50}

	Sorafenib (nM)	Giletritinib (nM)
Wt FLT3	6.0	4.0
MV411 (30bp ITD)	6.0	4.0
Seatle Patient (24bpITD)	4.0	2.0
P4	12.5	11.0
P5	18.0	16.0
D835Y	12.0	0.5
P4/D835Y	7.0	7.0
P5/D835Y	8.0	8.0
P4/F691L	9.0	6.0


Dapi
WGA-647
Con A-594
FLT3
Overlay
Wt FLT3
Sorafenib
[6 nM]
Gilteritinib
[4 nM]


Dapi
WGA-647
Con A-594
FLT3
Overlay
MV411
(30bp ITD)
Sorafenib
[6 nM]
Gilterintinib
[4 nM]


Dapi
WGA-647
Con A-594
FLT3
Overlay
Seattle Patient (24bp) ITD
Sorafenib [4 nM]
Gilterintinib [2 nM]
45


Dapi
WGA-647
Con A-594
FLT3
Overlay
D835Y
Sorafenib
[12 nM]
Gilterintinib
[0.5 nM]

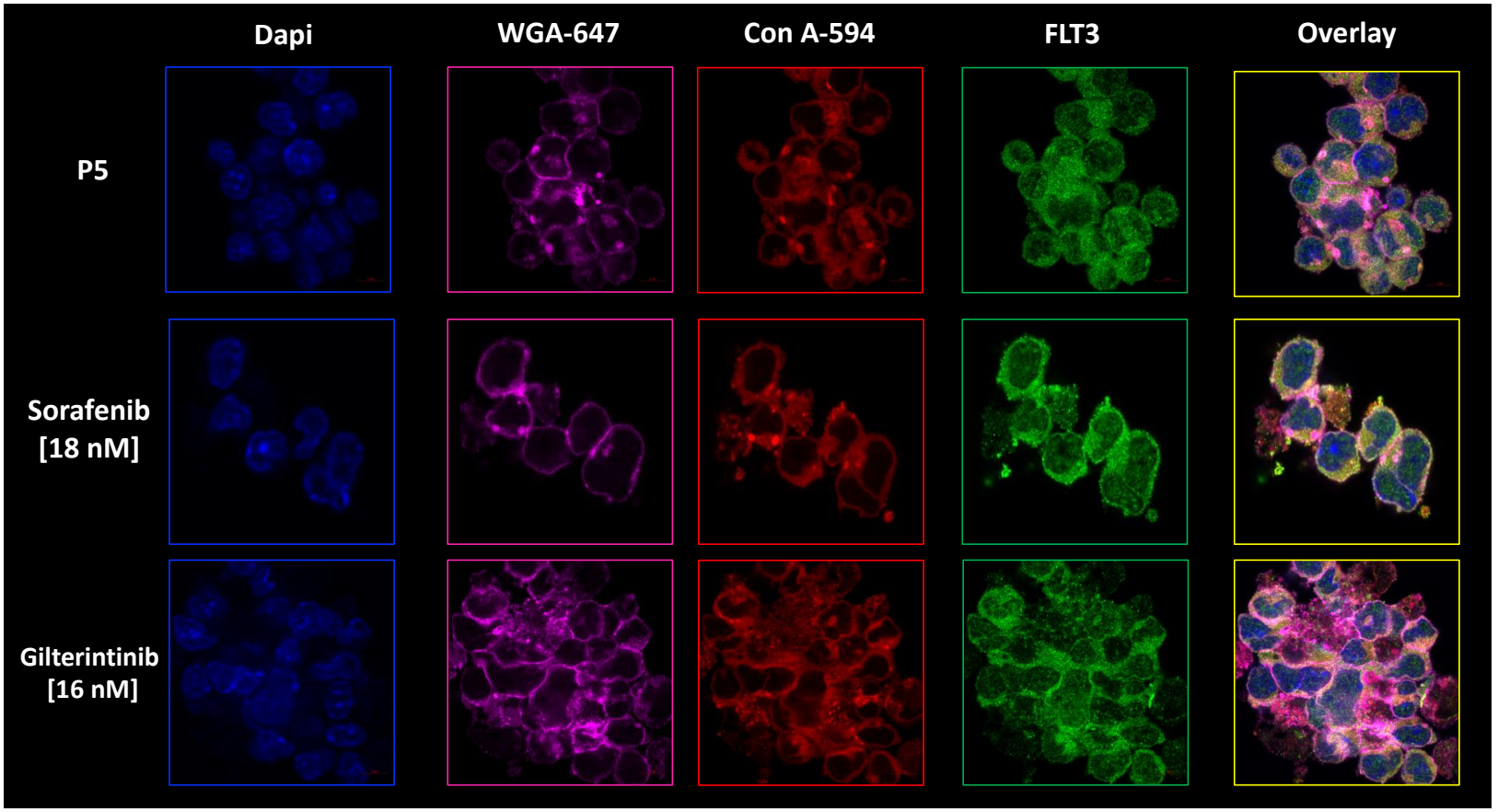


Figure 48. FLT3 Localization under TKI treatment.
Immunofluorescence staining of Baf/3 FLT3 expressing cells and un-transfected cells with and without TKI treatment.

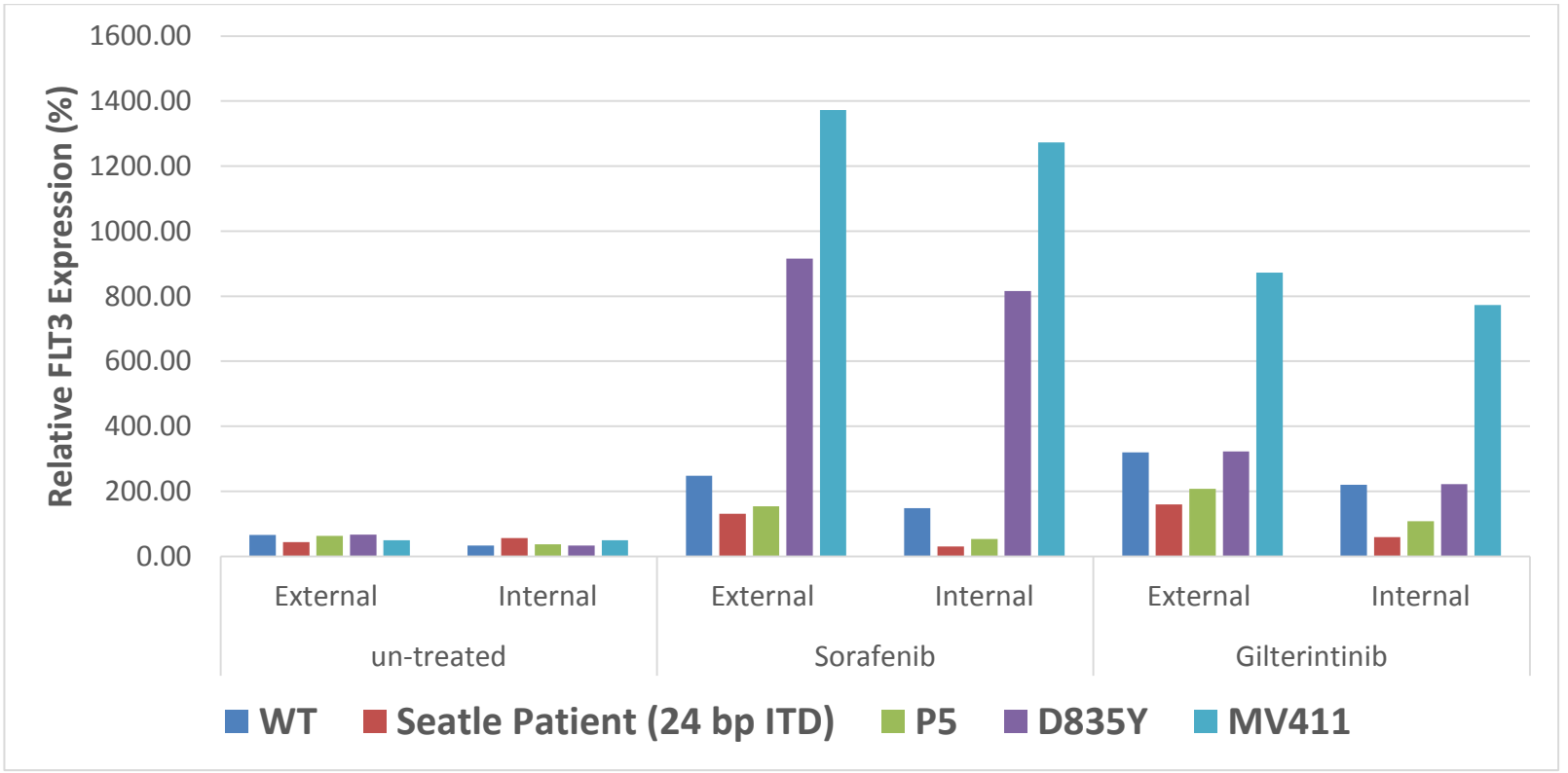


Figure 49. Relative FLT3 Expression in the presence of TKIs treatment.
Bar graph showing the relative FLT3 expression in Baf/3 cells, harboring different FLT3 genotypes, with or without TKI treatment.

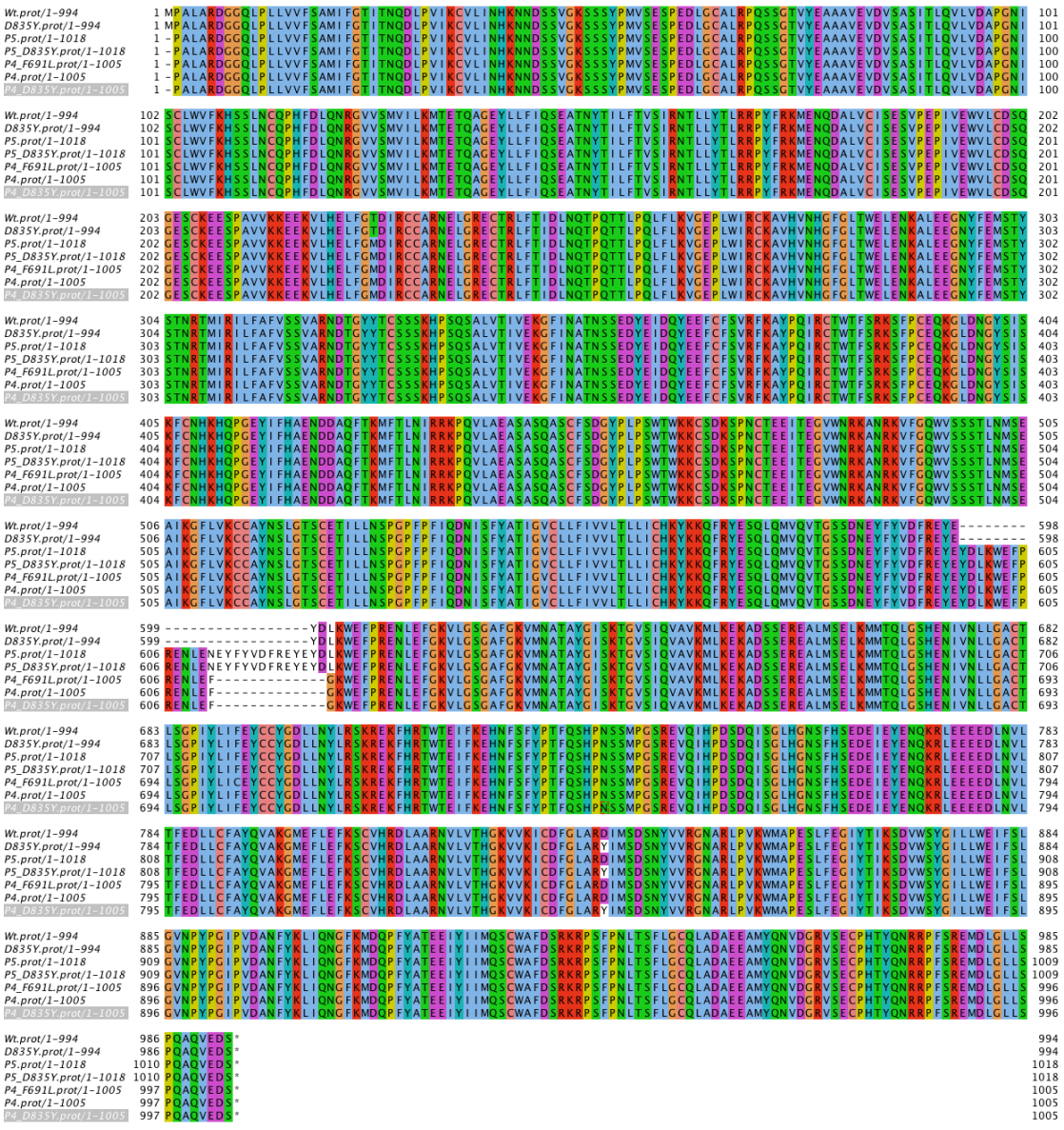


Figure 50. Clustal Omega Results.

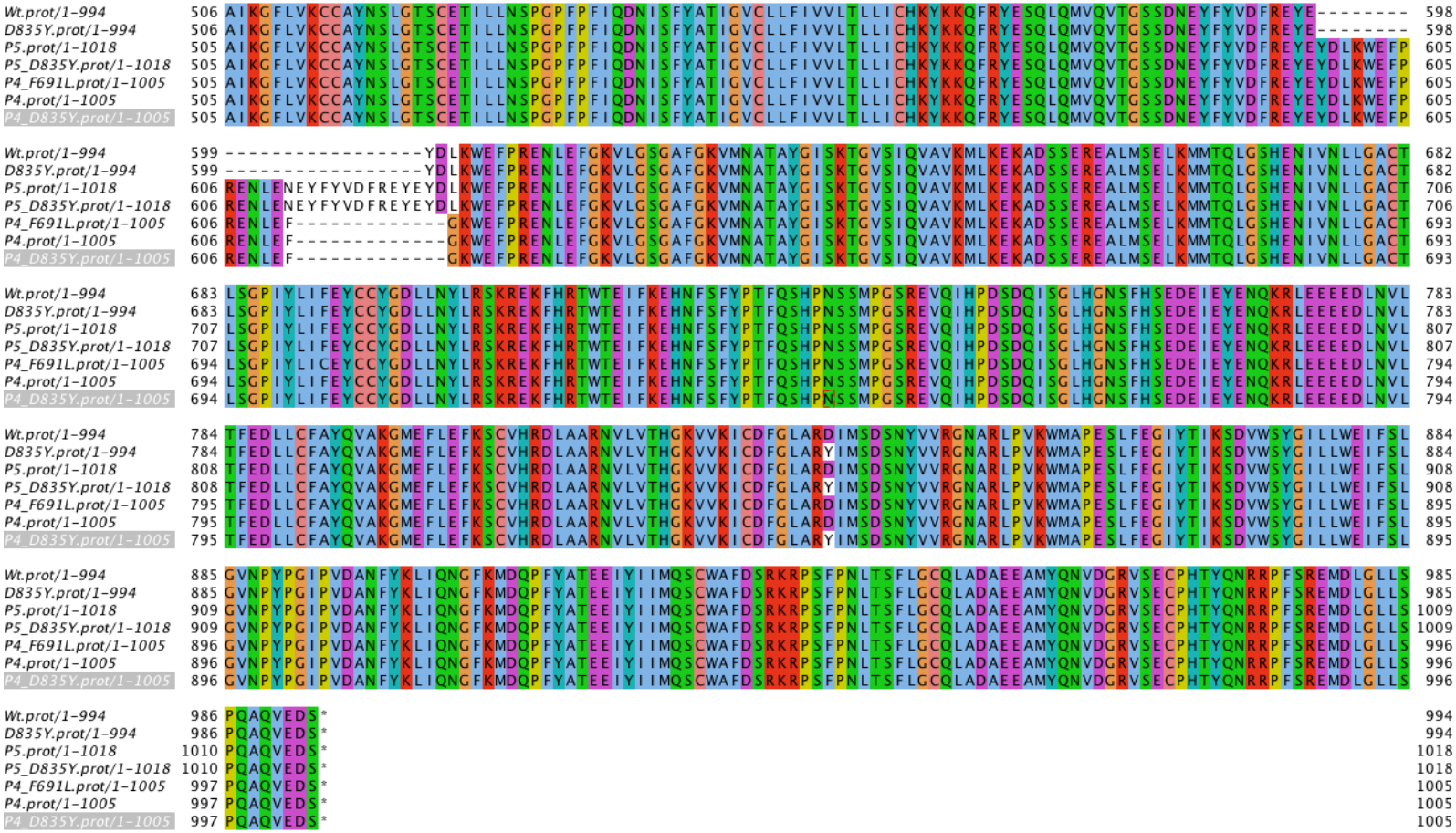
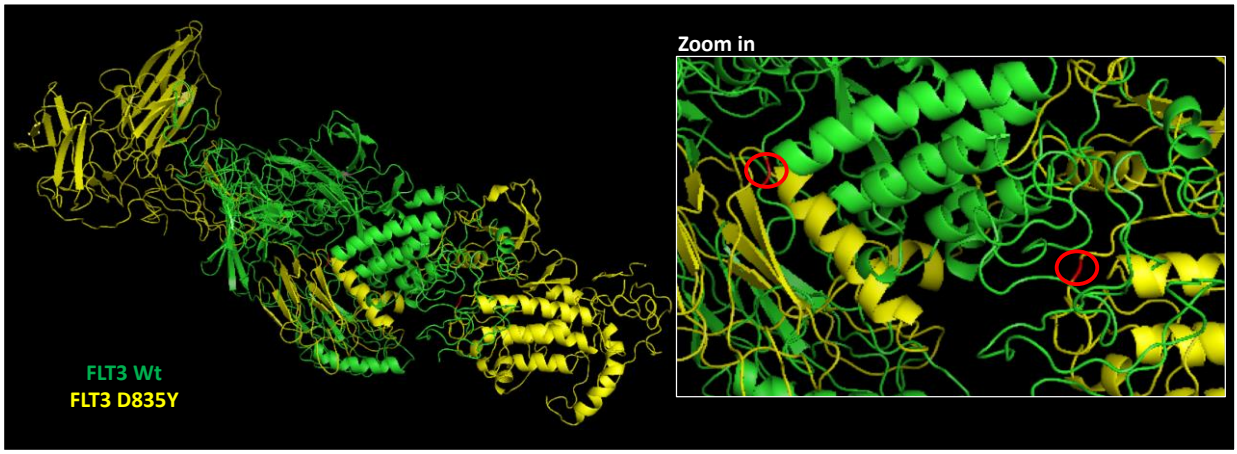
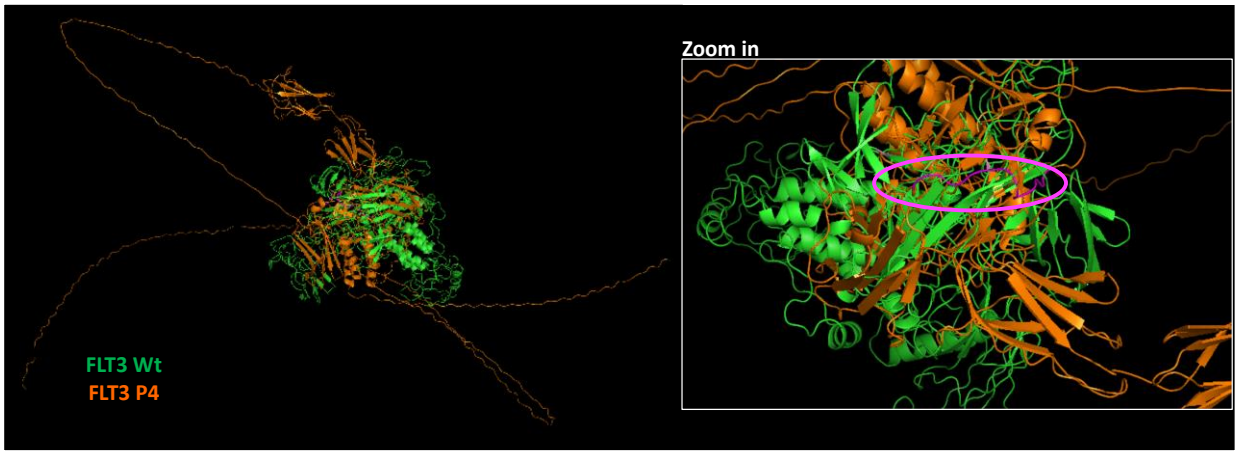
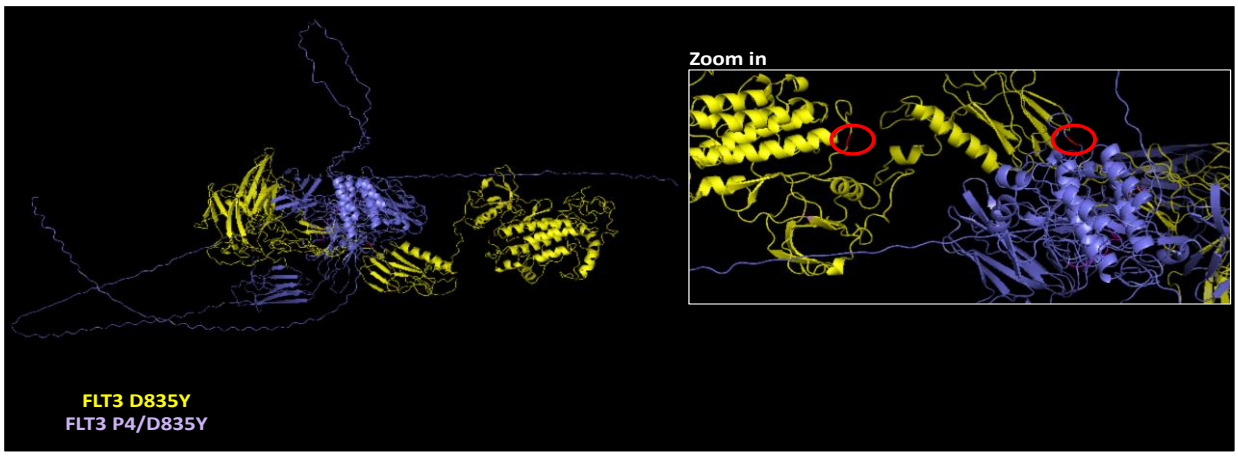
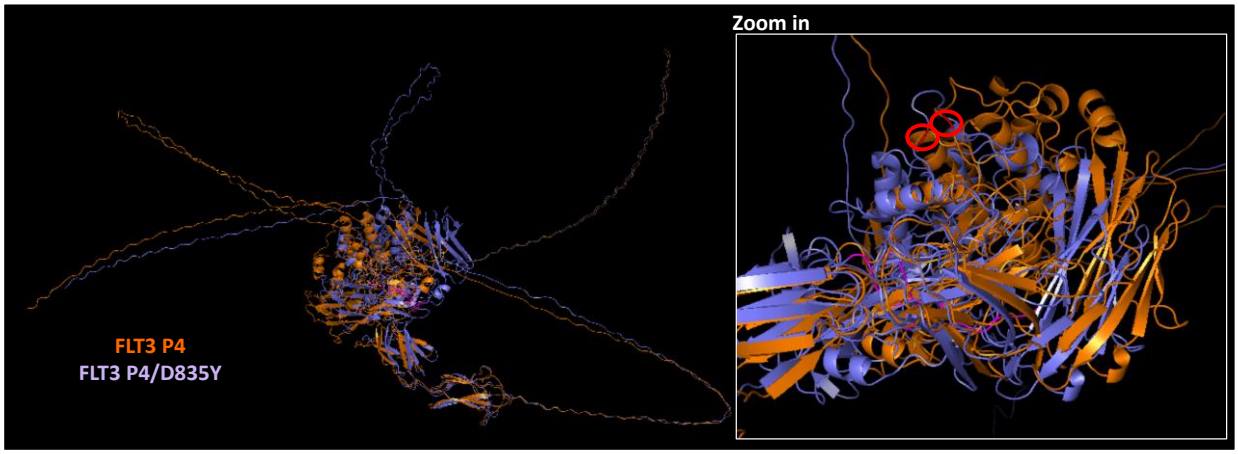


Figure 50. Clustal Omega Results.
Multiple Sequence Alignment


Zoom in
FLT3 Wt
FLT3 D835Y

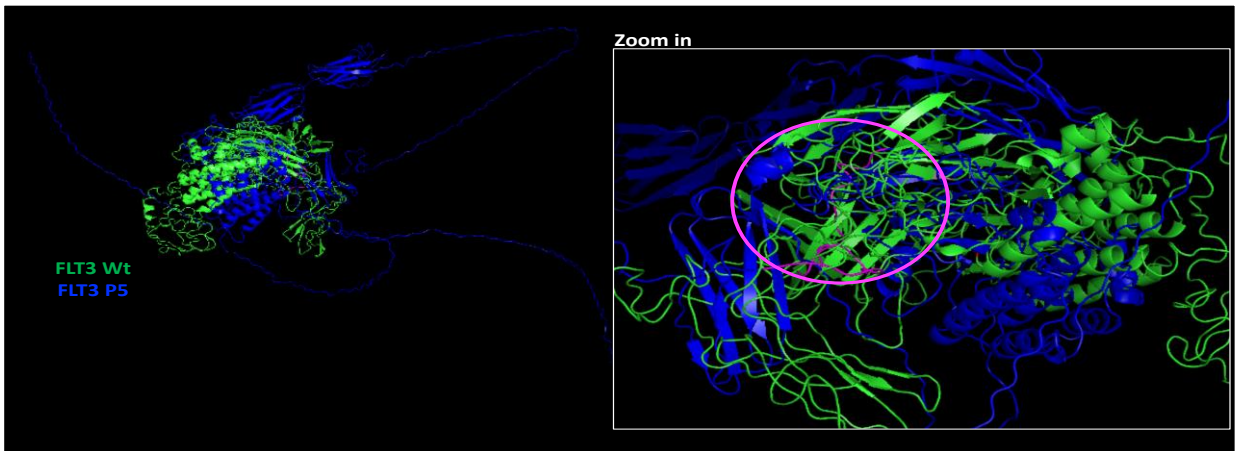

Zoom in
FLT3 Wt
FLT3 P4

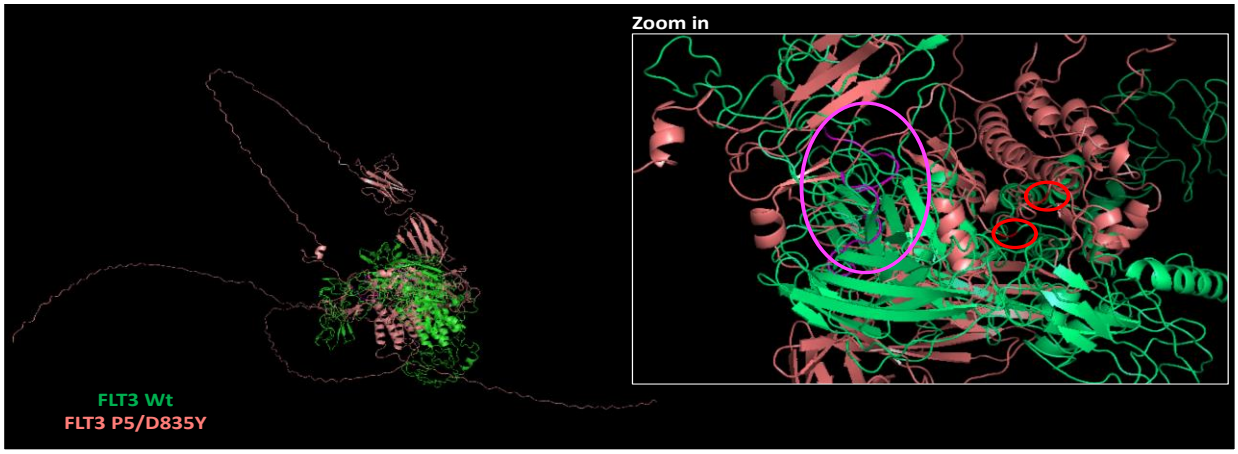

Zoom in
FLT3 D835Y
FLT3 P4/D835Y


Zoom in
FLT3 P4
FLT3 P4/D835Y


Zoom in
FLT3 Wt
FLT3 P4/F691L


Zoom in
FLT3 P4
FLT3 P4/F691L


Zoom in
FLT3 Wt
FLT3 P5


Zoom in
FLT3 Wt
FLT3 P5/D835Y

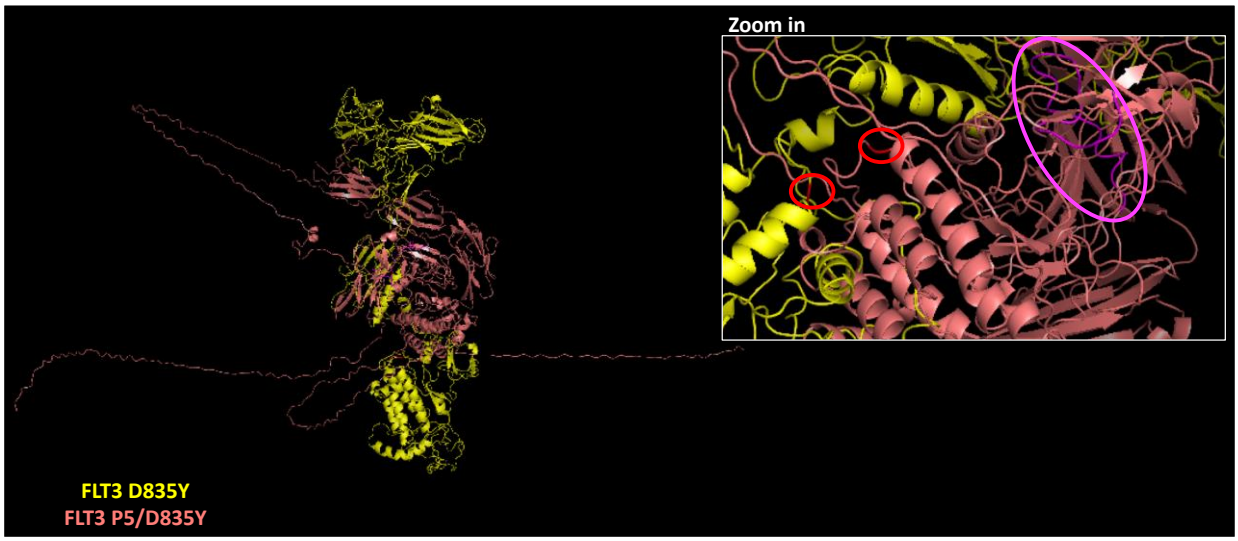


Figure 51. Mutant FLT3 Protein Structure Modeling.

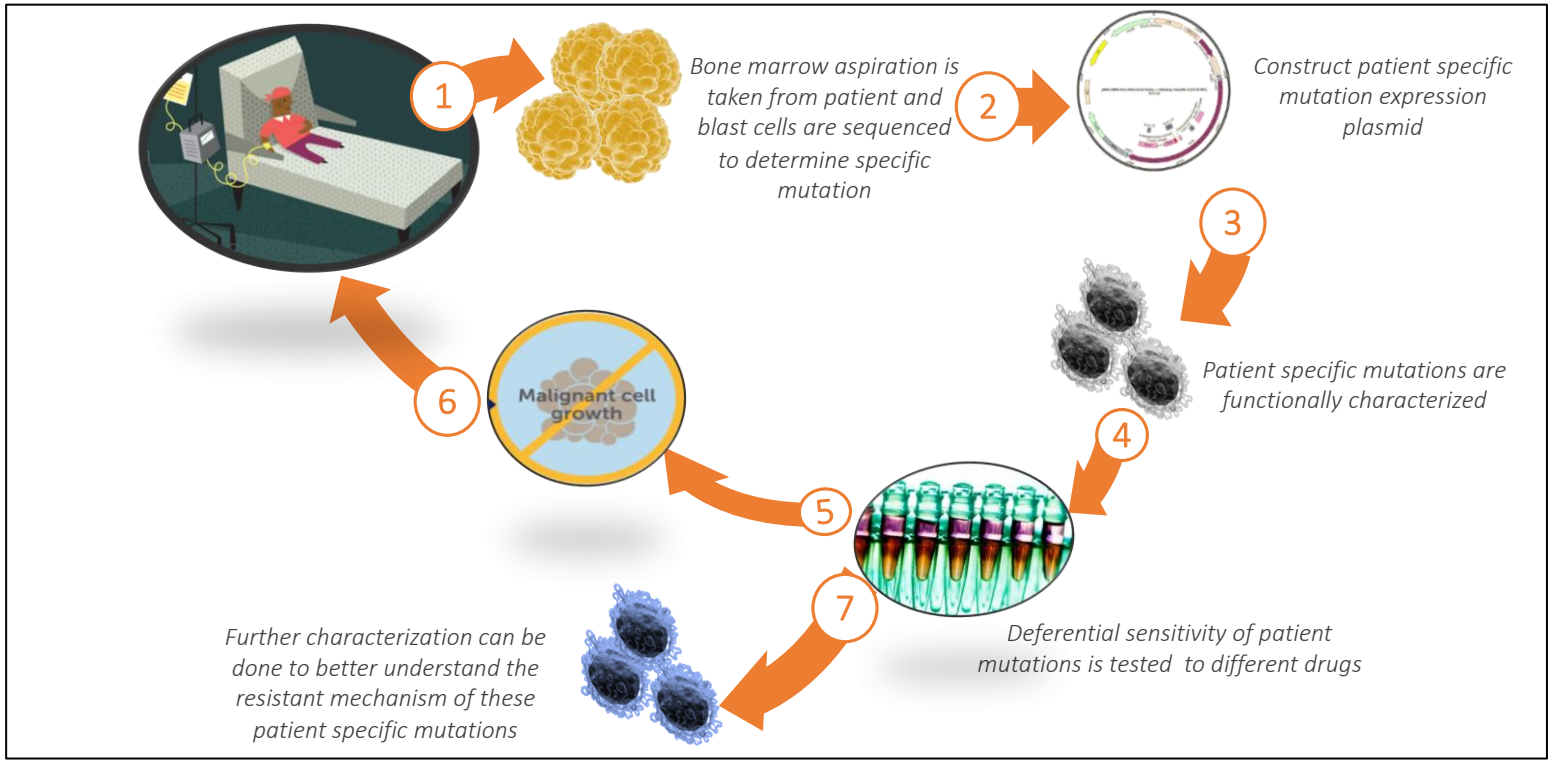


Figure 52. Modeling Pediatric AML FLT3 Mutations Mediated by Gene Editing

Chapter 8

SUMMARY AND PERSPECTUS

Single base mutations can be repaired by introducing single stranded DNA oligonucleotides (ssODN) into a target cell [3–5]. The frequency at which this occurs is dependent on several of factors: the length of ssODN, the position of the cell in its proliferative cycle [6,7], and the presence of double-stranded DNA breaks in the host genome [8,9]. Three biomolecules used to catalyze specific double stranded DNA breaks: Zinc-Finger Nucleases (ZFNs), Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 and Transcription Activator-Like Effector Nucleases (TALENs) [10–15]. Genome editing offers a promising strategy for gene repair and correction by overcoming difficulties associated with lack of precision. CRISPR/Cas has increased the pace and lowered the cost of research, allowing the genetic manipulation even in organisms that have historically been difficult to modify [68]. Furthermore, the combinatorial approach uniting ssODNs and CRISPR/Cas9 has emerged as a feasible therapeutic approach [189].

In the work presented in this dissertation I focused on the mechanism and application of gene editing utilizing CRISPR systems. In Chapter 2-5 the HCT 116-19 cells ,which are a well-established gene editing model system, were used to elucidate the mechanism of action [21,34,40]. This model system enabled the correlation between genotypic and phenotypic changes with functional protein activity which are a critical component of reaction optimization or characterization studies.

In **Chapter 2** I tested the combinatorial approach of utilizing CRISPR/Cas9 system along with ssODN to promote single base pair correction. While the traditional endpoint for genetic engineering utilizing CRISPR/Cas9 has been to disrupt or disable a gene through complete knockout, as I demonstrate, it is now possible to direct single nucleotide exchange in efficient manner. This approach will permit gene editing to generate inheritable nucleotide changes at higher efficiencies and, as such, could be used in the treatment of genetic disorders. We explored this combinatorial approach of CRISPR/Cas9 and ssODN, in more detail in **Chapter 3**. Here the oligonucleotide was designed so that it invaded the target duplex with a 3′ to 5′ polarity, consistent with traditional reaction mechanics of homologous pairing [190–193]. A critical reaction intermediate, a D-loop, is resolved and the resolution directs base exchange to proceed. In addition, cells undergoing DNA replication, passing through S phase, were found to be more amenable to gene editing activity [5,7,86,107,123,194], consistent with the mechanism of action. The pathway of gene editing utilizes the process of Homology Directed Repair (HDR) [40,92,93,103].

Since the active of CRISPR/Cas9 complex consists of RNA and protein, one approach is to target cells with a preformed Ribonucleoprotein (RNP) may exhibit nonspecific mutagenesis [105,112,195–199] (**Chapter 4)**. While analyses of off-site mutagenesis occupies the attention of a majority of workers in the field, some reports have focused on mutagenesis at the target site [114,200]. Recently, our laboratory analyzed a population of cells bearing a single base change induced by the combination of CRISPR/Cas9 and ssODNs for altered DNA sequence of the beta globin gene [115]. These findings indicate that point mutation repair directed by these gene editing tools leave a mutagenic footprint. We find that both insertions and

deletions accompany single base repair as judged by allelic analysis of clonally expanded cell populations. These results prompted us to investigate the type of DNA heterogeneity created at the site of single base repair in both corrected and uncorrected cell populations in more detail. The data collected and analyzed in **Chapter 4** confirm that the well-established gene editing system, provides foundational information with regard to the generation of genetic lesions and the process of on-site mutagenesis. CRISPR/Cas9 and single- stranded oligonucleotide donor DNA molecules working in tandem can lead to the precise repair of the point mutation in the eGFP gene. These observations led us to propose a new model for the repair of point mutations, a molecular pathway in which the donor DNA acts as a replication template for the repair of the mutant base, a process we have termed ExACT [45]. The targeted population, not exhibiting the corrected phenotype, displayed a variety of cells containing heterogeneous and widely ranging DNA indels surrounding the target site. Since no previous report has indicated that single- stranded oligonucleotides acting as single-agent gene-editing tools can induce indels at the target site, we concluded that CRISPR/Cas9 activity is responsible for these mutations.

A robust and consensus delivery strategy that can guide investigators interested in studying the mechanism of CRISPR/Cas9 uptake and its subsequent action is lacking. Some studies outline mechanisms and formats for delivering gene editing tools, but none of them provide a primary quantitative analysis of efficiency of delivery [117,201,202]. The experimental readout is simply an indirect measure of gene editing activity that, in some cases, could be unrelated to the efficiency of vehicle transfection. The protocols and mode of vehicle delivery used for gene editing are often described with minimal detail that often does not provide experimental evidence

of uptake efficiency that would enable other workers to reproduce or improve upon the effective protocol. In **Chapter 5** the relationship between transfection efficiency and gene editing activity was based on experimental and visual data evaluated. These data suggest that gene editing activity is facilitated when a significant number of RNP particles have entered an individual cell and that high levels of transfection frequency in the absence of highly intense individual cells bearing the RNP is not a sufficient nor predictable measure of the outcome of a gene editing reaction. These observations lead to the hypothesis that there is no direct correlation between efficient cellular uptake and genome modification directed by an RNP.

By understanding the mechanisms by which CRISPR/Cas executes gene editing in human cells, a more efficacious and potential approach to drug development can be undertaken. In **Chapter 6** and **7** we explore the application of the CRISPR gene editing system in two different approaches to study pediatric Leukemia. The first application, **Chapter 6**, pediatric patient specific ALL chromosomal translocation were re-created. Chromosomal translocations are a hallmark of cancer cells which have been shown to result from mis-repair of simultaneous double-strand breaks (DSBs) on two different chromosomes through NHEJ [203–206]. The evidence that DSBs on two different chromosomes can cause translocations came from studies in which recognition sites for specific nucleases were introduced into two different chromosomes [203]. Engineered nuclease platforms have been used to engineer translocations and chromosomal rearrangements of various cancers [138,205]. We find that the use of CRISPR-Cas9 can provide a convenient platform for modeling of cancer-related genetic mutations in vitro. These studies, which model how leukemia might occur in humans through the generation of patient specific translocations

involving the endogenous genes can provide the genetic foundation for studying leukemogenesis. The second application, **Chapter 7**, presents the use of a novel gene editing approach to create expression vectors that harbor patient specific mutations. In this approach we studied a subtype of pediatric AML that evades treatment and presents poor prognosis, through the involvement of FLT3 mutations. Therefore, this type of AML needs relevant cell-based models to enable the discovery of new therapeutic targets through the screening of anti-leukemic drugs designed to reverse the negative prognosis encountered in infants and adults with leukemia. We have developed a diagnostic system to monitor the impact of mutant FLT3 ITDs on the progression of oncogenesis and to evaluate the efficacy of novel AML drugs.

REFERENCES

1. Parekh-Olmedo H, Ferrara L, Brachman E, Kmiec EB (2005) Gene therapy progress and prospects: targeted gene repair. Gene Ther 12: 639–646. Available: http://dx.doi.org/10.1038/sj.gt.3302511. Accessed 12 August 2014.

2. Tomanin R, Scarpa M (2004) Why Do We Need New Gene Therapy Viral Vectors? Characteristics, Limitations and Future Perspectives of Viral Vector Transduction. Curr Gene Ther 4: 357–372. Available: http://www.ingentaconnect.com/content/ben/cgt/2004/00000004/00000004/art00002. Accessed 27 March 2015.

3. Parekh-Olmedo H, Kmiec EB (2007) Progress and prospects: targeted gene alteration (TGA). Gene Ther 14: 1675–1680. Available: http://www.ncbi.nlm.nih.gov/pubmed/17972921. Accessed 20 February 2014.

4. Engstrom JU, Suzuki T, Kmiec EB (2009) Regulation of targeted gene repair by intrinsic cellular processes. Bioessays 31: 159–168. Available: http://www.ncbi.nlm.nih.gov/pubmed/19204988. Accessed 2 July 2014.

5. Aarts M, te Riele H (2011) Progress and prospects: oligonucleotide-directed gene modification in mouse embryonic stem cells: a route to therapeutic application. Gene Ther 18: 213–219. Available: http://www.ncbi.nlm.nih.gov/pubmed/21160530. Accessed 24 February 2014.

6. Brachman EE, Kmiec EB (2005) Gene repair in mammalian cells is stimulated by the elongation of S phase and transient stalling of replication forks. DNA Repair (Amst) 4: 445–457. Available: http://www.ncbi.nlm.nih.gov/pubmed/15725625. Accessed 12 April 2013.

7. Olsen PA, Randol M, Luna L, Brown T, Krauss S (2005) Genomic sequence correction by single-stranded DNA oligonucleotides: role of DNA synthesis and chemical modifications of the oligonucleotide ends. J Gene Med 7: 1534–1544. Available: http://www.ncbi.nlm.nih.gov/pubmed/16025558. Accessed 2 July 2014.

8. Ferrara L, Parekh-Olmedo H, Kmiec EB (2004) Enhanced oligonucleotide-directed gene targeting in mammalian cells following treatment with DNA damaging agents. Exp Cell Res 300: 170–179. Available:

http://www.ncbi.nlm.nih.gov/pubmed/15383324. Accessed 26 February 2014.

9. Radecke F, Peter I, Radecke S, Gellhaus K, Schwarz K, et al. (2006) Targeted chromosomal gene modification in human cells by single-stranded oligodeoxynucleotides in the presence of a DNA double-strand break. Mol Ther 14: 798–808. Available: http://www.ncbi.nlm.nih.gov/pubmed/16904944. Accessed 14 February 2014.

10. Carroll D (2011) Genome engineering with zinc-finger nucleases. Genetics 188: 773–782. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3176093&tool=pmcentrez&rendertype=abstract. Accessed 13 February 2014.

11. Mak AN-S, Bradley P, Bogdanove AJ, Stoddard BL (2013) TAL effectors: function, structure, engineering and applications. Curr Opin Struct Biol 23: 93–99. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3572262&tool=pmcentrez&rendertype=abstract. Accessed 19 February 2014.

12. Cho SW, Kim S, Kim JM, Kim J-S (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31: 230–232. Available: http://dx.doi.org/10.1038/nbt.2507. Accessed 23 May 2014.

13. Mali P, Yang L, Esvelt KM, Aach J, Guell M, et al. (2013) RNA-guided human genome engineering via Cas9. Science 339: 823–826. Available: http://www.sciencemag.org/content/339/6121/823. Accessed 23 May 2014.

14. Wood AJ, Lo T-W, Zeitler B, Pickle CS, Ralston EJ, et al. (2011) Targeted genome editing across species using ZFNs and TALENs. Science 333: 307. Available: http://www.sciencemag.org/content/333/6040/307. Accessed 28 May 2014.

15. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, et al. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 29: 149–153. Available: http://dx.doi.org/10.1038/nbt.1775. Accessed 8 June 2014.

16. Wright D a, Li T, Yang B, Spalding MH (2014) TALEN-mediated genome editing: prospects and perspectives. Biochem J 462: 15–24. Available: http://www.ncbi.nlm.nih.gov/pubmed/25057889. Accessed 28 July 2014.

17. Kim H, Kim J-S (2014) A guide to genome engineering with programmable nucleases. Nat Rev Genet 15: 321–334. Available: http://dx.doi.org/10.1038/nrg3686. Accessed 13 July 2014.

18. Gaj T, Gersbach CA, Barbas 3rd CF, Barbas CF (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31: 397–405. Available: http://www.cell.com/article/S0167779913000875/fulltext. Accessed 9 July 2014.

19. Hsu PD, Lander ES, Zhang F (2014) Development and Applications of CRISPR-Cas9 for Genome Engineering. Cell 157: 1262–1278. Available: http://www.cell.com/article/S0092867414006047/fulltext. Accessed 9 July 2014.

20. Yang L, Guell M, Byrne S, Yang JL, De Los Angeles A, et al. (2013) Optimization of scarless human stem cell genome editing. Nucleic Acids Res 41: 9049–9061. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3799423&tool=pmcentrez&rendertype=abstract. Accessed 25 January 2014.

21. Sun N, Liang J, Abil Z, Zhao H (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 8: 1255–1263. Available: http://pubs.rsc.org/en/content/articlehtml/2012/mb/c2mb05461b. Accessed 2 July 2014.

22. Agrawal RK, Patel RK, Shah V, Nainiwal L, Trivedi B (2014) Hydroxyurea in sickle cell disease: drug review. Indian J Hematol Blood Transfus 30: 91–96. Available: http://www.ncbi.nlm.nih.gov/pubmed/24839362. Accessed 5 August 2014.

23. Creary SE, Gladwin MT, Byrne M, Hildesheim M, Krishnamurti L (2014) A pilot study of electronic directly observed therapy to improve hydroxyurea adherence in pediatric patients with sickle-cell disease. Pediatr Blood Cancer 61: 1068–1073. Available: http://www.ncbi.nlm.nih.gov/pubmed/24436121. Accessed 1 August 2014.

24. Vannocci T, Kurata H, de la Fuente J, Roberts IA, Porter ACG (2014) Nuclease-stimulated homologous recombination at the human β-globin gene. J Gene Med 16: 1–10. Available: http://doi.wiley.com/10.1002/jgm.2751. Accessed 5 August 2014.

25. Yin H, Xue W, Chen S, Bogorad RL, Benedetti E, et al. (2014) Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol 32: 551–553. Available: http://dx.doi.org/10.1038/nbt.2884. Accessed 9 July 2014.

26. Patel N, Reiss U, Davidoff AM, Nathwani AC (2014) Progress towards gene

therapy for haemophilia B. Int J Hematol 99: 372–376. Available: http://www.ncbi.nlm.nih.gov/pubmed/24500177. Accessed 10 July 2014.

27. Tebas P, Stein D, Tang WW, Frank I, Wang SQ, et al. (2014) Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. N Engl J Med 370: 901–910. Available: http://www.ncbi.nlm.nih.gov/pubmed/24597865.

28. Wright AF (2015) Long-Term Effects of Retinal Gene Therapy in Childhood Blindness. N Engl J Med 372: 1954–1955. Available: http://www.nejm.org/doi/10.1056/NEJMe1503419. Accessed 1 May 2019.

29. Fischer A (2017) FDA approves novel gene therapy to treat patients with a rare form of inherited vision loss. Available: https://www.fda.gov/news-events/press-announcements/fda-approves-novel-gene-therapy-treat-patients-rare-form-inherited-vision-loss. Accessed 2 May 2019.

30. Second generation CFTR gene repair (n.d.). Available: https://www.cysticfibrosis.org.uk/the-work-we-do/research/research-areas/gene-therapy/second-generation-cftr-gene-repair#. Accessed 5 February 2019.

31. Yao S, He Z, Chen C (2015) CRISPR/Cas9-Mediated Genome Editing of Epigenetic Factors for Cancer Therapy. Hum Gene Ther 26: 463–471. Available: http://www.liebertpub.com/doi/10.1089/hum.2015.067. Accessed 5 February 2019.

32. Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS (2015) Cas9-mediated targeting of viral RNA in eukaryotic cells. Proc Natl Acad Sci U S A 112: 6164–6169. Available: http://www.pnas.org/lookup/doi/10.1073/pnas.1422340112. Accessed 5 February 2019.

33. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, et al. (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science (80-) 353: aaf5573. Available: http://www.ncbi.nlm.nih.gov/pubmed/27256883. Accessed 5 February 2019.

34. Papaioannou I, Disterer P, Owen JS (2009) Use of internally nuclease-protected single-strand DNA oligonucleotides and silencing of the mismatch repair protein, MSH2, enhances the replication of corrected cells following gene editing. J Gene Med 11: 267–274. Available: http://www.ncbi.nlm.nih.gov/pubmed/19153972. Accessed 2 July 2014.

35. Liu L, Rice MC, Drury M, Cheng S, Gamper H, et al. (2002) Strand Bias in

Targeted Gene Repair Is Influenced by Transcriptional Activity. Mol Cell Biol 22: 3852–3863. doi:10.1128/MCB.22.11.3852.

36. Brachman EE, Kmiec EB (2002) The “biased” evolution of targeted gene repair. Curr Opin Mol Ther 4: 171–176. Available: http://europepmc.org/abstract/MED/12044039. Accessed 5 August 2014.

37. Strouse B, Bialk P, Niamat R a, Rivera-Torres N, Kmiec EB (2014) Combinatorial gene editing in mammalian cells using ssODNs and TALENs. Sci Rep 4: 3791. Available: http://www.ncbi.nlm.nih.gov/pubmed/24445749. Accessed 22 January 2014.

38. Olsen PA, Randol M, Krauss S (2005) Implications of cell cycle progression on functional sequence correction by short single-stranded DNA oligonucleotides. Gene Ther 12: 546–551. Available: http://dx.doi.org/10.1038/sj.gt.3302454. Accessed 2 July 2014.

39. Ferrara L, Kmiec EB (2004) Camptothecin enhances the frequency of oligonucleotide-directed gene repair in mammalian cells by inducing DNA damage and activating homologous recombination. Nucleic Acids Res 32: 5239–5248. Available: http://nar.oxfordjournals.org/content/32/17/5239.long. Accessed 17 June 2014.

40. Engstrom JU, Kmiec EB (2008) DNA replication, cell cycle progression and the targeted gene repair reaction. Cell Cycle: 1402–1414.

41. Huen MSY, Li X, Lu L-Y, Watt RM, Liu D-P, et al. (2006) The involvement of replication in single stranded oligonucleotide-mediated gene repair. Nucleic Acids Res 34: 6183–6194. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1693898&tool=pmcentrez&rendertype=abstract. Accessed 26 February 2014.

42. Engstrom JU, Kmiec EB (2007) Manipulation of cell cycle progression can counteract the apparent loss of correction frequency following oligonucleotide-directed gene repair. BMC Mol Biol 8: 9. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1797188&tool=pmcentrez&rendertype=abstract. Accessed 26 February 2014.

43. Aarts M, te Riele H (2010) Parameters of oligonucleotide-mediated gene modification in mouse ES cells. J Cell Mol Med 14: 1657–1667. doi:10.1111/j.1582-4934.2009.00847.x.

44. Rivera-Torres N, Strouse B, Bialk P, Niamat RA, Kmiec EB (2014) The position of DNA cleavage by TALENs and cell synchronization influences the

frequency of gene editing directed by single-stranded oligonucleotides. PLoS One 9: e96483. Available: http://www.ncbi.nlm.nih.gov/pubmed/24788536. Accessed 5 May 2014.

45. Rivera-Torres N, Banas K, Bialk P, Bloh KM, Kmiec EB (2017) Insertional Mutagenesis by CRISPR/Cas9 Ribonucleoprotein Gene Editing in Cells Targeted for Point Mutation Repair Directed by Short Single-Stranded DNA Oligonucleotides. PLoS One 12: e0169350. Available: http://dx.plos.org/10.1371/journal.pone.0169350. Accessed 17 January 2017.

46. Bonner M, Strouse B, Applegate M, Livingston P, Kmiec EB (2012) DNA damage response pathway and replication fork stress during oligonucleotide directed gene editing. Mol Ther Nucleic Acids 1: e18. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3381643&tool=pmcentrez&rendertype=abstract. Accessed 30 March 2013.

47. Chen S, Oikonomou G, Chiu CN, Niles BJ, Liu J, et al. (2013) A large-scale in vivo analysis reveals that TALENs are significantly more mutagenic than ZFNs generated using context-dependent assembly. Nucleic Acids Res 41: 2769–2778. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3575824&tool=pmcentrez&rendertype=abstract. Accessed 23 January 2014.

48. Lamb BM, Mercer AC, Barbas CF (2013) Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. Nucleic Acids Res 41: 9779–9785. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3834825&tool=pmcentrez&rendertype=abstract. Accessed 25 February 2014.

49. Briggs AW, Rios X, Chari R, Yang L, Zhang F, et al. (2012) Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res 40: e117. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3424587&tool=pmcentrez&rendertype=abstract. Accessed 23 February 2014.

50. Wefers B, Meyer M, Ortiz O, Hrabé de Angelis M, Hansen J, et al. (2013) Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. Proc Natl Acad Sci U S A 110: 3782–3787. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3593923&tool=pmcentrez&rendertype=abstract. Accessed 20 January 2014.

51. Ding Q, Lee Y-K, Schaefer E a K, Peters DT, Veres A, et al. (2013) A TALEN

genome-editing system for generating human stem cell-based disease models. Cell Stem Cell 12: 238–251. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3570604&tool=pmcentrez&rendertype=abstract. Accessed 28 May 2014.

52. Bedell VM, Wang Y, Campbell JM, Poshusta TL, Starker CG, et al. (2012) In vivo genome editing using a high-efficiency TALEN system. Nature 491: 114–118. Available: http://dx.doi.org/10.1038/nature11537. Accessed 26 May 2014.

53. Liu J, Majumdar A, Liu J, Thompson LH, Seidman MM (2010) Sequence conversion by single strand oligonucleotide donors via non-homologous end joining in mammalian cells. J Biol Chem 285: 23198–23207. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2906313&tool=pmcentrez&rendertype=abstract. Accessed 26 February 2014.

54. Cong L, Ran FA, Cox D, Lin S, Barretto R, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339: 819–823. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3795411&tool=pmcentrez&rendertype=abstract. Accessed 9 July 2014.

55. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, et al. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31: 227–229. Available: http://dx.doi.org/10.1038/nbt.2501. Accessed 9 July 2014.

56. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA (2013) RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31: 233–239. Available: http://dx.doi.org/10.1038/nbt.2508. Accessed 9 July 2014.

57. Jinek M, East A, Cheng A, Lin S, Ma E, et al. (2013) RNA-programmed genome editing in human cells. Elife 2: e00471. Available: http://elifesciences.org/content/2/e00471.abstract. Accessed 10 July 2014.

58. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature 507: 62–67. Available: http://www.ncbi.nlm.nih.gov/pubmed/24476820. Accessed 2 August 2018.

59. Her J, Bunting SF (n.d.) How cells ensure correct repair of DNA double-strand breaks. Available: http://www.jbc.org/cgi/doi/10.1074/jbc.TM118.000371. Accessed 1 April 2019.

60. Brandsma I, Gent DC (2012) Pathway choice in DNA double strand break repair: observations of a balancing act. Genome Integr 3: 9. Available: http://www.ncbi.nlm.nih.gov/pubmed/23181949. Accessed 1 April 2019.

61. Selle K, Barrangou R (2015) CRISPR-Based Technologies and the Future of Food Science. J Food Sci 80: R2367–R2372. Available: http://doi.wiley.com/10.1111/1750-3841.13094. Accessed 1 April 2019.

62. Niler E (2018) How Crispr Could Transform Our Food Supply. Available: https://www.nationalgeographic.com/environment/future-of-food/food-technology-gene-editing/. Accessed 1 April 2019.

63. Donohoue PD, Barrangou R, May AP (2018) Advances in Industrial Biotechnology Using CRISPR-Cas Systems. Trends Biotechnol 36: 134–146. Available: https://www.sciencedirect.com/science/article/pii/S0167779917301877. Accessed 1 April 2019.

64. Charpentier E (2015) CRISPR-Cas9: how research on a bacterial RNA-guided mechanism opened new perspectives in biotechnology and biomedicine. EMBO Mol Med 7: 363–365. Available: http://www.ncbi.nlm.nih.gov/pubmed/25796552. Accessed 1 April 2019.

65. Niu Y, Li T, Ji W (2017) Paving the road for biomedicine: genome editing and stem cells in primates. Natl Sci Rev 4: 543–549. Available: https://academic.oup.com/nsr/article/4/4/543/4093906. Accessed 1 April 2019.

66. Ahmed Khan F, Shinta Pandupuspitasari N, ChunJie H, Ishfaq Ahmad H, Wang K, et al. (n.d.) UNCORRECTED PROOF Applications of CRISPR/Cas9 in Reproductive Biology. Available: https://doi.org/10.21775/9781910190630.08. Accessed 1 April 2019.

67. Cavaliere G (2018) Genome editing and assisted reproduction: curing embryos, society or prospective parents? Med Heal Care Philos 21: 215–225. Available: http://link.springer.com/10.1007/s11019-017-9793-y. Accessed 1 April 2019.

68. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol advance on. Available: http://dx.doi.org/10.1038/nbt.2842. Accessed 2 March 2014.

69. Boroviak K, Doe B, Banerjee R, Yang F, Bradley A (2016) Chromosome engineering in zygotes with CRISPR/Cas9. Genesis 54: 78–85. Available: http://doi.wiley.com/10.1002/dvg.22915. Accessed 5 February 2019.

70. Dow LE (2015) Modeling Disease In Vivo With CRISPR/Cas9. Trends Mol Med 21: 609–621. Available: http://www.ncbi.nlm.nih.gov/pubmed/26432018. Accessed 5 February 2019.

71. Sheridan C (2017) CRISPR therapeutics push into human testing. Nat Biotechnol 35: 3–5. Available: http://www.nature.com/doifinder/10.1038/nbt0117-3. Accessed 1 April 2019.

72. Baylis F, McLeod M (2017) First-in-human Phase 1 CRISPR Gene Editing Cancer Trials: Are We Ready? Curr Gene Ther 17: 309–319. Available: http://www.ncbi.nlm.nih.gov/pubmed/29173170. Accessed 1 April 2019.

73. Musunuru K (2013) Genome editing of human pluripotent stem cells to generate human cellular disease models. Dis Model Mech 6: 896–904. Available: http://www.ncbi.nlm.nih.gov/pubmed/23751357. Accessed 5 February 2019.

74. Peacock E (n.d.) Gene Therapy: Bringing Hope to the Rare Disease Community — Rare Disease Review. Available: https://www.rarediseasereview.org/publications/2018/3/12/gene-therapy-bringing-hope-to-the-rare-disease-community. Accessed 1 April 2019.

75. Rossi K (2018) Developing Personalized CRISPR-Cas9 Genome Editing Therapy for Pompe Disease | MD Magazine. Available: https://www.mdmag.com/medical-news/developing-personalized-crisprcas9-genome-editing-therapy-for-pompe-disease. Accessed 1 April 2019.

76. Boycott KM, Vanstone MR, Bulman DE, MacKenzie AE (2013) Rare-disease genetics in the era of next-generation sequencing: discovery to translation. Nat Rev Genet 14: 681–691. Available: http://www.nature.com/articles/nrg3555. Accessed 1 April 2019.

77. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, et al. (2011) Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12: 745–755. Available: http://www.nature.com/articles/nrg3031. Accessed 1 April 2019.

78. Baliou S, Adamaki M, Kyriakopoulos A, Spandidos D, Panagiotidis M, et al. (2018) CRISPR therapeutic tools for complex genetic disorders and cancer (Review). Int J Oncol 53: 443–468. Available: http://www.spandidos-publications.com/10.3892/ijo.2018.4434. Accessed 5 April 2019.

79. Cai L, Fisher AL, Huang H, Xie Z (2016) CRISPR-mediated genome editing and human diseases. Genes Dis 3: 244. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6150104/. Accessed 5 April 2019.

80. Picher-Martel V, Valdmanis PN, Gould P V., Julien JP, Dupré N (2016) From

animal models to human disease: A genetic approach for personalized medicine in ALS. Acta Neuropathol Commun 4: 70. Available: http://actaneurocomms.biomedcentral.com/articles/10.1186/s40478-016-0340-5. Accessed 5 April 2019.

81. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, et al. (2016) Rationally engineered Cas9 nucleases with improved specificity. Science (80-) 351: 84–88. Available: http://www.ncbi.nlm.nih.gov/pubmed/26628643.

82. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, et al. (2016) High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature advance on. Available: http://dx.doi.org/10.1038/nature16526. Accessed 6 January 2016.

83. Ma H, Tu L-C, Naseri A, Huisman M, Zhang S, et al. (2016) Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nat Biotechnol 34: 528–530. Available: http://www.nature.com/articles/nbt.3526. Accessed 5 February 2019.

84. Structure et instabilité des génomes - Sorbonne Universités - MNHN - CNRS 7196 / INSERM U1154 | Genome Editing, DNA double-strand break Repair and cellular Responses (GE2R) (n.d.). Available: https://biophysique.mnhn.fr/site/Genome+Editing,+DNA+double-strand+break+Repair+and+cellular+Responses+(GE2R). Accessed 30 April 2019.

85. Bonner M, Kmiec EB (2009) DNA breakage associated with targeted gene alteration directed by DNA oligonucleotides. Mutat Res 669: 85–94. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2749079&tool=pmcentrez&rendertype=abstract. Accessed 26 February 2014.

86. Radecke S, Radecke F, Peter I, Schwarz K (2006) Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. J Gene Med 8: 217–228. Available: http://www.ncbi.nlm.nih.gov/pubmed/16142817. Accessed 23 June 2014.

87. Pierce EA, Liu Q, Igoucheva O, Omarrudin R, Ma H, et al. (2003) Oligonucleotide-directed single-base DNA alterations in mouse embryonic stem cells. Gene Ther 10: 24–33. Available: http://dx.doi.org/10.1038/sj.gt.3301857. Accessed 2 March 2015.

88. Bertoni C, Rustagi A, Rando TA (2009) Enhanced gene repair mediated by methyl-CpG-modified single-stranded oligonucleotides. Nucleic Acids Res 37: 7468–7482. Available:

http://nar.oxfordjournals.org/content/early/2009/10/23/nar.gkp757.short. Accessed 12 August 2014.

89. Yamamoto T, Moerschell RP, Wakem LP, Komar-Panicucci S, Sherman F (1992) Strand-specificity in the transformation of yeast with synthetic oligonucleotides. Genetics 131: 811–819. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1205094&tool=pmcentrez&rendertype=abstract. Accessed 2 March 2015.

90. Bertoni C, Morris GE, Rando TA (2004) Strand bias in oligonucleotide-mediated dystrophin gene editing. Hum Mol Genet 14: 221–233. Available: http://hmg.oxfordjournals.org/content/14/2/221. Accessed 2 March 2015.

91. Ferrara L, Kmiec EB (2006) Targeted gene repair activates Chk1 and Chk2 and stalls replication in corrected cells. DNA Repair (Amst) 5: 422–431. Available: http://www.sciencedirect.com/science/article/pii/S1568786405003186. Accessed 12 August 2014.

92. Ferrara L, Engstrom JU, Schwartz T, Parekh-Olmedo H, Kmiec EB (2007) Recovery of cell cycle delay following targeted gene repair by oligonucleotides. DNA Repair (Amst) 6: 1529–1535. Available: http://www.sciencedirect.com/science/article/pii/S1568786407001735. Accessed 12 August 2014.

93. Lin S, Staahl B, Alla RK, Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. Elife 3: e04766. Available: http://elifesciences.org/content/3/e04766.abstract. Accessed 18 December 2014.

94. Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, et al. (2004) Mutation detection using Surveyor^{TM} nuclease. Biotechniques 36: 702–707.

95. Brinkman EK, Chen T, Amendola M, van Steensel B (2014) Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 42: e168-. Available: http://nar.oxfordjournals.org/content/42/22/e168.short. Accessed 11 October 2014.

96. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, et al. (2013) Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. Cell 154: 1380–1389. Available: http://www.ncbi.nlm.nih.gov/pubmed/23992846. Accessed 10 July 2014.

97. Majumdar A, Puri N, Cuenoud B, Natt F, Martin P, et al. (2003) Cell cycle

modulation of gene targeting by a triple helix-forming oligonucleotide. J Biol Chem 278: 11072–11077. Available: http://www.jbc.org/content/278/13/11072.short. Accessed 12 August 2014.

98. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, et al. (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8: 2281–2308. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3969860&tool=pmcentrez&rendertype=abstract. Accessed 9 July 2014.

99. Davis L, Maizels N (2014) Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. Proc Natl Acad Sci U S A 111: E924-32. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3956201&tool=pmcentrez&rendertype=abstract. Accessed 28 July 2014.

100. Wu X-S, Xin L, Yin W-X, Shang X-Y, Lu L, et al. (2005) Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. Proc Natl Acad Sci U S A 102: 2508–2513. Available: http://www.pnas.org/content/102/7/2508.short. Accessed 12 August 2014.

101. Bialk P, Rivera-Torres N, Strouse B, Kmiec EB (2015) Regulation of Gene Editing Activity Directed by Single-Stranded Oligonucleotides and CRISPR/Cas9 Systems. PLoS One 10: e0129308. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4459703&tool=pmcentrez&rendertype=abstract. Accessed 9 June 2015.

102. Hu Y, Parekh-Olmedo H, Drury M, Skogen M, Kmiec EB (2005) Reaction parameters of targeted gene repair in mammalian cells. Mol Biotechnol 29: 197–210. Available: http://www.ncbi.nlm.nih.gov/pubmed/15767697.

103. Liu L, Parekh-Olmedo H, Kmiec EB (2003) The development and regulation of gene repair. Nat Rev Genet 4: 679–689. Available: http://dx.doi.org/10.1038/nrg1156. Accessed 12 August 2014.

104. Vouillot L, Thélie A, Pollet N (2015) Comparison of T7E1 and Surveyor Mismatch Cleavage Assays To Detect Mutations Triggered by Engineered Nucleases. G3 (Bethesda) 5: 407–415. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4349094&tool=pmcentrez&rendertype=abstract. Accessed 3 February 2015.

105. Schumann K, Lin S, Boyer E, Simeonov DR, Subramaniam M, et al. (2015) Generation of knock-in primary human T cells using. Proc Natl Acad Sci U S A 112: 10437–10442. Available:

http://www.pnas.org/content/112/33/10437.abstract. Accessed 28 July 2015.

106. Nur-E-Kamal A, Li T-K, Zhang A, Qi H, Hars ES, et al. (2003) Single-stranded DNA induces ataxia telangiectasia mutant (ATM)/p53-dependent DNA damage and apoptotic signals. J Biol Chem 278: 12475–12481. Available: http://www.ncbi.nlm.nih.gov/pubmed/12540848. Accessed 22 April 2015.

107. Olsen PA, Solhaug A, Booth JA, Gelazauskaite M, Krauss S (2009) Cellular responses to targeted genomic sequence modification using single-stranded oligonucleotides and zinc-finger nucleases. DNA Repair (Amst) 8: 298–308. Available: http://www.sciencedirect.com/science/article/pii/S1568786408003959. Accessed 6 August 2014.

108. Kmiec EB (2015) Is the age of genetic surgery finally upon us? Surg Oncol. Available: http://www.sciencedirect.com/science/article/pii/S0960740415000249. Accessed 8 May 2015.

109. Rivera-Torres N, Kmiec EB (2015) Genetic spell-checking: gene editing using single-stranded DNA oligonucleotides. Plant Biotechnol J: n/a-n/a. Available: http://doi.wiley.com/10.1111/pbi.12473.

110. Wang Z, Zhou Z-J, Liu D-P, Huang J-D (2006) Single-stranded oligonucleotide-mediated gene repair in mammalian cells has a mechanism distinct from homologous recombination repair. Biochem Biophys Res Commun 350: 568–573. Available: http://www.ncbi.nlm.nih.gov/pubmed/17026965. Accessed 14 December 2015.

111. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, et al. (2015) Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol 33: 538–542. doi:10.1038/nbt.3190.

112. Kouranova E, Forbes K, Zhao G, Warren J, Bartels A, et al. (2016) CRISPRs for Optimal Targeting : Delivery of CRISPR Components as DNA , RNA , and Protein into Cultured Cells and Single-Cell Embryos. 27: 464–475. doi:10.1089/hum.2016.009.

113. Hou Z, Zhang Y, Propson NE, Howden SE, Chu L-F, et al. (2013) Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci 110: 15644–15649. Available: http://www.pnas.org/cgi/doi/10.1073/pnas.1313587110. Accessed 29 August 2016.

114. Merkle FT, Neuhausser WM, Santos D, Valen E, Gagnon JA, et al. (2015) Efficient CRISPR-Cas9-Mediated Generation of Knockin Human Pluripotent Stem Cells Lacking Undesired Mutations at the Targeted Locus. Cell Rep 11: 875–883. Available: http://dx.doi.org/10.1016/j.celrep.2015.04.007.

115. Bialk P, Sansbury B, Rivera-Torres N, Bloh K, Man D, et al. (2016) Analyses of point mutation repair and allelic heterogeneity generated by CRISPR/Cas9 and single-stranded DNA oligonucleotides. Sci Rep 6: 32681. Available: http://www.nature.com/articles/srep32681.

116. Andersen MS, Sorensen CB, Bolund L, Jensen TG (2002) Mechanisms underlying targeted gene correction using chimeric RNA/DNA and single-stranded DNA oligonucleotides. J Mol Med (Berl) 80: 770–781. Available: http://www.ncbi.nlm.nih.gov/pubmed/12483462. Accessed 26 February 2014.

117. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol 34: 339–344. Available: http://www.ncbi.nlm.nih.gov/pubmed/26789497. Accessed 28 July 2016.

118. Paquet D, Kwart D, Chen A, Sproul A, Jacob S, et al. (2016) Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature 533: 125–129. Available: http://www.ncbi.nlm.nih.gov/pubmed/27120160. Accessed 5 October 2016.

119. Valerie K, Povirk LF (2003) Regulation and mechanisms of mammalian double-strand break repair. Oncogene 22: 5792–5812. Available: http://www.ncbi.nlm.nih.gov/pubmed/12947387. Accessed 5 October 2016.

120. Davis AJ, Chen DJ (2013) DNA double strand break repair via non-homologous end-joining. Transl Cancer Res 2: 130–143. Available: http://www.ncbi.nlm.nih.gov/pubmed/24000320. Accessed 5 October 2016.

121. Saintigny Y, Dumay A, Lambert S, Lopez BS (2001) A novel role for the Bcl-2 protein family: specific suppression of the RAD51 recombination pathway. EMBO J 20: 2596–2607. Available: http://www.ncbi.nlm.nih.gov/pubmed/11350949. Accessed 5 October 2016.

122. Saintigny Y, Makienko K, Swanson C, Emond MJ, Monnat RJ (2002) Homologous recombination resolution defect in werner syndrome. Mol Cell Biol 22: 6971–6978. Available: http://www.ncbi.nlm.nih.gov/pubmed/12242278. Accessed 5 October 2016.

123. Radecke F, Radecke S, Schwarz K (2004) Unmodified oligodeoxynucleotides require single-strandedness to induce targeted repair of a chromosomal EGFP gene. J Gene Med 6: 1257–1271. Available: http://www.ncbi.nlm.nih.gov/pubmed/15459968. Accessed 12 August 2014.

124. Rivera-Torres N, Kmiec EB (2017) A Standard Methodology to Examine On-site Mutagenicity As a Function of Point Mutation Repair Catalyzed by CRISPR/Cas9 and SsODN in Human Cells. J Vis Exp: e56195–e56195. Available: https://www.jove.com/video/56195/a-standard-methodology-to-examine-on-site-mutagenicity-as-function. Accessed 28 August 2017.

125. Meyer C, Hofmann J, Burmeister T, Gröger D, Park TS, et al. (2013) The MLL recombinome of acute leukemias in 2013. Leukemia 27: 2165–2176. Available: http://www.nature.com/doifinder/10.1038/leu.2013.135. Accessed 21 February 2017.

126. Santoli D, Yang YC, Clark SC, Kreider BL, Caracciolo D, et al. (1987) Synergistic and antagonistic effects of recombinant human interleukin (IL) 3, IL-1 alpha, granulocyte and macrophage colony-stimulating factors (G-CSF and M-CSF) on the growth of GM-CSF-dependent leukemic cell lines. J Immunol 139: 3348–3354. Available: http://www.ncbi.nlm.nih.gov/pubmed/3500218. Accessed 12 February 2019.

127. Lange B, Valtieri M, Santoli D, Caracciolo D, Mavilio F, et al. (1987) Growth factor requirements of childhood acute leukemia: establishment of GM-CSF-dependent cell lines. Blood 70: 192–199. Available: http://www.ncbi.nlm.nih.gov/pubmed/3496132. Accessed 12 February 2019.

128. Romano P, Manniello A, Aresu O, Armento M, Cesaro M, et al. (2009) Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. Nucleic Acids Res 37: D925-32. Available: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkn730. Accessed 12 February 2019.

129. Reichel M, Gillert E, Angermüller S, Hensel JP, Heidel F, et al. (2001) Biased distribution of chromosomal breakpoints involving the MLL gene in infants versus children and adults with t(4;11) ALL. Oncogene 20: 2900–2907. Available: http://www.nature.com/onc/journal/v20/n23/full/1204401a.html#bib14. Accessed 18 September 2015.

130. Marschalek R (2011) Mechanisms of leukemogenesis by MLL fusion proteins. Br J Haematol 152: 141–154. doi:10.1111/j.1365-2141.2010.08459.x.

131. Liu H, Cheng EHY, Hsieh JJD (2009) MLL fusions: pathways to leukemia. Cancer Biol Ther 8: 1204–1211. Available: http://www.ncbi.nlm.nih.gov/pubmed/19729989. Accessed 7 July 2016.

132. Collins EC, Pannell R, Simpson EM, Forster A, Rabbitts TH (2000) Inter-chromosomal recombination of Mll and Af9 genes mediated by cre-loxP in mouse development. EMBO Rep 1: 127–132. Available: http://www.ncbi.nlm.nih.gov/pubmed/11265751. Accessed 14 February 2019.

133. Van Deursen J, Fornerod M, Van Rees B, Grosveld G (1995) Cre-mediated site-specific translocation between nonhomologous mouse chromosomes. Proc Natl Acad Sci U S A 92: 7376–7380. Available: http://www.ncbi.nlm.nih.gov/pubmed/7638200. Accessed 14 February 2019.

134. Smith AJH, De Sousa MA, Kwabi-Addo B, Heppell-Parton A, Impey H, et al. (1995) A site–directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. Nat Genet 9: 376–385. Available: http://www.ncbi.nlm.nih.gov/pubmed/7795643. Accessed 14 February 2019.

135. Jiang J, Zhang L, Zhou X, Chen X, Huang G, et al. (2016) Induction of site-specific chromosomal translocations in embryonic stem cells by CRISPR/Cas9. Sci Rep 6: 21918. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4761995&tool=pmcentrez&rendertype=abstract. Accessed 29 February 2016.

136. Reimer J, Knoess S, Labuhn M, Charpentier EM, Göhring G, et al. (2017) CRISPR-Cas9-induced t(11;19)/MLL-ENL translocations initiate leukemia in human hematopoietic progenitor cells in vivo. Haematologica: haematol.2017.164046. Available: http://www.haematologica.org/lookup/doi/10.3324/haematol.2017.164046.

137. Lagutina I V, Valentine V, Picchione F, Harwood F, Valentine MB, et al. (2015) Modeling of the human alveolar rhabdomyosarcoma Pax3-Foxo1 chromosome translocation in mouse myoblasts using CRISPR-Cas9 nuclease. PLoS Genet 11: e1004951. Available: http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004951. Accessed 15 October 2015.

138. Choi PS, Meyerson M (2014) Targeted genomic rearrangements using CRISPR/Cas technology. Nat Commun 5: 3728. Available: http://www.nature.com/ncomms/2014/140424/ncomms4728/abs/ncomms4728.html. Accessed 28 February 2015.

139. Ghezraoui H, Piganeau M, Renouf B, Renaud J-B, Sallmyr A, et al. (2014)

Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell 55: 829–842. Available: http://www.sciencedirect.com/science/article/pii/S1097276514006352. Accessed 24 November 2015.

140. Meshinchi S, Arceci RJ (2007) Prognostic factors and risk-based therapy in pediatric acute myeloid leukemia. Oncologist 12: 341–355. Available: http://www.ncbi.nlm.nih.gov/pubmed/17405900. Accessed 18 September 2017.

141. Lee BH, Tothova Z, Levine RL, Anderson K, Buza-Vidas N, et al. (2007) FLT3 Mutations Confer Enhanced Proliferation and Survival Properties to Multipotent Progenitors in a Murine Model of Chronic Myelomonocytic Leukemia. Cancer Cell 12: 367–380. Available: http://ac.els-cdn.com/S1535610807002632/1-s2.0-S1535610807002632-main.pdf?_tid=b404dab6-035b-11e7-93e5-00000aab0f26&acdnat=1488907985_8882395aca7bcce61ed6888fab498abb. Accessed 7 March 2017.

142. Sakamoto KM, Grant S, Saleiro D, Crispino J, Hijiya N, et al. (n.d.) Targeting Novel Signaling Pathways for Resistant Acute Myeloid Leukemia. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4355162/pdf/nihms-647532.pdf. Accessed 31 October 2017.

143. de Rooij J, Zwaan C, van den Heuvel-Eibrink M, van den Heuvel-Eibrink M (2015) Pediatric AML: From Biology to Clinical Management. J Clin Med 4: 127–149. Available: http://www.mdpi.com/2077-0383/4/1/127/. Accessed 9 August 2018.

144. Gary Gilliland D, Griffin JD (2002) The roles of FLT3 in hematopoiesis and leukemia. Blood 100: 1532–1542. Available: http://www.bloodjournal.org/content/bloodjournal/100/5/1532.full.pdf. Accessed 9 October 2017.

145. Cauchy P, James SR, Zacarias-Cabeza J, Ptasinska A, Imperato MR, et al. (2015) Chronic FLT3-ITD Signaling in Acute Myeloid Leukemia Is Connected to a Specific Chromatin Signature. Cell Rep 12: 821–836. Available: http://ac.els-cdn.com/S221112471500707X/1-s2.0-S221112471500707X-main.pdf?_tid=5f18cce8-0b43-11e7-bc69-00000aacb35e&acdnat=1489777144_5d2d7e55a76965b004925f52ba63f5c6. Accessed 17 March 2017.

146. Leung AYHH, Man C-HH, Kwong Y-LL (2013) FLT3 inhibition: A moving and evolving target in acute myeloid leukaemia. Leukemia 27: 260–268.

Available: http://www.ncbi.nlm.nih.gov/pubmed/22797419. Accessed 5 May 2017.

147. Stirewalt DL, Radich JP (2003) The role of FLT3 in haematopoietic malignancies. Nat Rev Cancer 3: 650–665. Available: http://www.ncbi.nlm.nih.gov/pubmed/12951584. Accessed 5 May 2017.

148. Heidel F, Solem FK, Breitenbuecher F, Lipka DB, Kasper S, et al. (2005) Clinical resistance to the kinase inhibitor PKC412 in acute myeloid leukemia by mutation of Asn-676 in the FLT3 tyrosine kinase domain. Blood 107. Available: http://www.bloodjournal.org/content/107/1/293.long. Accessed 12 September 2017.

149. Fathi AT, Chen Y-B (2011) Treatment of FLT3-ITD acute myeloid leukemia. Am J Blood Res 1: 175–189. Available: http://www.ncbi.nlm.nih.gov/pubmed/22432079. Accessed 5 May 2017.

150. Kottaridis PD, Gale RE, Frew ME, Harrison G, Langabeer SE, et al. (2001) The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: Analysis of 854 patients from the United King. Blood 98: 1752–1759. Available: http://www.bloodjournal.org/cgi/doi/10.1182/blood.V98.6.1752. Accessed 19 April 2017.

151. Small D (2006) FLT3 Mutations: Biology and Treatment. Hematology 2006: 178–184. Available: http://www.ncbi.nlm.nih.gov/pubmed/17124058. Accessed 5 May 2017.

152. Libura M, Asnafi V, Tu A, Delabesse E, Tigaud I, et al. (2003) FLT3 and MLL intragenic abnormalities in AML reflect a common category of genotoxic stress. Blood 102: 2198–2204. Available: http://www.ncbi.nlm.nih.gov/pubmed/12791658. Accessed 5 May 2017.

153. Hartmann J, Haap M, Kopp H-G, Lipp H-P (2009) Tyrosine Kinase Inhibitors – A Review on Pharmacology, Metabolism and Side Effects. Curr Drug Metab 10: 470–481. Available: http://www.ncbi.nlm.nih.gov/pubmed/19689244. Accessed 13 September 2017.

154. Kiyoi H (2015) FLT3 INHIBITORS: RECENT ADVANCES AND PROBLEMS FOR CLINICAL APPLICATION. Nagoya J Med Sci 77: 7–17. Available: http://www.ncbi.nlm.nih.gov/pubmed/25797966. Accessed 12 September 2017.

155. Arora A, Scholar EM (n.d.) Role of Tyrosine Kinase Inhibitors in Cancer Therapy. Available: http://jpet.aspetjournals.org/content/jpet/315/3/971.full.pdf. Accessed 13 September 2017.

156. Alvarado Y, Kantarjian HM, Ravandi F, Luthra R, Borthakur G, et al. (2015) FLT3 Inhibitor Treatment in FLT3-Mutated AML Is Associated with Development of Secondary FLT3-TKD Mutations. Blood 118. Available: http://www.bloodjournal.org/content/118/21/1493?sso-checked=true. Accessed 15 September 2017.

157. Brown P, Meshinchi S, Levis M, Alonzo TA, Gerbing R, et al. (2004) Pediatric AML primary samples with FLT3/ITD mutations are preferentially killed by FLT3 inhibition. 104: 1841–1849. Available: www.bloodjournal.org. Accessed 11 April 2019.

158. Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, et al. (2014) Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. Nat Commun 5: 3964. Available: http://www.ncbi.nlm.nih.gov/pubmed/24888982.

159. Li B, Zhao W, Luo X, Zhang X, Li C, et al. (2017) Engineering CRISPR–Cpf1 crRNAs and mRNAs to maximize genome editing efficiency. Nat Biomed Eng 1: 0066. Available: http://www.nature.com/articles/s41551-017-0066. Accessed 9 November 2017.

160. Chylinski K, Le Rhun A, Charpentier E (2013) The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol 10: 726–737. doi:10.4161/rna.24321.

161. Sansbury BM, Wagner AM, Nitzan E, Tarcic G, Kmiec EB (2018) CRISPR-Directed *In Vitro* Gene Editing of Plasmid DNA Catalyzed by Cpf1 (Cas12a) Nuclease and a Mammalian Cell-Free Extract. Cris J 1: 191–202. Available: http://www.liebertpub.com/doi/10.1089/crispr.2018.0006. Accessed 19 December 2018.

162. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2014) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 539–539. Available: http://www.ncbi.nlm.nih.gov/pubmed/21988835. Accessed 7 May 2019.

163. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, et al. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38: W695–W699. Available: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkq313. Accessed 7 May 2019.

164. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10: 845–858. Available: http://www.nature.com/articles/nprot.2015.053. Accessed 11 May 2019.

165. Crowgey EL, Kolb A, Wu CH (2015) Development of Bioinformatics Pipeline for Analyzing Clinical Pediatric NGS Data. AMIA Jt Summits Transl Sci Proc AMIA Summit Transl Sci 2015: 207–211. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4525226/pdf/2091834.pdf. Accessed 5 September 2017.

166. Baker SD, Zimmerman EI, Wang Y-D, Orwick S, Zatechka DS, et al. (2013) Emergence of polyclonal FLT3 tyrosine kinase domain mutations during sequential therapy with sorafenib and sunitinib in FLT3-ITD-positive acute myeloid leukemia. Clin cancer Res 19: 5758–5768. Available: http://www.ncbi.nlm.nih.gov/pubmed/23969938. Accessed 25 August 2017.

167. Zhang W, Konopleva M, Jacamo RO, Borthakur G, Chen W, et al. (2015) Acquired Point Mutations of TKD Are Responsible for Sorafenib Resistance in FLT3-ITD Mutant AML,. Blood 118. Available: http://www.bloodjournal.org/content/118/21/3505. Accessed 25 August 2017.

168. Fischer M, Schnetzke U, Spies-Weisshart B, Walther M, Fleischmann M, et al. (2017) Impact of FLT3-ITD diversity on response to induction chemotherapy in patients with acute myeloid leukemia. Haematologica 102: e129–e131. Available: http://www.ncbi.nlm.nih.gov/pubmed/28034991. Accessed 9 October 2017.

169. Schwartz GW, Manning B, Zhou Y, Velu P, Bigdeli A, et al. (2018) Classes of ITD predict outcomes in AML patients treated with FLT3 inhibitors. Available: http://clincancerres.aacrjournals.org.udel.idm.oclc.org/content/clincanres/early/2018/09/01/1078-0432.CCR-18-0655.full.pdf. Accessed 19 September 2018.

170. Warmuth M, Kim S, Gu X, Xia G, Adrián F (2007) Ba/F3 cells and their use in kinase drug discovery. Curr Opin Oncol 19: 55–60. Available: https://insights.ovid.com/pubmed?pmid=17133113. Accessed 26 September 2017.

171. Pradhan A, Lambert QT, Reuther GW, Goff SP (n.d.) Transformation of hematopoietic cells and activation of JAK2-V617F by IL-27R, a component of a heterodimeric type I cytokine receptor. Available: http://www.pnas.org/content/104/47/18502.full.pdf. Accessed 3 October 2017.

172. Tse K-F, Mukherjee G, Small D (2000) Constitutive activation of FLT3

stimulates multiple intracellular signal transducers and results in transformation. Leukemia 14: 1766–1776. Available: https://www.nature.com/leu/journal/v14/n10/pdf/2401905a.pdf. Accessed 21 September 2017.

173. Nguyen B, Williams AB, Young DJ, Ma H, Li L, et al. (2017) FLT3 activating mutations display differential sensitivity to multiple tyrosine kinase inhibitors. Oncotarget 8: 10931–10944. Available: www.impactjournals.com/oncotarget. Accessed 28 August 2017.

174. Walters DK, Stoffregen EP, Heinrich MC, Deininger MW, Druker BJ (n.d.) RNAi-induced down-regulation of FLT3 expression in AML cell lines increases sensitivity to MLN518. Available: http://www.bloodjournal.org/content/bloodjournal/105/7/2952.full.pdf. Accessed 5 October 2017.

175. Hou P, Wu C, Wang Y, Qi R, Bhavanasi D, et al. (2017) A Genome-Wide CRISPR Screen Identifies Genes Critical for Resistance to FLT3 Inhibitor AC220. 77: 4402–4413. Available: http://cancerres.aacrjournals.org.udel.idm.oclc.org/content/canres/77/16/4402.full.pdf. Accessed 25 August 2017.

176. Reiter K, Polzer H, Krupka C, Maiser A, Vick B, et al. (2018) Tyrosine kinase inhibition increases the cell surface localization of FLT3-ITD and enhances FLT3-directed immunotherapy of acute myeloid leukemia. Leukemia 32: 313–322. Available: http://dx.doi.org/10.1038/leu.2017.257.

177. Zimmerman EI, Turner DC, Buaboonnam J, Hu S, Orwick S, et al. (n.d.) Crenolanib is active against models of drug-resistant FLT3-ITD2positive acute myeloid leukemia. Available: http://www.bloodjournal.org/content/bloodjournal/122/22/3607.full.pdf. Accessed 24 April 2018.

178. Galanis A, Ma H, Rajkhowa T, Ramachandran A, Small D, et al. (2014) Crenolanib is a potent inhibitor of flt3 with activity against resistance-Conferring point mutants. Blood 123: 94–100. Available: www.bloodjournal.org. Accessed 31 July 2018.

179. Moloney JN, Stanicka J, Cotter TG (2017) Subcellular localization of the FLT3-ITD oncogene plays a significant role in the production of NOX-and p22 phox-derived reactive oxygen species in acute myeloid leukemia. Leuk Res 52: 34–42. Available: http://dx.doi.org/10.1016/j.leukres.2016.11.006. Accessed 9 January 2019.

180. Schmidt-Arras D, Böhmer S-A, Koch S, Müller JP, Blei L, et al. (2009) Anchoring of FLT3 in the endoplasmic reticulum alters signaling quality. Blood 113: 3568–3576. Available: http://www.ncbi.nlm.nih.gov/pubmed/19204327. Accessed 30 October 2018.

181. Koch S, Jacobi A, Ryser M, Ehninger G, Thiede C (2008) Abnormal Localization and Accumulation of FLT3-ITD, a Mutant Receptor Tyrosine Kinase Involved in Leukemogenesis. Cells Tissues Organs 188: 225–235. Available: www.karger.comwww.karger.com/cto. Accessed 13 December 2018.

182. Meshinchi S, Stirewalt DL, Alonzo TA, Boggon TJ, Gerbing RB, et al. (n.d.) Structural and numerical variation of FLT3/ITD in pediatric AML. Available: http://www.bloodjournal.org/content/bloodjournal/111/10/4930.full.pdf. Accessed 29 August 2017.

183. Tarlock K, Hansen ME, Hylkema T, Ries R, Farrar JE, et al. (2015) Discovery and Functional Validation of Novel Pediatric Specific FLT3 Activating Mutations in Acute Myeloid Leukemia: Results from the COG/NCI Target Initiative. Blood 126. Available: http://www.bloodjournal.org/content/126/23/87?sso-checked=true. Accessed 29 August 2017.

184. Xuan L, Wang Y, Huang F, Jiang E, Deng L, et al. (2018) Effect of sorafenib on the outcomes of patients with FLT3-ITD acute myeloid leukemia undergoing allogeneic hematopoietic stem cell transplantation. Cancer 124. doi:10.1002/cncr.31295.

185. Liu T, Ivaturi V, Sabato P, Gobburu JVS, Greer JM, et al. (2018) Sorafenib Dose Recommendation in Acute Myeloid Leukemia Based on Exposure-FLT3 Relationship. Clin Transl Sci 11: 435–443. Available: http://doi.wiley.com/10.1111/cts.12555. Accessed 27 February 2019.

186. Press Release. U.S. FDA Grants Orphan-Drug Designation to Astellas for Development of FLT3 Inhibitor Gilteritinib in Acute Myeloid Leukemia (n.d.). Available: http://www.prnewswire.com/news-releases/us-fda-grants-orphan-drug-designation-to-astellas-for-development-of-flt3-inhibitor-gilteritinib-in-acute-myeloid-leukemia-300491268.html. Accessed 10 October 2017.

187. Hassanein M, Almahayni MH, Ahmed SO, Gaballa S, Fakih R El (2016) FLT3 Inhibitors for Treating Acute Myeloid Leukemia. Clin Lymphoma, Myeloma Leuk 16: 543–549. Available: http://ac.els-cdn.com.udel.idm.oclc.org/S2152265016301070/1-s2.0-S2152265016301070-

main.pdf?_tid=7e71a606-8c25-11e7-9225-00000aab0f27&acdnat=1503948012_898f854da9035a13252a8924b2ab544d. Accessed 28 August 2017.

188. Sansbury B, Wagner A, Kmiec EB (2017) CRISPR-directed gene editing in vitro: A novel system for biomedical engineering and site-directed mutagenesis.

189. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, et al. (2016) CRISPR/Cas9 β-globin gene targeting in human haematopoietic stem cells. Nature 539: 384–389. Available: http://www.ncbi.nlm.nih.gov/pubmed/27820943. Accessed 8 September 2017.

190. Hotchkiss RD (1971) Toward A General Theory of Genetic Recombination in Dna. Adv Genet 16: 325–348. doi:10.1016/S0065-2660(08)60362-4.

191. Hotchkiss RD (1974) Models of genetic recombination. Annu Rev Microbiol 28: 445–468. doi:10.1146/annurev.mi.28.100174.002305.

192. Miller RC (1975) Replication and molecular recombination of T-phage. Annu Rev Microbiol 29: 355–376. doi:10.1146/annurev.mi.29.100175.002035.

193. Radding CM (1973) Molecular mechanisms in genetic recombination. Annu Rev Genet 7: 87–111. Available: http://www.annualreviews.org/doi/abs/10.1146/annurev.ge.07.120173.000511. Accessed 14 December 2015.

194. Brachman EE, Kmiec EB (2004) DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. J Cell Sci 117: 3867–3874. Available: http://jcs.biologists.org/content/117/17/3867.short. Accessed 24 July 2014.

195. Cho SW, Lee J, Carroll D, Kim J-S, Lee J (2013) Heritable gene knockout in Caenorhabditis elegans by direct injection of Cas9-sgRNA ribonucleoproteins. Genetics 195: 1177–1180. Available: http://www.ncbi.nlm.nih.gov/pubmed/23979576. Accessed 29 August 2016.

196. Gagnon JA, Valen E, Thyme SB, Huang P, Ahkmetova L, et al. (2014) Efficient Mutagenesis by Cas9 Protein-Mediated Oligonucleotide Insertion and Large-Scale Assessment of Single-Guide RNAs. PLoS One 9: e98186. Available: http://dx.plos.org/10.1371/journal.pone.0098186. Accessed 29 August 2016.

197. Sung YH, Kim JM, Kim H-T, Lee J, Jeon J, et al. (2014) Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. Genome Res

24: 125–131. Available: http://genome.cshlp.org/cgi/doi/10.1101/gr.163394.113. Accessed 29 August 2016.

198. Kim S, Kim D, Cho SW, Kim J-SJS, Kim J-SJS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24: 1012–1019. Available: http://genome.cshlp.org/content/24/6/1012.short. Accessed 9 July 2014.

199. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, et al. (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol 208: 44–53. Available: http://www.ncbi.nlm.nih.gov/pubmed/26003884%5Cnhttp://ac.els-cdn.com/S016816561500200X/1-s2.0-S016816561500200X-main.pdf?_tid=0c528f42-bae7-11e5-b035-00000aacb35e&acdnat=1452793901_25a6583e84069a4dd170b63c5a6406e8.

200. Cho SW, Kim S, Kim Y, Kweon J, Kim HS, et al. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24: 132–141. Available: http://www.ncbi.nlm.nih.gov/pubmed/24253446. Accessed 29 August 2016.

201. Liang Q, Huashan L, Yunhan J, Chunsheng D (2015) The molecular mechanism of CRISPR/Cas9 system and its application in gene therapy of human diseases. Yi chuan = Hered 37: 974–982. Available: http://www.ncbi.nlm.nih.gov/pubmed/26496749. Accessed 8 September 2017.

202. Gundry MC, Brunetti L, Lin A, Mayle AE, Kitano A, et al. (2016) Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. Cell Rep 17: 1453–1461. Available: http://www.ncbi.nlm.nih.gov/pubmed/27783956. Accessed 8 September 2017.

203. Weinstock DM, Brunet E, Jasin M (2008) Induction of Chromosomal Translocations in Mouse and Human Cells Using Site-Specific Endonucleases. JNCI Monogr 2008: 20–24. Available: http://www.ncbi.nlm.nih.gov/pubmed/18647997. Accessed 13 April 2019.

204. Brunet E, Simsek D, Tomishima M, DeKelver R, Choi VM, et al. (2009) Chromosomal translocations induced at specified loci in human stem cells. Proc Natl Acad Sci 106: 10620–10625. Available: http://www.ncbi.nlm.nih.gov/pubmed/19549848. Accessed 13 April 2019.

205. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, et al. (2013) Cancer translocations in human cells induced by zinc finger and TALE

nucleases. Genome Res 23: 1182–1193. Available: http://www.ncbi.nlm.nih.gov/pubmed/23568838. Accessed 23 February 2015.

206. Iliakis G, Wang H, Perrault AR, Boecker W, Rosidi B, et al. (2004) Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenet Genome Res 104: 14–20. Available: http://www.ncbi.nlm.nih.gov/pubmed/15162010. Accessed 13 April 2019.

207. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, et al. (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 31: 827–832. Available: http://dx.doi.org/10.1038/nbt.2647. Accessed 21 January 2014.

Appendix A

A STANDARD METHODOLOGY TO EXAMINE ON-SITE MUTAGENICITY AS A FUNCTION OF POINT MUTATION REPAIR CATALYZED BY CRISPR/CAS9 AND SSODN IN HUMAN CELLS

If you only have one appendix, change the styles of the previous two paragraphs to *Appendix - one* and *APPENDIX TITLE - one*, respectively.

DEVELOPED PROTOCOL

The video component of this protocol can be found at https://www.jove.com/video/56195/

Cell Line and Culture Conditions

1. Make 500 mL of medium for the culture of HCT 116 cells: McCoy's 5A modified medium supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 1% penicillin (this is complete medium).

NOTE: Grow HCT 116-19 cells in a T-75 or T-175 flask prior to plating. When 90% confluent, each T-75 flask will yield 8.4 x 10^6 cells, roughly five 10-cm plates, and each T-175 will yield 18.4 x 10^6 cells, roughly fifteen 10-cm plates.

Harvesting Cells from the Flask

1. Aspirate away the medium, wash with Dulbecco's Phosphate-Buffered Saline without calcium and magnesium (PBS) (10 mL for T-75 or 25 mL for T-175), and aspirate.
2. Add trypsin dropwise to the flask using a 2-mL pipette (2 mL for T-175 or 1 mL for T-75). Place the flask in an incubator at 37 °C and 5% CO_2 for 5 min to allow the cells to detach.
3. Tap the flask to make sure that all cells are dislodged and then quench with complete medium by dispersing it over the entire surface of the flask (8 mL for T-175 or 4 mL T-75).
4. Pipette up and down multiple times to break up cell clumps and transfer the cells to a 15-mL conical tube.
5. Before spinning the cells down, take 10 µL from the 15-mL conical and combine it with 10 µL of trypan blue to count the cells. Pellet the cells by spinning for 5 min at 125 x g and 16 °C.

Counting the Cells

1. Transfer 10 µL of the cells mixed with trypan blue to the hemocytometer. Count the 4 grids around the outside (each grid contains 16 squares).
    1. Take the average cell count from each set of sixteen corner squares.
    2. Multiply by 10,000 (10^4).
    3. Multiply by the total volume of medium used to harvest the cells to correct for the dilution from the trypan blue addition.

NOTE: The equation format to calculate the volume to resuspend the cells follows:

Plating the Cells

1. For each 10-cm plate of cells to be synchronized, add 5 mL of complete medium and 6 µM of aphidicholin (12 µL of a 2.5 mM stock in 200-proof ethanol).
2. Transfer 100 µL of re-suspended cell pellet, 2.5 x 10^6 cells, to each 10-cm plate and swirl gently to mix.
3. Incubate the plates at 37 °C and 5% CO_2 for 16-24 h to synchronize cells at the G1/S border.

Releasing the Cells from Aphidicolin Synchronization

1. 4 h prior to targeting, aspirate the medium, wash with PBS, aspirate the PBS, and add 5 mL of complete medium
2. Place it back in the incubator at 37 °C and 5% CO_2 for 4 h

RNA Complexing

1. Enter the mutant eGFP gene sequence into the Zhang Lab's online generator[207] (http:// crispr.mit.edu/) and choose the CRISPR guide sequences that bind with close proximity to the target site. Obtain the CRISPR guide sequences from a commercial source.
2. Store the CRISPR RNA (crRNA), trans-activating crRNA (tracrRNA), and Cas9 protein at 20 °C and use according to manufacturer suggestions.
    1. Mix the RNA in equimolar concentrations to 45 µM. Add 6.75 µL of a 200 µM stock of crRNA and 6.75 µL of a 200 µM stock of tracrRNA to a 1.5-mL centrifuge tube. Add 16.50 µL of TE Buffer to make a final volume of 30 µL.

3. Heat at 95 °C for 5 min in a heat block or PCR machine.
Caution: Hot!

4. Allow to cool to room temperature. NOTE: If using a PCR machine, set the cooling to 0.2 °C/s.

5. Perform the following steps for each sample.
    1. Dilute 2.22 µL of crRNA:tracrRNA complex in 2.78 µL of TE Buffer (10 mM Tris, pH 8.0 and 0.1 mM EDTA; pH 8.0) to a final volume of 5 µL.
    2. Dilute 1.67 µL of Cas9 Protein from a 60 µM stock in 3.33 µL of low-serum medium to final volume of 5 µL.

6. Mix 5 µL of Cas9 protein with 5 µL of complexed RNA

Harvesting the Cells for Targeting

1. Aspirate the medium, wash with 5 mL of PBS, aspirate the PBS, and add 1 mL of pre-warmed trypsin to each 10-cm plate. Put the plates in the incubator at 37 °C and 5% CO_2 for 5 min.

2. Tap on the 10-cm plate to make sure all cells are dislodged and then quench with 4 mL of complete medium by dispersing it over the entire surface of the plate.

3. Pipette up and down multiple times to break up cell clumps and transfer the cells to a 15-mL conical tube.

4. Before spinning the cells down, take 10 µL from the 15-mL conical tube and combine it with 10 µL of trypan blue to count the cells. Pellet the cells by spinning for 5 min at 125 x g and room temperature.

5. Aspirate the medium and wash with 5 mL of PBS. Pellet the cells by spinning at 125 x g for 5 min at room temperature.

Counting the Cells

1. Transfer 10 µL of the cells mixed with trypan blue to the hemocytometer. Count the 4 grids around the outside (each grid contains 16 squares).
    1. Take the average cell count from each set of sixteen corner squares.
    2. Multiply by 10,000 (10^4).
    3. Multiply by the total volume of medium used to harvest the cells to correct for the dilution from the trypan blue addition.

NOTE: Following is the equation format to calculate the volume to resuspend the cells. Re-suspend the required number of cells in serum-free McCoy's 5A modified medium.

Targeting Samples

1. Transfer 100 µL of cell suspension (5×10^5 cells) from step 8.1 to each electroporation 4-mm gap cuvette. Add 10 µL of RNP complex from step 6.7 to 100 µL of cells at a 5×10^5 cell density. Add ODN (2 µM) to each sample. NOTE: For a positive control, add 1 µL of eGFP at 1 µg/µL expressing plasmid
    1. Take the rack to an electroporation machine, lightly flick each sample, and place them in the chamber. Electroporate at 250 V, LV; 2 pulses, 1 s; 13 ms; unipolar pulse.
    2. Transfer the rack back to the hood. Transfer each sample to a well containing 2 mL of complete medium in a 6-well plate. Incubate at 37 °C and 5% CO_2 for 72 h before checking for correction levels.

Analysis of Gene Edited Cells and Transfection Efficiency

1. Aspirate the medium and wash the cells with 2 mL of PBS. Aspirate the PBS and add 500 µL of pre-warmed trypsin to each well of the 6-well plate. Put the plates in the incubator at 37 °C and 5% CO_2 for 5 min.
2. Tap the plate to make sure all cells are dislodged and then quench with 1 mL of complete medium by dispersing it over the entire surface of the well.
3. Pass the cells into a 1.5-mL centrifuge tube and pellet at 5,000 x g for 5 min at room temperature.
4. Aspirate the medium. Re-suspend the cell pellet in 500 µL of FACS buffer (0.5% BSA, 2 mM EDTA, and 2 µg/mL propidium iodide in PBS).
5. Measure the cell fluorescence (eGFP+) by flow cytometry.
    1. Calculate the correction efficiency as the percentage of the total live eGFP-positive cells over the total number of live cells in each sample, as described in Rivera-Torres et al[45].

DNA Sequence Analysis

1. Electroporate the synchronized and released HCT 116-19 cells at a concentration of 5 x 10^5 cells/100 µL, with RNP complex at 100 pmols and 72NT ODN at 2.0 µM.
2. Transfer the cells to 6-well plates and allow them to recover for 72 h.
3. Sort the cells individually into 96-well plates using a FACS sorter with a 488-nm (100 mw) laser for eGFP+/-, as described in Rivera-Torres et al[45]. NOTE: Not all wells will successfully grow.

4. Expand the cells over 6 weeks and harvest as described above.
5. From the wells that have growth, isolate cellular gDNA using a commercially available DNA isolation kit and amplify the region surrounding the target base via PCR (718 bp; forward primer 5'-ATGGTGAGCAAGGGCGAGGA-3' and reverse primer 5'-ACTTGTACAGCTCGTCCATGC-3').
6. Perform DNA sequencing analysis on the samples.

Appendix B

List of Manuscripts

1. Bialk P, **Rivera-Torres N**, Strouse B, Kmiec EB (2015) Regulation of Gene Editing Activity Directed by Single-Stranded Oligonucleotides and CRISPR/Cas9 Systems. PLoS One 10: e0129308.

2. Bialk P, Sansbury B, **Rivera-Torres N**, Bloh K, Man D, et al. (2016) Analyses of point mutation repair and allelic heterogeneity generated by CRISPR/Cas9 and single-stranded DNA oligonucleotides. Sci Rep 6: 32681.

3. **Rivera-Torres N**, Banas K, Bialk P, Bloh KM, Kmiec EB (2017) Insertional Mutagenesis by CRISPR/Cas9 Ribonucleoprotein Gene Editing in Cells Targeted for Point Mutation Repair Directed by Short Single-Stranded DNA Oligonucleotides. PLoS One 12: e0169350.

4. **Rivera-Torres N**, Kmiec EB (2017) A Standard Methodology to Examine On-site Mutagenicity As a Function of Point Mutation Repair Catalyzed by CRISPR/Cas9 and SsODN in Human Cells. J Vis Exp: e56195–e56195.

5. Modarai SR, Man D, Bialk P, **Rivera-Torres N**, Bloh K, et al. (2017) Efficient delivery and nuclear uptake is not sufficient to enable detectable gene editing in CD34+ cells directed by a ribonucleoprotein complex.

6. **Rivera-Torres N**, Kmiec EB (2019) Modeling Pediatric AML FLT3 Mutations with CRISPR-Cas12a Mediated Gene Editing (***in preparation***)