Site-Specific Genome Engineering in Mouse Primary Fibroblasts

APPROVED BY SUPERVISORY COMMITTEE

Matthew Porteus, M.D., Ph.D.

_____

Lily Huang, Ph.D.

_____

James Amatruda, M.D., Ph.D.

_____

Diego Castrillon, M.D., Ph.D.

_____

DEDICATION

First, this work is dedicated to my parents, my sister and my husband. I have had the most supportive family anyone could hope for. Because of them, it never occurred to me to aim for anything less than the highest goals I could set for myself. Secondly, I dedicate this work to the influential mentors that have guided me in each step of my career. Education of future generations is the most important aspect of science. Without teachers, the discoveries we make are meaningless. I am fortunate to have encountered these people, who have been truly the best mentors anyone could ask for. Because of the time they invested in my learning, I am infinitely better equipped to pursue a career as a physician scientist. In turn, education of the next generation will always be a priority for me.

SITE-SPECIFIC GENOME ENGINEERING IN MOUSE PRIMARY FIBROBLASTS

by

JENNY BARKER

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center

Dallas, TX

June, 2012

SITE-SPECIFIC GENOME ENGINEERING IN MOUSE PRIMARY FIBROBLASTS

Jenny Barker

The University of Texas Southwestern Medical Center at Dallas, 2012

Matthew H. Porteus, M.D., Ph.D.

Site-specific genome engineering is a powerful tool for medical therapeutics and basic scientific research. As the name implies, site-specific genome engineering describes the ability to add or subtract nucleic acid information in a precise, controlled manner within the genome. This technology has developed as a result of two key discoveries. The first is that a cell's double strand break machinery can be highjacked to affect a desired change in the genome. Double stranded breaks are typically repaired by a pathway called non-homologous end joing (NHEJ) which can allow for disruption of an endogenous locus, or the pathway of homologous recombination (HR) which can allow for insertion of sequences if a homologous donor is supplied. The second major discovery is that chimeric enzymes can be engineered to create site-specific double stranded breaks, and that these enzymes can dramatically stimulate the frequency of gene targeting at a given locus. Recently, there has been an outpouring of studies performing site-specific genome engineering in human cell lines and primary cells, including embryonic stem cells, induced pluripotent stem

cells and CD34+ hematopoietic stem cells. However, very little has been accomplished in terms of animal modeling of these principles. The work described in this thesis seeks to address some of these issues in a reporter mouse model that we have developed to study genome engineering. With this model we have demonstrated gene correction of an endogenous locus and also site-specific gene addition through a novel strategy which does not require disruption of the gene product at the site of the insertion. We have accomplished gene correction in several cell types including embryonic and adult fibroblasts, astrocytes and embryonic stem cells. For gene addition, we have demonstrated site-specific insertion of multiple transgenes including human growth hormone and human platelet derived growth factor-B, as well as a surface selectable marker (the truncated nerve growth factor receptor) and drug selectable marker. The tantamount goal for *ex vivo* genome engineering is to be able to modify a patient's cells and then re-transplant them into the patient for a therapeutic outcome. In our mouse model, we have described a strategy where primary fibroblasts can be modified and then transplanted back into a recipient mouse. We are currently investigating the ability of these fibroblasts to serve as vehicles for local protein delivery to augment biological processes, such as tissue repair and wound healing. This work contributes to the body of knowledge that will enable translation of genome engineering from a scientific endeavor into a clinical reality.

TABLE OF CONTENTS

.

PRIOR PUBLICATIONS

Connelly, J.P., Barker, J.C., Pruett-Miller, S., and Porteus, M.H. Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. Mol Ther *18*, 1103-1110.

Barker, J.C., Voit, R.A., Porteus, M.H. "Nuclease mediated targeted genome modification in mammalian cells" in *Site-Directed Insertion of Transgenes (Topics in Current Genetics)*. Springer. July 2012.

Barker, J.C., Huang, L.J., Porteus M.H. "Zinc finger nuclease and TAL effector nuclease mediated safe harbor gene addition without safe harbor gene disruption" (Manuscript in Preparation)

LIST OF FIGURES

LIST OF DEFINITIONS

AAV- Adeno-associated virus

ADA- Adenosine Deaminase

CCR5- C-C Chemokine Receptor type 5

CGD- Chronic Granulomatous Disease

CMV- Cytomegalovirus

DSB- Double Strand Break

DNA- Deoxyribonucleic Acid

Epo- Erythropoietin

ES- Embryonic Stem Cell

FokI- Restriction Enzyme from Flavobacterium okeanokoites

GFP- Green Fluorescent Protein

HR- Homologous Recombination

IL2R-γ– Interleukin-2 receptor gamma

I-SceI- Homing Endonuclease from *Saccharomyces cervisiae*

iPS- Induced Pluripotent Stem Cell

MSC- Mesenchymal Stem Cell

NHEJ- Non-homologous end joining

PDGF-B- Platelet Derived Growth Factor B

PolyA- Polyadenylation

PPP1R12C- Protein phosphatase 1 regulatory subunit 12C

ROSA26- Reverse-Orientation Splice Acceptor

SCID- Severe Combined Immunodeficiency

TAL- Transcription activator like

TALEN- TAL effector nuclease

tdTomato- tandem dimer tomato, fluorescent protein

Ubc- Ubiquitin C

Ubx- Ultrabithorax

ZFN- Zinc Finger Nuclease

Chapter I- INTRODUCTION TO SITE SPECIFIC GENOME ENGINEERING, A REVIEW OF THE LITERATURE

Portions previously published and co-written with Richard Voit

Current Treatment Options for Monogenic Diseases

The fundamental basis of medicine is to provide therapy for human diseases and a major aspect of medical therapy is to provide medicines derived from small molecules to do this. These small molecule based therapies have had a dramatic impact on the human condition and provide curative therapy for many diseases. Nonetheless, for the preponderance of diseases, small molecule based therapy provides symptomatic rather than curative relief. In particular, for diseases caused by genetic mutations, small molecule based therapy provides treatment for the consequences of the underlying genetic defect rather than directly addressing the root cause. In the last several decades, however, new classes of therapeutics have either been developed or are being developed that attempt to treat genetic diseases at their root cause, including the use of protein replacement therapy, bone marrow transplantation, and gene therapy.

Enzyme replacement therapy for the lysosomal storage diseases and severe combined immunodeficiency caused by mutations in the adenosine deaminase (ADA) gene, as well as clotting factor replacement for hemophilia are examples of protein therapeutics in which protein is administered directly to replace the missing protein. Protein replacement therapy does provide lifer-altering treatment for patients for which it is available, but does not cure the disease as the underlying genetic defect remains. Further, protein replacement is limited in both cost and effectiveness. The high cost of protein therapy (hundreds

of thousands of dollars per year per patient for some diseases) means this type of therapy is not available to patients from resource poor parts of the world. Finally, despite attempts at prolonging the half-life of therapeutic proteins, protein replacement therapy requires life-long, repeated infusions at daily, weekly or monthly intervals.

The administration of recombinant proteins is not limited to the treatment of genetic diseases, and can be used to enhance biologic functions as well. One example includes the use of recombinant PDGF-BB (Becaplermin gel or Regranex) to augment the wound healing process in diabetic ulcers (Smiell et al., 1999). However, peptide-based therapy is limited in efficacy within the destabilized chronic wound environment, and other, cell-based strategies may be preferable. Cell-based delivery platforms may have the advantage of providing protein delivery in a stable, constituitive or regulated fashion, and within biologic context.

The only broadly used cell-based therapeutic is allogeneic bone marrow transplantation. Bone marrow transplantation is a highly effective therapy for patients with genetic diseases of the blood such as sickle cell disease, thalassemia and SCID, where the hematopoietic system containing the disease-causing mutation is replaced with the hematopoietic system free of the genetic defect from an immunologically matched donor. The limitations, however, of bone marrow transplantation are several-fold. First, like protein therapeutics, bone marrow transplantation is a highly technical and expensive procedure and so is not available in many parts of the world. Second, many patients do not have a

suitable immunologically-matched donor. Third, there are long-term side-effects from bone marrow transplantation including graft versus host disease and side-effects of high doses of myeloablative conditioning therapy that can result in conditions worse than the original disease.

An alternative to protein replacement therapy or allogeneic stem cell therapy is to use genome engineering to modify the patient's own cells and use those gene modified cells as treatment for the disease ("gene therapy"). This can be accomplished by disruption of a gene product, correction of a gene where modified cells replace diseased cells, or by addition of a gene where modified cells functionally participate in the correction of the diseased state or simply act as vehicles for protein delivery. The modification of the patient's cells can be done *in vivo* in which the modification takes place without the cells leaving body or *ex vivo* in which the cellular modification takes place outside the body and then the cells are transplanted back into the patient. Unlike small molecule therapy or protein replacement therapy for genetic diseases, this strategy is directed at the root cause of the disease and if successful, cure could be achieved with a single intervention. Unlike allogeneic bone marrow transplantation, this approach uses the patient's own cells and so the problems of finding matched donors and graft versus host disease are solved.

There have been several gene therapy clinical trials for genetic diseases of the blood including SCID-X1, ADA-SCID, Wiskot-Aldrich syndrome, chronic granulomatous disease, and thalassemia (Abuljadayel et al., 2006; Aiuti et al., 2009a; Bank et al., 2005; Boztug et al., 2010; Cavazzana-Calvo et al., 2000;

Cavazzana-Calvo et al., 2005; Gaspar et al., 2006; Kang et al., 2010; Malech et al., 1997). In each of these trials, the patient's bone marrow cells have been removed, transduced with a retroviral or lentiviral vector that carries a wild-type copy of the mutated gene and then transplanted back into the patient. Tens of patients of worldwide have benefited from this therapy. A fundamental feature of this viral based strategy is that the transgene integrates in an uncontrolled fashion. An unfortunate consequence of this uncontrolled integration is that the integration can activate a latent oncogene and this has occurred in a number of patients, causing frank leukemia, myelodysplasia, or a clonal proliferation (Hacein-Bey-Abina et al., 2008; Hacein-Bey-Abina et al., 2003b; Ott et al., 2006). Overall, the frequency of uncontrolled integration seems to be greater than 10% although there are vector and disease specific differences and this relatively high frequency may make this approach unsustainable for the treatment of genetic diseases.

Targeted genome engineering, a technology that enables site-specific, controlled insertion of a transgene, allows for the benefits of transgene technology while avoiding the complications of random integration. Targeted genome engineering strategies in human cells have been developed to stimulate sequence specific DNA modifications to achieve one of three therapeutic endpoints. First, the targeted disruption of an endogenous locus can provide a rapid, efficient means for researchers to achieve permanent gene knockouts in human cells, a powerful research tool that has heretofore been restricted to model organisms. In addition, the ability to disrupt a gene at its endogenous

locus provides a new therapeutic approach in which targeted gene mutation can be used to treat a disease. The second use of targeted genome engineering is the direct correction of disease-causing mutations at endogenous loci. Third, targeted genome modification can be used to target transgene addition to precise locations in the genome, including "safe harbors." Such targeted transgene addition has important ramifications both therapeutically, to eliminate insertional oncogenesis, and as a scientific tool. With such potentially precise methods of genomic manipulation available, researchers now have the ability to begin to address the root cause of many of the more than 10,000 monogenic diseases.

Mammalian Double Strand Break Repair

Most natural double strand breaks (DSBs) are the result of ongoing environmental stresses such as ionizing radiation and metabolically derived reactive oxygen species (**Figure 1**). Unrepaired DSBs lead to cell death and multiple, redundant cellular DNA repair pathways have evolved to repair these important genetic lesions. The two most important DSB repair mechanisms are non-homolgous end joining (NHEJ) and homologous recombination (HR).

NHEJ is performed by re-ligation of DSB ends in either a non-mutagenic or mutagenic fashion. This pathway is initiated by the end-binding of Ku70 and Ku80 and the subsequent recruitment of DNA Protein Kinase catalytic subunit (DNA-PKcs). Once the ends are in juxtaposition, Artemis proceeds with nucleolytic end processing. DNA Ligase IV, XRCC4 and XLF complete ligation of the ends (Wyman and Kanaar, 2006) (**Figure 1**, **Left Panel**).

A second pathway by which DSB repair can occur is through homologous recombination. As the name implies, HR requires the presence of a homologous template, which is typically the sister chromatid. As such, HR is thought to occur mostly in the S/G2 phase of the cell cycle. Indeed, inducing an S/G2 cell cycle arrest in certain cell types allows for increases in the frequency of an HR event (Connelly et al., 2010). The HR pathway is initiated by the MRN complex (Mre11/Rad50/Nbs1) binding the DSB ends and generating 3' overhangs by 5' end resection. BRCA1 binds ssDNA and facilitates Rad51 assembly who mediates strand invasion of the homologous template. The repair process then proceeds with either DNA synthesis, Holliday junction formation and resolution or through the synthesis dependent strand annealing (SDSA) pathway (Szostak et al., 1983; Wyman and Kanaar, 2006) (**Figure 1, Right Panel**). In the case where the sequence of the homologous template differs, an altered sequence can be incorporated into the original locus. A few examples of this include loss of heterozygosity (LOH) where one wild-type allele is converted to a mutant allele or an intentional sequence altering event through a genome engineering strategy.

Gene Targeting by Homologous Recombination

As suggested, the HR pathway can be manipulated to precisely modify a locus. This is accomplished by designing an exogenous plasmid with regions of homology surrounding an intentional sequence alteration. When the HR pathway utilizes the exogenous DNA as the homologous template, instead of the sister chromatid, conversion of the target locus can occur. This strategy was first demonstrated in S. cerevisiae by gene targeting of a Leu2+ sequence into the

Leu2 gene in a Leu2- strain of yeast (Hinnen et al., 1978). Later, several groups demonstrated that similar targeting events could also occur in mammalian cells, including integration at the endogenous human beta-globin locus, albeit with an extremely low frequency (Smith and Berg, 1984; Smithies et al., 1985; Smithies et al., 1984). Gene targeting is currently best known for its use in the generation of genetically altered mice through embryonic stem cell knock-in technology (Thompson et al., 1989). Precise genomic modification certainly held promise for use in the treatment of genetic diseases, but the low frequency of occurrence for targeting events limited it's practicality.

Because DSBs are the initiating event for HR, it is logical that site-specific generation of a DSB could stimulate HR at a particular locus. Indeed, in 1994 Rouet *et al* demonstrated that integration of an 18 base pair recognition site for the homing endonuclease I-Sce into a non-functional neomycin resistance gene could stimulate HR and gene targeting at the locus, restoring gene expression (Rouet et al., 1994a). This work was a critical stepping stone for the translation of gene targeting into gene therapy. However, because this strategy required integration of the I-Sce site into the target locus, it was also not practical for gene therapy.

Type IIS restriction enzymes such as FokI contains two separate domains, the C-terminal of which is a non-specific nuclease domain. The N-terminus is responsible for DNA binding and specificity (Li et al., 1992; Wah et al., 1997). This observation led to the hypothesis that the nuclease domain could be isolated and fused to a new DNA binding domain to confer specificity at a new

site. Indeed, Kim and Chandrasegaran *et al* demonstrated feasibility of this concept by fusing the FokI nuclease domain to the DNA binding domain of Ultrabithorax (Ubx), a Drosophila transcription factor. This chimeric protein was able to bind and create a site-specific DSB at the Ubx locus *in vivo (*Kim and Chandrasegaran, 1994). Subsequent DNA binding domains (such as zinc-fingers or TAL effectors) have been fused with the modular FokI nuclease domain, and this has led to the development of chimeric proteins that can create site-specific DSBs at virtually any locus in the genome.

Methods for Site-Specific Genome Engineering

While AAV vectors can be used for targeted gene addition and gene correction (Khan et al., 2011; Khan et al.), the focus of this section will be on using site-specific nucleases to perform targeted gene mutation, gene correction and targeted gene addition. Central to targeted genome engineering is the ability to design highly specific nucleases. There are now three general classes of sequence specific nucleases: 1) homing endonucleases 2) zinc-finger nucleases (ZFNs) and 2) TAL effector nucleases (TALENs). As mentioned, ZFNs and TALENs share a common general architecture in which an engineered DNA binding domain is fused to the non-specific nuclease domain from the FokI restriction enzyme. The genome specificity of ZFNs and TALENs derives from both the sequence specificity of the DNA binding domain and the need for the FokI nuclease domain to dimerize (Bitinaite et al., 1998) in order to create a DSB at a specific target site. As previously discussed, the DSB is the initiating even in generating targeted genome modification.

Zinc-Finger Nucleases

Zinc fingers have a conserved ββα structure and the α helix fits into the major groove of DNA, interacting with three nucleotides. The first six amino acids of the α helix, along with the immediately preceding amino acid, give the zinc finger its nucleotide specificity. By stringing multiple zinc fingers in series, proteins can be designed with 9- or 12-base pair (or longer) recognition sequences. When these zinc finger motifs are fused at their C-termini to the FokI nuclease domain, the resulting chimeric protein is known as a zinc finger nuclease (ZFN) (Kim et al., 1996). Study of the activity of the wild type FokI endonuclease revealed that efficient cleavage is mediated by the dimerization of nuclease domains (Bitinaite et al., 1998). The dimerization requirement for efficient activity is maintained for chimeric nucleases including ZFNs. For ZFNs, dimerization is achieved by having two different ZFNs bind in an inverted orientation on opposite DNA strands thereby allowing the nuclease domain to dimerize and cut the DNA in the spacer region between the two ZFN binding sites (Bibikova et al., 2001). Each individual ZFN has a 9- or 12- base pair target binding site, a length not sufficient to confer genomic specificity, but a pair of ZFNs has an 18- or 24- base pair target site, a length which is sufficient to confer genomic specificity.

There are several approaches to the design of ZFNs. The simplest is to use modular-assembly in which individual fingers are strung together in an array of three or more. This approach is relatively straightforward and has been useful in several different studies (Beerli et al., 2000; Dreier et al., 2001; Dreier et al.,

2005). The major disadvantage is that it has a low success rate when target sites do not fit a certain type of sequence and even with potentially ideal target sites, the zinc finger nucleases made using this strategy have lower activity and higher off-target effects than proteins made using other strategies. A similar strategy is to use modular-assembly to string together two-finger modules into an array (Perez et al., 2008; Urnov et al., 2005). This strategy was first developed by Gendaq, then bought by Sangamo Biosciences, and now sold to the research community through Sigma-Aldrich Pharmaceuticals. Because this method has not been publicly described, the overall success rate has not been published but the strategy has had a high success rate for targeting multiple genes. This method is based on a proprietary archive of two-finger modules and thus is not freely available to researchers. A derivative of this approach is to assemble a ZFN using two-finger modules and then to subsequently refine the original protein for improved properties ("Sangamo approach"). It is through this method that the best published ZFNs from Sangamo Biosciences have been made but is also not an approach that is freely available to others. An alternative approach to assembling pre-determined modules, is to use selection strategies, either by phage display or using bacterial based methods (including the Oligomerized Pool Engineering or "OPEN" method), to select for zinc finger domains that bind the desired target sequence out of a randomized library of proteins. These selection approaches are difficult to use but do result in nucleases with properties of high activity and lower off-target effects. Interestingly, ZFNs made using the Sangamo approach and using the OPEN approach have been compared for their off-target

effects and both methods seem to generate relatively specific ZFNs but they achieve the specificity through different mechanisms. Finally, there are two derivates of the OPEN approach to making ZFNs. In context dependent assembly (CoDA), modular assembly is used to assemble ZFNs using zinc finger modules that have been previously identified using the OPEN system (Sander et al., 2011). In a hybrid approach individual modules can be combined with the OPEN system to generate proteins that have high activity and relatively low toxicity (Wilson and Porteus, unpublished data).

Homing Endonucleases and TAL Effector Nucleases (TALENs)

Zinc finger nucleases have had the widest application in targeted genome modifications, but other classes of nucleases are also being developed to stimulate targeted DNA double stand breaks. The first of these classes includes meganucleases, also called homing endonucleases. The major family of meganucleases, called LAGLIDADG endonucleases because of the presence of a conserved amino acid motif, are derived from the mitochodria and chloroplasts of eukaryotic unicellular organisms such as yeast. As their name suggests, meganucleases have very long DNA recognition sites (between 12 and 40 base pairs), allowing them to bind DNA with very high specificity. The first example of a meganuclease used in gene targeting is I-SceI. In 1994, Maria Jasin and colleagues stably integrated two I-SceI sites into mouse chromosomes and demonstrated cleavage at one or both of those sites by transient expression of I-SceI (Rouet et al., 1994b). To use meganucleases in targeted genome modification, however, the nuclease must be re-engineered to recognize a new

target site. Several different academic groups and companies (Cellectis and Precision Biosciences) have had some success with the re-engineering of meganucleases to recognize new target sites for the purposes of genome engineering using targeted genome modification (Arnould et al., 2006; Arnould et al., 2011; Grizot et al., 2010; Smith et al., 2006). However, the re-engineering of meganucleases to new target sites is a challenging endeavor and has not been widely adopted.

Recently, another DNA-binding domain has been identified as a potential motif for use in chimeric nucleases. Transcription activator-like (TAL) effectors are virulence factors in the phytopathogenic bacteria *Xanthomonas* and bind to DNA through a series of nearly identical repeats. Each repeat of 34 amino acids has two key residues that recognize and bind one nucleotide, and proteins with 11 to 17 repeats can be designed *de novo* to bind very specific DNA sequences (Boch et al., 2009; Moscou and Bogdanove, 2009). TAL effector repeats, like zinc finger proteins, can be fused to the FokI nuclease domain to generate TAL effector nucleases (TALENs) which dimerize and effectively cleave expected DNA sequences in yeast and plants (Cermak et al., 2011; Christian et al., 2010; Mahfouz et al., 2011). Notably, TALENs have also been developed which stimulate homologous recombination at an endogenous human locus at a rate of 16% (Miller et al., 2011). A significant advantage of TALENs compared to ZFNs is their seemingly total modularity; the same repeat recognizes and binds the same nucleotide regardless of the context of its neighboring repeats. This modularity eliminates most of the screening steps required in the synthesis of

other chimeric nucleases. In fact, the Voytas and Bogdanove groups have established a rapid method to design and synthesize novel TALENs and has made the necessary reagents publicly available (Cermak et al., 2011). Though most current literature contains ZFN data, in early comparisons, TALENs seems to have comparable activity, yet fewer off-target effects than ZFNs, and may soon be the more widely adopted method for targeted genome engineering strategies.

Targeted Genome Modification using non-Nuclease Approaches

In addition to nuclease-mediated genome editing strategies, adeno-associated virus (AAV) has been used for targeted genome modifications (Khan et al., 2011). Through mechanisms still not clear, at high multiplicities of infection, AAV can stimulate targeted transgene integration or gene correction at frequencies up to 1% under certain experimental conditions. While this frequency is a log lower than what can be achieved using nuclease mediated genome modification, it does have the advantage of not needing to engineer site-specific nucleases.

Papapetrou *et al* have described another site-specific viral targeting approach in induced pluripotent stem cells (iPS). This lentiviral-mediated approach depends upon retrospective analysis and selection of optimal integration events at a “safe harbor” locus. The authors define a safe harbor locus as a locus that meets the following citeria: a locus that is 50kb from the 5’ end of any gene, 300kb from any cancer-related gene, 300kb from any microRNA and one that is located outside transcription units or “ultraconserved”

regions. The iPS cells for this experiment were derived from patients with β -thalassemia and consequently did not express functional β -globin when differentiated into the erythroid lineage. The iPS cells were infected at a low multiplicity of infection with a β-globin expressing lentivirus. Single copy integrant clones were then selected and screened for the genomic location of integration. After screening, the authors found that they could isolate an iPS clone expressing β -globin efficiently from a locus that met the “safe harbor” criteria (Papapetrou et al., 2011). Although this retrospective strategy differs from many of the prospective targeting approaches, the end result is similar and perhaps provides a plausible alternative for targeted genome engineering.

Gene Disruption, Gene Correction and Gene Addition

The following describes the types of genome modifications that can be effected using site-specific nucleases. These can be classified into three major categories that include 1) endogenous gene disruption 2) endogenous gene correction 3) safe harbor gene addition. The type of modification selected depends upon the etiology of the disease state, i.e. a protein that contributes to a disease state can be removed by endogenous gene disruption, a missing protein can be replaced by safe harbor gene addition or the defect can be reversed at the endogenous site through gene correction.

Gene Disruption

The simplest method of targeted genome modification is gene disruption by mutagenic NHEJ after the induction of a DSB by the engineered nuclease. As described, in NHEJ the DNA ends created by a DSB are ligated back to one

another. Often, this process occurs without altering the DNA sequence near the break site (**Figure 1a**), but sometimes repair by NHEJ introduces insertions or deletions at the site of the break ("mutagenic NHEJ") (**Figure 1b**). These insertions or deletions, if they occur in the coding region of a gene, can result in frameshift mutations creating a non-functional gene product. If the DSB is created by an engineered nuclease and that DSB is repaired by mutagenic NHEJ, inactivating gene specific frameshift mutations can be generated. Thus, by combining the engineering of gene specific nuclease with the intrinsic mutagenic property of NHEJ, researchers can create targeted gene mutations. In this method of targeted genome modification, the site of the mutation is controlled by the specificity of the nuclease but the specific genomic modification at that site is uncontrolled. Depending on the activity and expression level of the nuclease, targeted gene mutation frequencies of greater than 50% can be generated (Perez et al., 2008).

Gene disruption has been used for a variety of experimental purposes including the creation of knockout mammalian cell lines as well as primary cells including stem cells (Alwin et al., 2005; DeKelver et al., 2010; Hockemeyer et al., 2009; Kim et al., 2009; Perez et al., 2008; Porteus and Baltimore, 2003; Urnov et al., 2005; Zou et al., 2009). Further, whole animal knockout has been achieved by this method in model organisms that include flies (Beumer et al., 2006; Beumer et al., 2008; Bibikova et al., 2002; Bozas et al., 2009), zebrafish (Doyon et al., 2008; Foley et al., 2009; Meng et al., 2008), mice (Carbery et al., 2010; Cui et al., 2011; Meyer et al., 2010), rats (Geurts et al., 2009; Mashimo et al., 2010),

rabbits (Flisikowska et al., 2011), frogs (Young et al., 2011), sea urchins (Ochiai et al., 2010), plants (Lloyd et al., 2005; Wright et al., 2005) and nematodes (Morton et al., 2006) and worms (Takasu et al., 2010). This is particularly useful in species where classic transgenesis is more difficult, such as in zebrafish and rats. To create gene specific knockouts, site-specific nucleases have been introduced as either DNA expression plasmids or as mRNA through a variety of delivery techniques including standard transfection, nucleofection, adenoviral infection, integration-defective lentiviral infection, or by glass-needle microinjection. From a clinical gene therapy perspective the most important use of this approach has been to create HIV resistant cells through the ZFN induced mutation in the CCR5 gene. CCR5 is a necessary co-receptor for infection of CD4+ T-cells by HIV. People with homozygous Δ32 mutations in CCR5, which results in non-functional CCR5 proteins, are resistant to HIV infection, but are otherwise phenotypically normal except for an increased risk of infection by West Nile Virus and Listeria monocytogenes. The hypothesis based on these finding is that by creating an immune system that is CCR5 mutated in a patient who was already infected with HIV, one could alter the course of the disease. In an exciting proof-of-principle experiment, an HIV infected patient underwent a bone marrow transplant for acute myelogenous leukemia using donor marrow that was homozygous CCR5 Δ32. Remarkably, after the transplant the patient's HIV viral load became undetectable and his CD4 count, derived from the CCR5 Δ32 donor, increased (Allers et al., 2011; Hutter et al., 2009). The results of this patient's treatment demonstrated that it might be possible to engineer the

immune system to become HIV resistant with a beneficial clinical outcome. Unfortunately, bone marrow transplantation using homozygous Δ32 donors is not a broadly feasible strategy for the treatment of HIV because most patients will not have an HLA matched homozygous Δ32 donor and because of the general toxicity from allogeneic bone marrow transplantation.

An alternative strategy would be to engineer the patient's own immune system to become HIV resistant by creating ZFN mediated knockout of the CCR5 gene in autologous cells, either CD4+ T-cells or hematopoietic stem cells. In pre-clinical studies, Carl June and co-workers have demonstrated that using CCR5 directed ZFNs, primary T-cell populations can be generated in which >50% of the CCR5 alleles have been mutated and that these T-cells are resistant to a laboratory CCR5 tropic strain of HIV (Perez et al., 2008). Similarly, Paula Cannon and her co-workers have demonstrated that the same CCR5 ZFNs can cause inactivation of the CCR5 gene in human CD34+ cells and that T-cells derived from these gene modified cells are resistant to HIV (Holt et al., 2010). Recently, it has been demonstrated that ZFNs can also be used to disrupt the cxcr4 locus (the co-receptor for r4 tropic HIV) in CD4+ T-cells and that transplantation of this cells into mice can confer some resistance to HIV infection (Yuan et al., 2012). Based on these pre-clinical findings for CCR5, clinical trials have opened based on this strategy. In the first clinical trials to open, patient derived CD4+ T-cells are infected with an adenovirus that expresses the CCR5 ZFNs causing CCR5 gene modification. The population of autologous T-cells is then expanded *in vitro* and then transplanted back into the patient. These trials have enrolled a number of

patients and the results, including whether the modified T-cells persist and are able to confer functional clinical protection against HIV, are pending.

Of the handful of published ZFN pairs designed to disrupt human genes, the CCR5 ZFNs have had the most pre-clinical and clinical success. Others, however, are being developed that are perhaps promising therapeutic tools. One such pair of ZFNs has been developed to recognize CAG triplet repeat tracts. CAG tracts are physiologically normal in multiple genes when the number of repeats in fewer than 30. However, expansion of repeat number leads to debilating diseases such as myotonic dystrophy, Huntigton's disease and spinocerebellar ataxia. ZFNs that bind these repeats were designed by modular assembly, and demonstrate the ability to contract CAG repeats in a reporter assay in human cells (Mittelman et al., 2009).

Though endogenous gene disruption is clinically relevant in perhaps only a few scenarios, an important point to consider is the impact that endogenous gene disruption could have for studying the biology of disease. As demonstrated, highly efficient targeting strategies now allow for robust and precise genetic manipulation in previously inaccessible cells types such as hESC/iPSC. Moreover, ZFNs have been used to create targeted gene mutations to create genetically modified organisms in a wide variety of species for which it previously was not possible to perform targeted gene modification. The ability to create precise genetically modified lines in species other than mouse is certainly going to result in improved models of human disease, improved species for agricultural

purposes, and in the long-term general improvements in human health as a result.

Gene Correction

By harnessing the capability of nucleases to stimulate gene targeting through homologous recombination, small, precise changes can be introduced into the genome including the correction of disease causing mutations. To achieve gene correction at an endogenous locus, site-specific nucleases must be designed to a sequence within the gene through one of the strategies described above. A donor DNA template must also be designed that is homologous to the endogenous gene and contains the small changes that are to be introduced into the genome (**Figure 1e**). The donor DNA template will be used by the homologous recombination machinery as the substrate to synthesize new DNA that will then be used repair the DSB. Finally, both components (the nucleases and the donor DNA template) must be able to be delivered to a cell-type that is capable of reconstituting the diseased tissue or organ. Translation of this therapeutic paradigm to the clinic has not yet been accomplished, but progress has been made *in vitro* towards its end, and work is currently underway to demonstrate the feasibility of the strategy in animal models.

The first examples of nuclease mediated gene correction of a chromosomally integrated gene in human cells were described in 2003 and 2005 (Porteus and Baltimore, 2003; Urnov et al., 2005). In these first examples of ZFN mediated gene correction, a mutated GFP gene was introduced as a single copy into the genome of cells to create a reporter line. The frequency of gene

correction was measured by determining the frequency that cells became GFP positive. When the donor DNA template was introduced without nucleases, the frequency of gene correction was on the order of 0.0001% (10^{-6}). In HEK-293 cells, the frequency of gene correction, when ZFNs that target the integrated reporter gene were co-transfected with the donor DNA template, increased to 0.5-2.2% (Porteus and Baltimore, 2003; Urnov et al., 2005). This frequency increased to ~10% when the cells were transiently arrested in G2/M by vinblastine treatment (Urnov et al., 2005). A similar stimulation in gene correction using the GFP reporter system was also found in human embryonic stem cells (hESCs) and human induced pluripotent stem cells (iPS) (Zou et al., 2009). Importantly, the gene corrected hESCs and iPS cells retained their pluripotency and did not develop any overt karyotypic abnormalities. Finally, gene correction rates of ~2% were obtained in primary mouse fibroblasts using a transgenic gene targeting mouse reporter line in which the mutated GFP gene was "knocked-in" to the ROSA26 locus (Connelly et al., 2010).

Urnov *et al* published the first example of gene correction in an endogenous human gene (Urnov et al., 2005). Mutations in the IL2RG gene cause SCID-X1, the most common form of severe combined immunodeficiency (SCID). ZFNs to exon 5 were engineered using the Sangamo strategy. Transfection of these ZFNs along with a gene correction DNA donor template, demonstrated gene correction efficiencies of ~20% in K562 cells and ~5% in primary T-cells. Lombardo *et al* built on this work and demonstrated that using integration defective lentivirus rather than standard transfection techniques,

similar rates of gene targeting could be obtained in HEK-293 cells, K562 cells, and an Epstein Barr Virus transformed lymphoblastoid cell line (Lombardo et al., 2007).

Recently, Sun *et al* has published the use of TALENs for targeting the human β-globin locus. In this work, the β-globin target sequence was inserted into a GFP gene. Generation of the double strand break by the β-globin TALENs and correction of the GFP gene from an exogenous plasmid template was used as the readout in this reporter assay (Sun et al., 2012). Though gene correction of the sickle cell mutation would be the ultimate goal for gene therapy, no targeting at the endogenous locus was demonstrated in this scenario. Another emerging concept in gene therapy is that instead of targeting small gene corrective mutations, one can target full-length cDNA within a disease locus which would provide generic protection for all mutations that may be found at a given locus for a given disease. Excitingly, recent work in our lab has demonstrated targeting of the β-globin cDNA at the β-globin locus in a cell line using another pair of β-globin TALENs (Richard Voit *et al*, manuscript submitted).

In addition to ZFNs, homing endonucleases have also been designed to stimulate HR at chromosomal loci of human diseases. The first example of this is an I-CreI derived homing endonuclease that targets XPC, a gene involved in the skin disorder Xeroderma Pigmentosum (Arnould et al., 2007). The endonuclease-binding site from XPC was cloned into the lacZ gene, disrupting expression of lacZ. Then the mutated lacZ construct was integrated into the genome of CHO-K1 cells, transfected with the designed homing endonuclease

and a lacZ repair template, and shown to successfully stimulate HR. Another gene relevant to human disease that has been targeted with a homing endonuclease is RAG1. RAG1 is essential for VDJ recombination in B- and T-cells, and mutations in RAG1 cause Omenn syndrome, a SCID variant. RAG1 homing endonucleases stimulate HR at the endogenous locus in HEK-293 cells at rates of up to 6% (Grizot et al., 2009; Munoz et al., 2011). Although the results of these two studies with homing endonucleases show promising initial targeting results, there have not yet been reports of the correction of disease-causing mutations in primary human cells.

Endogenous gene correction has several advantageous features. First, with modification of the endogenous gene, the resultant transgene product remains under the control of endogenous regulatory elements at its correct chromosomal location. This is crucial for maintaining the biological optimum of protein expression that may be lost with ectopically located transgene integrations. Second, direct gene correction, rather than virally mediated transgene addition, eliminates the concern regarding insertional onocgenesis from uncontrolled viral integration sites. Third, gene correction, in contrast to gene addition, could be used to treat dominant diseases or diseases in which the mutant protein has dominant features. An example of the latter is sickle cell disease which is caused by a point mutation in the beta-globin gene. In a gene addition strategy, the transgene will have to out-compete the two mutant genes that continue to express large amounts of mutant beta-globin. In contrast, after gene correction, one would be turning the homozygous state into a heterozygous

state, a non-disease causing condition, and issues of competition would be avoided.

The major challenges to translating nuclease mediated gene correction into therapy include developing high quality nucleases to disease causing genes, demonstrating that clinically relevant levels of gene correction can occur in the appropriate cell type and finally demonstration that gene corrected cells can correct the disease phenotype.

Safe Harbor Gene Addition

An alternative to direct gene correction is to use homologous recombination to target transgene addition to a safe harbor in the genome. While the defining features of a safe genetic safe harbor are still being established, in principle a safe harbor locus can be generally defined as one in which the insertion of a transgene has no aberrant physiologic consequences. Another feature of a safe harbor locus, in addition to its safety, is it must be a location in which a transgene can be expressed at the levels necessary to achieve therapeutic efficacy. A robust example of safe harbor utilization is the frequent targeting of the murine ROSA26 locus for knock-in mouse technology.

Targeted gene addition by homologous recombination can be achieved using nucleases engineered to recognize the specific safe harbor site. In addition, a donor DNA plasmid is engineered by inserting a transgene cassette (driving the expression of one or several genes) between arms of homology. The arms of homology can be as short as 0.5-1 kilobase in length and flank the site of the nuclease site. Targeted gene addition occurs when the homologous

recombination machinery uses the donor DNA template to repair the nuclease induced DSB. When the homologous recombination machinery uses the donor DNA template, it results in the insertion of the transgene cassette at the site of the break (**Figure 1f**). Moehle *et al* first described this technology using the previously described IL2R-γ ZFNs. The authors began by inserting a very small, in-frame tag of only 4 amino acids (RAKR, furin cleavage site) into the IL2R-γ gene. They found that they could achieve targeting up to 15% by this strategy in K562 cells. Then, to investigate the full potential of gene addition, the authors targeted a 900 base pair GFP cassette, an 1100 base pair GFP-PolyA cassette and a full transcriptionally active unit of a promoter-GFP that measured 1500 base pairs. The authors demonstrated targeting frequencies of 6%, 3% and 6% for these three constructs, respectively. Finally, the authors generated a donor plasmid with a very large insert (8kb) between homology arms that contained three independent, transcriptionally active units. Remarkably, the authors found by Southern blot analysis that 5-8% of the IL2R-γ alleles had been targeted (Moehle et al., 2007). These data highlight the potential for nuclease-mediated gene addition strategies by suggesting that even very large payloads can be efficiently and precisely targeted to safe harbor loci.

Lombardo *et al* described gene addition using the IL2R-γ ZFNs in a panel of human cell types using integration deficient lentivirus (IDLV) delivery of targeting components. First, the authors attempted targeting with a donor construct where GFP was flanked by IL2R-γ homology arms. They reported targeting efficiencies of up to 6.6% and up to 2.4% in K562s and in an Epstein

Barr virus transformed lymphoblastoid cell line, respectively. Of note, in the latter cell line, gamma chain expression was lost in GFP+ cells. Because SCID can be caused by many mutations in the IL2R-γ gene, the authors thought a potential treatment for the disease could be to target normal IL2R-γ cDNA to the endogenous locus, as a one shot approach to cover all of the different mutations (instead of designing novel ZFNs to target each mutated exon separately). The authors found that they could target IL2R-γ cDNA to the endogenous IL2R-γ locus at a frequency of up to 6% in their lymphoblastoid cell line and that IL2R-γ expression was then restored (Lombardo et al., 2007).

One site being developed as a potential safe harbor is the AAVS1 integration site on human chromosome 19. The AAVS1 locus encodes the ubiquitously expressed *PPP1R12C* gene. Because AAV integration at this site does not seem to be deleterious and because the AAVS1 site is ubiquitously expressed, it has characteristics that may make it a good safe harbor. To this end, Sangamo Biosciences developed a pair of zinc-finger nucleases that recognize the AAVS1 locus and found that they could target several transgenes to the AAVS1 locus in various cell types, including hESCs (DeKelver et al., 2010; Hockemeyer et al., 2009). In Hockemeyer et al, the authors demonstrated that thye could achieve three methods of transgene expression within this locus. First, they could successfully target a puromycin resistance cassette driven by the endogenous *PPP1R12C* promoter through a splice acceptor. Second, they could achieve constitutively expressed eGFP by targeting an eGFP cassette with cytomegalovirus (CMV) promoter with a chicken β-actin enhancer. Lastly, the

authors targeted the eGFP gene with a tetracycline response element and demonstrated that eGFP expression could be induced by administration of doxycycline (Hockemeyer et al., 2009). Later, the authors demonstrated that they could achieve similar targeting in iPS and hESCs with TALENs designed to the AAVS1 locus as compared to the ZFNs (Hockemeyer et al., 2011).

Recently, Zou *et al* has described an AAVS1 ZFN-mediated targeting strategy in iPS cells derived from a patient with X-linked chronic granulomatous disease (CGD). Patients with CGD have a deficiency of the p91-phox gene product and lack reactive oxygen species (ROS)-mediated microbicidal activity in neutrophils. As a result, these patients present with severe, recurrent, and atypical infections. In this paper, the authors described iPS cell derivation from patients with CGD, that, when differentiated into neutrophils, expressed no functional p91-phox protein and stained for reduced ROS production, compared to normal iPS cells. The authors targeted a p91-phox minigene ($gp91^{phox}$) to the AAVS1 locus in the iPS cells using mRNA for the AAVS1 ZFNs and achieved targeted transgene insertion in 75% (15/20) of the iPS clones they examined. The authors then excluded clones with targeting at more than one AAVS1 allele and any clones that had experienced small NHEJ-mediated deletions at the other, non-targeted AAVS1 allele. Of the three (3/20) remaining clones, all retained normal iPS morphology, stable expression of the $gp91^{phox}$ transgene (for more than 4 months) and pluripotency (as measured by embryoid body and teratoma formation assays). Importantly, when differentiated into neutrophils, oxidase activity was restored, correcting the original disease phenotype. The

authors also performed an interesting and important control for this experiment. Using the same CGD iPS cells, the authors infected a gp91^{phox} expressing lentivirus into the cells in an attempt to correct the disease phenotype as might have been attempted as an earlier non-targeted gene therapy strategy. Though gp91^{phox} was initially expressed at very high levels, only 4% of the resultant differentiated neutrophils were weakly oxidase positive at the end of the experiment. The authors suggested that this was because “a significant proportion of lentivector genomic inserts are subject to accelerated silencing” when the iPS cells were differentiated into neutrophils (Zou et al., 2011). This data highlights the importance of moving towards precise, site-specific genome engineering and away from random integration and virally-mediated genome engineering.

Another example of a potential safe harbor locus is CCR5. In addition to using ZFNs to create mutations in the CCR5 gene, there have also been reports using those ZFNs to target transgenes to that locus. One example shows the successful gene addition of erythropoietin (Epo) to CCR5 in up to 40% of human mesenchymal stromal cells (MSCs). Epo is a cytokine that drives the maturation of red blood cells in the bone marrow, and loss of Epo leads to severe anemia. Injection of the MSCs into immunodeficient mice resulted in increased levels of circulating Epo and higher hematocrit compared to controls (Benabdallah et al., 2010). These results indicate that ZFN-mediated targeting of plasma-soluble factors to a safe harbor locus may be a viable therapeutic option.

One question about safe harbor loci was recently addressed by van Rensberg *et al* concerning chromatin status of the locus in different cell types. Using chromatin immunoprecipitation and analysis of different histone modifications, the authors examined the chromatin state of the AAVS1 and CCR5 loci in iPS cells and CD34+ cells. They found that the AAVS1 site has a relatively open chromatin status in both cell types, while the CCR5 locus was predominately closed (van Rensburg et al., 2012). They hypothesized that this may have implications for targeting efficiency, simply in terms of nuclease accessibility to the site. In addition however, it could be theorized that transgene expression may suffer as well from targeting to a relatively inactive region. As a result, it might be preferable to select a robustly expressed target site through a different safe harbor gene addition strategy. As mentioned, definition of a safe harbor locus heretofore included the necessity that when the locus is disrupted, there should be no physiologic consequence. Recently, our lab has demonstrated that by targeting a coding sequence which results in the same amino acid sequence, but is not homologous to the target site at the nucleotide level, gene expression at the safe harbor locus can be preserved. Thus, any targetable locus in the genome can now act as a safe harbor site, without regard to endogenous gene disruption. This may allow for selection of target sites which are more accessible by nucleases and also provide superior transgene expression. This work is further discussed in Chapter 3.

Though endogenous gene correction could be considered the ultimate and paramount goal for gene therapy for loss-of-function genetic diseases, safe

harbor gene addition provides a significant amount of flexibility when compared to gene correction. For safe harbor targeting through nuclease-mediated strategies, only one pair of nucleases need be designed and optimized to achieve targeting of any number of therapeutic transgenes (as opposed to the individual tailoring that is required for endogenous gene correction). As briefly discussed, the effort put forth to bringing a pair of nucleases to clinical trials are substantial and a strategy that necessitates only one pair could be advantageous. Secondly, as highlighted by Benabdallah *et al* (Benabdallah et al., 2010), alternative cell types (such as MSCs) can be manipulated to secrete ectopic transgene products (such as Epo) and this strategy might only be available through gene addition. Lastly, gene addition provides the option to express transgenes through any number of mechanisms, including splice acceptor/endogenous promoter approaches, constitutive expression or response-element regulatable expression, as demonstrated by Hockemeyer *et al* (Hockemeyer et al., 2009).

Non-Clinical Applications for Targeted Genome Engineering

Site-specific genome engineering holds promise for use as a clinical therapeutic for many monogenic diseases. However, precise, genetic modification of human cells can also be developed as a tool to better understand more basic biological questions. One such example was published by DeKelver *et al* in 2010 (DeKelver et al., 2010). For this study, the authors utilized ZFNs directed to the AAVS1 locus to generate a panel of stable human cell lines. Typically, stable cell lines are generated by DNA transfection or viral infection of

a gene of interest along with a selectable marker (usually drug selection). The gene randomly integrates into the genome, integrants are selected for using the marker, and a clonal population is isolated. Unfortunately, many confounding factors can arise from the random integration and clonal selection that may alter the true phenotype that would have resulted from the genetic manipulation. Further, the transgene expression can be unstable due to silencing effects over time, lending to unpredictable results. DeKelver *et al* postulated that targeting trangenes to a safe harbor locus would circumvent these complications and allow for generation of more predictable and "isogenic" stable cell lines. To this end, the authors demonstrated targeting at the AAVS1 locus in a panel of cell lines (including K562, Hep3B, HEK293 and U2OS) and in two primary human cell types, fibroblasts and hESC. The authors targeted a variety of constructs to the AAVS1 locus that included 1) a GFP reporter in K562, Hep3B, and HEK293 cells where targeting frequencies ranged from 3-10% 2) glucocorticoid receptor response element luciferase reporter constructs in U2OS and 3) shRNA cassettes targeted to CD58 and to components in the mTOR pathway in K562, HEK293 and hESC. Though the scientific potential for targeted "isogenic" cell lines is clearly evident, and the ability to target a variety of constructs to many cell types was thoroughly demonstrated, a comparison between targeted "isogenic" stable cells and stable cell lines generated by conventional means was lacking in this report.

Similar to the generation of targeted stable cells lines is the application of genome targeting for protein tagging. Classically, cellular trafficking studies have

been accomplished by overexpression of a tagged fusion protein from plasmid DNA. Often, however, protein overexpression can perturb the cell's normal physiology and confound experimental results. A recent example of targeted protein tagging demonstrated that clarithin-mediated endocytosis could better be studied by the ZFN-mediated generation of fluorescent fusion proteins at their endogenous loci (Doyon et al., 2011a). Other studies have also demonstrated that ZFN-mediated gene targeting can be used to create tagged proteins that allow for lineage tracing. One example involved tagging Oct4 with eGFP as a reporter to monitor the pluripotent state of hESCs. In this same publication, the authors also targeted a non-hESC expressed locus (PITX3) with eGFP in hESC and iPS cells. They suggested that this strategy could be used as a reporter when the hESC or iPS cells were differentiated into neurons (where PITX3 is expressed), though this data was not reported (Hockemeyer et al., 2009). Interestingly, the authors provided a follow-up study and demonstrated that both the Oct4 and PITX3 experiments could be repeated with TALENs instead of ZFNs with similar efficacy (Hockemeyer et al., 2011).

Another non-therapeutic role of ZFNs is the demonstration that two distinct pairs of nucleases can stimulate chromosomal translocations when administered to cells simultaneously (**Figure 1c**). Using first the nuclease combination of I-SceI and ZFNs to the IL2R-γ locus and then the combination of the IL2R-γ and AAVS1 ZFNs, Brunet *et al* demonstrated the translocation of large chromosomal fragments. In the IL2R-γ /AAVS1 experiments, reciprocal translocations between the X-chromosome (site of IL2R-γ) and chromosome 19 (AAVS1 site) were

observed at a frequency of 10^{-4} in HEK-293 cells and $2x10^{-6}$ in human embryonic stem cells (Brunet et al., 2009). Although these rates are several orders of magnitude lower than rates of gene targeting at a single locus, this strategy provides a framework to study the mechanisms of chromosomal rearrangement and the factors required for this process. Additionally, developing nucleases to chromosomal sites that have been implicated in oncogenic translocations would provide a way to study the expression and regulation of the resulting fusion proteins from their endogenous loci. The demonstration that ZFNs can induce translocations, albeit at very low frequencies, also has implications for how to evaluate the safety of nucleases as they proceed into human clinical trials.

Another application of ZFNs for chromosomal manipulation is the use of two distinct pairs that recognize sites on the same chromosome to induce deletions. Using ZFNs designed to the adjacent genes CCR2 and CCR5, Lee *et al* used a simple PCR approach to detect deletions of up to 15 kb within the gene cluster. Further, when the authors utilized ZFNs upstream to the CCR5 locus, they were able to induce extremely large chromosomal deletions of up to 15 Mbp (Lee et al., 2010). Similarly, Sollu *et al* described chromosomal deletions at a frequency of 10% using two pairs of nucleases designed to different sites within the HOXB13 locus (Sollu et al., 2010). In this way, it may now be possible to achieve the targeted deletion of entire gene clusters, introns or even specific exons in human cells. While the ability to create specific chromosomal rearrangements by nucleases has important research implications, it also has important implications for their translation into a clinical tool as the induction of

unintended gross chromosomal rearrangements would be a serious adverse event.

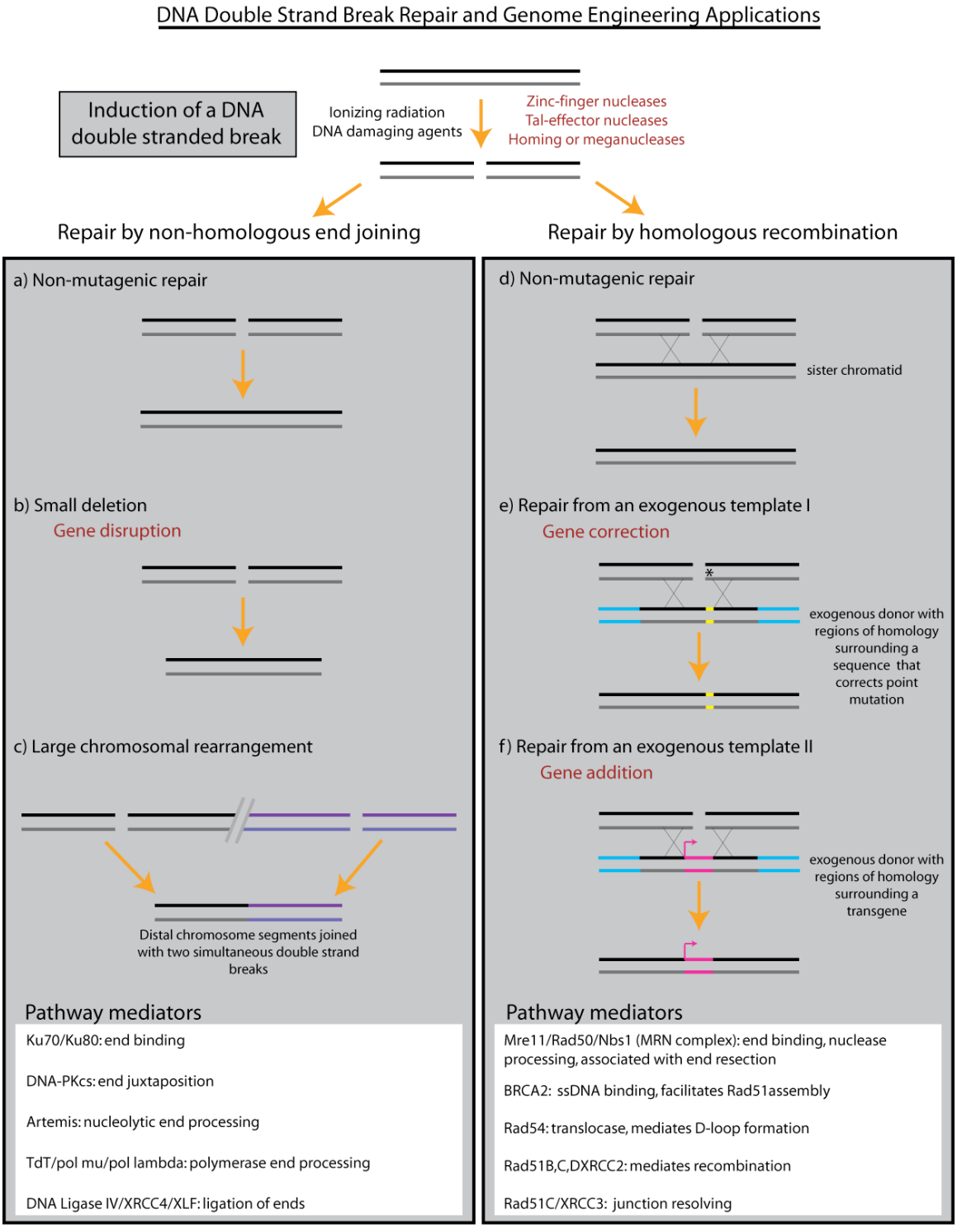


Figure 1. After the induction of a double strand break, repair can proceed down 2 pathways, non-homologous end joining (left panel) or homologous recombination (right panel). For genome engineering purposes, repair by NHEJ can be used to intentionally disrupt a gene (b) or create chromosomal rearrangements by simultaneously introducing two pair of nucleases (c). If repair occurs by homologous recombination, exogenous donor DNA can be used to direct correction of an endogenous locus (e) or insert an entirely new gene into a safe harbor locus (f).

Obstacles and Safety Concerns for Targeted Genome Engineering

Since the first report on using ZFNs to create targeted genome modifications, it has been evident that expression of ZFNs in mammalian cells can result in reduced cell viability(Porteus and Baltimore, 2003). The reduced cell viability is directly correlated with the expression level of the ZFN as transfecting more ZFN expression plasmids results in a higher degree of cell toxicity (Pruett-Miller et al., 2008). The mechanism of reduced cell viability is directly related to the creation of off-target DNA double-strand breaks. A point mutation in the nuclease domain that inactivates its catalytic activity, for example, eliminates the cytotoxicity of ZFNs. Moreover, staining for 53BP1 foci, a marker for double-strand breaks, demonstrates that increased toxicity is directly correlated with increased numbers of 53BP1 foci (Pruett-Miller et al., 2008). Elimination of increased 53BP1 foci also eliminates the reduced cell viability associated with ZFN cell expression (Pruett-Miller et al., 2009). Intriguingly, TALENs seem to have less cytotoxicity than ZFNs when expressed in cells, although they have not been studied to the same depth as ZFNs. Nonetheless, the apparent decreased cellular toxicity of TALENs is a promising feature of this alternative nuclease platform.

The identification of the sites of the ZFN off-target double-strand breaks has been a challenging problem. Gabriel *et al* have used the capture of extra-chromosomal DNA at sites of double-strand breaks in a non-homologous fashion to identify such sites and identified a number of genomic targets for a specific ZFN pair that had not been previously recognized. The nature of their approach, however, clearly biased against identifying rare off-target sites and underestimated the number of off-target sites (Gabriel et al., 2011). Pattanayak *et al* used an entirely different strategy to identify off-target sites for the same set of ZFNs studied by Gabriel *el al* in the same cell type (Pattanayak et al., 2011). Interestingly, while both groups confirmed a number of new off-target loci, the set of sites were completely non-overlapping from the two different approaches. Thus, the full set of off-target sites, even for this highly studied nuclease pair, has not been described and it is likely that further improvements will be needed to generate a comprehensive or near comprehensive list. Both of these studies evaluated the CCR5 ZFN pair that is now in clinical trials. Prior to entering clinical trials, this pair was introduced into large numbers of mouse hematopoietic progenitor cells, the ZFN exposed cells were then transplanted into mice and the transplanted mice were observed for tumors. There was no increase in tumor formation in the transplanted mice but it is difficult to assess the significance of this heroic study as there are no positive controls for such studies to validate how this assay would relate to the probability of tumor formation in human cells transplanted after exposure to ZFNs.

The reduced cell viability after ZFN exposure has prompted an investigation into strategies to reduce this effect. Two major strategies have been developed. The first is to modify the nuclease domain such that a nuclease pair can only cleave DNA as heterodimers ("obligate heterodimer nuclease domains). Theoretically and now confirmed experimentally, this reduces the number of off-target breaks by ~50%. The functional result is a marked improvement in cell viability when using nuclease pairs that contain the obligate heterodimer nuclease (Doyon et al., 2011b; Miller et al., 2007; Pruett-Miller et al., 2008; Szczepek et al., 2007). In some cases, however, the use of the obligate heterodimer nuclease domains also results in reduced on-target activity (Wilson and Porteus unpublished data). An alternative strategy is to regulate the expression level of the ZFNs. Since the targeted genome modification is a "hit and run" strategy, sustained nuclease expression is not needed and, in fact, sustained nuclease expression can result in increased cell toxicity. Pruett-Miller *et al* demonstrated that using small molecule regulation of ZFN protein level that on-target activity can be maintained while decreasing the cell toxicity and off-target foci formation to background levels (Pruett-Miller et al., 2009). The importance of regulating ZFN protein level both in amount and over time is confirmed by the off-target studies of Pattanayak *et al* (Pattanayak et al., 2011).

Summary and Research Goals

The ability to precisely modify the genome of human cells has the potential to shift the paradigm of how to treat monogenic and other human diseases. Instead of repeated infusions of deficient proteins like Factor VIII to

treat hemophilia or ADA to treat one form of SCID, which are very laborious and expensive regimens, gene therapy holds the promise of a robust, long-term cure. Additionally, gene therapy can address the root cause of other monogenic diseases like sickle cell disease and X-linked SCID, in which the deficiency is not a secreted protein, cannot be treated by protein infusion, and can currently only be cured through allogeneic stem cell transplants. Targeted modification of a patient's own stem cells avoids the side-effects and reliance on immunosuppressive drugs associated with these allogeneic transplants and is now under investigation as a way to cure HIV infection. The targeted therapy approaches highlighted in this introduction address and improve the problems associated with classic retroviral-mediated transgene addition. Targeted methods do not rely on random integration that can lead to oncogenesis, and they are more broadly applicable because they can be used not only for gene addition, but also for gene correction and gene disruption. As the specificity of these methods continues to be improved and more clinical data from patient trials becomes available, targeted genome engineering will perhaps on day fulfill the promise of a cure for monogenic diseases.

A second emerging concept is that patient-derived cells can be modified to secrete proteins that augment a fundamental biologic process. In this setting, patient cells are serving a vehicles for systemic or local protein delivery instead of participating in the correction of a genetic defect. This technology may have the advantage of constituitive or regulatable secretion of proteins within a biologic context, as opposed to high-dose bolus treatment with recombinant proteins.

The central aim of this thesis is to thoroughly develop a mouse model for the study of both gene correction and safe harbor gene addition without safe harbor disruption. Many aspects of genome engineering require extensive optimization which underscores the need for such a model. Examples of this include: 1) How can we isolate cells of interest for genome engineering and how best to introduce foreign DNA into them? 2) Which types and quantity of site-specific nucleases perform the best for gene targeting in these cell types? 3) Which donor plasmids are most effective for conferring the desired outcome, in terms of targeting efficiency and protein expression? 4) Which selection or purification strategies are most effective and allow for the best cell survival in these cell types? 5) What are the best techniques for introducing modified cells into a recipient? The mouse model we describe here is capable of addressing these detailed, technical questions through the use of an easily assayable and event-specific GFP reporter for both endogenous gene correction and safe-harbor gene addition, and in doing so, provides essential information to the field of genome engineering.

Chapter II- GENE CORRECTION BY HOMOLOGOUS RECOMBINATION WITH ZINC FINGER NUCLEASES IN PRIMARY CELLS FROM A MOUSE MODEL OF A GENERIC RECESSIVE GENETIC DISEASE

This data has been previously published in co-authorship with Jon Patrick Connelly, Ph.D.

Abstract

Zinc finger nucleases have been used in a wide variety of systems to create precise genome modifications at frequencies that might be therapeutically useful in the field of gene therapy. One barrier to translating the use of zinc finger nuclease mediated homologous recombination is the lack of an animal model to establish a potential therapeutic paradigm in humans. To address this limitation, we have created a mouse model of a generic recessive genetic disease. We first knocked a mutated GFP gene into the ROSA26 locus in murine embryonic stem cells. Using these knock-in embryonic stem cells we created a transgenic mouse and derived lines that were either heterozygous or homozygous for a knock-in of the mutated GFP gene at the ROSA26 locus. We used zinc finger nucleases that target the GFP gene to determine the rate of gene targeting in different primary cells from this mouse model. We achieved targeting rates from 0.17- 6% in different cell types, including primary fibroblasts and astrocytes. We demonstrated that *ex vivo* gene corrected fibroblasts could be transplanted back into a mouse where they retained the corrected phenotype. In addition, we achieved targeting rates of over 1% in embryonic stem cells and demonstrated that targeted embryonic stem cells retained the ability to differentiate into multiple cell types from all three germ line lineages. In summary, potentially therapeutically relevant rates of zinc finger nuclease

mediated gene targeting can be achieved in a variety of primary cells, and this work continues to support the possibility of using zinc finger nuclease technology for therapeutic purposes.

Introduction

Conceptually, the simplest application of gene therapy is for diseases caused by mutations in a single gene, the so-called monogenic diseases. While millions of people suffer from monogenic diseases, a cure is only possible for a small fraction for whom either hematopoietic stem cell transplantation or organ transplantation is available. In contrast, gene therapy uses the patient's own cells and has the potential to cure many of these diseases. In the last decade, several clinical trials have been carried out which have highlighted both the promise of gene therapy (the benefit of tens of patients with severe combined immunodeficiency (Aiuti et al., 2009b; Hacein-Bey-Abina et al., 2002) a handful of patients with Leber's congenital amaurosis (Bainbridge et al., 2008; Maguire et al., 2009) and two patients with X-linked adrenoleukodystrohy (Cartier et al., 2009) from gene therapy based on viral delivery), and the potential harm from the uncontrolled integrations of the viral vectors used to deliver the therapeutic transgene (Hacein-Bey-Abina et al., 2003a; Hacein-Bey-Abina et al., 2003b). An alternative to using integrating viruses is to use gene targeting by homologous recombination to precisely control the genomic modification either through directly correcting a mutation or through controlling the site of transgene integration (Lombardo et al., 2007; Porteus and Carroll, 2005).

The natural rate of gene targeting by homologous recombination, hereafter referred to as "gene targeting," is $1x10^{-5}$ -$1x10^{-8}$ (Doetschman et al., 1987; Hasty et al., 1991; Porteus and Baltimore, 2003; Sargent et al., 1997) and is too low to be therapeutically useful. This barrier has been overcome by the recognition that the creation of a gene specific DNA double-strand break can stimulate gene targeting several thousand fold (Brenneman et al., 1996; Choulika et al., 1995; Donoho et al., 1998; Porteus and Baltimore, 2003; Rouet et al., 1994b; Sargent et al., 1997; Smih et al., 1995), potentially to therapeutically relevant levels for certain diseases. To translate this finding to the field of gene therapy, it was necessary to devise a method to generate site specific DNA double-strand breaks. There have been two major approaches to this problem. The first is to re-design homing endonucleases to recognize target sites in endogenous genes (Arnould et al., 2006; Ashworth et al., 2006; Epinat et al., 2003; Grizot et al., 2009; Paques and Duchateau, 2007; Redondo et al., 2008). The second is to design zinc finger nucleases to recognize target sites in endogenous genes. Zinc finger nucleases (ZFNs) are artificial proteins in which the non-specific nuclease domain from the FokI restriction endonuclease is fused to a zinc finger DNA binding domain (reviewed in (Durai et al., 2005; Porteus and Carroll, 2005; Wu et al., 2007)). A ZFN can have 3-6 individual zinc finger domains arranged in tandem recognizing a target site 9-18 basepairs long. Additionally, the FokI nuclease domain functions as a dimer (Bitinaite et al., 1998; Smith et al., 2000). Therefore, a pair of ZFNs must be engineered to bind the target site in a way that permits the nuclease domain to dimerize and create

the double-strand break. Thus, even with a pair of 3-finger ZFNs, the full target site is 18 basepairs long. An 18 basepair sequences should only occur once in the mammalian genome based on probability and can be empirically determined for any given sequence by BLAST searches. There are a number of different approaches to engineer ZFNs, each of which has their advantages and disadvantages (reviewed in Cathomen and Joung (2008) (Cathomen and Joung, 2008)). Nonetheless, ZFNs have been successfully engineered to a wide variety of different gene targets in a range of different species (Beumer et al., 2006; Bibikova et al., 2003; Bibikova et al., 2002; Carroll, 2008; Doyon et al., 2008; Geurts et al., 2009; Meng et al., 2008; Morton et al., 2006; Porteus and Baltimore, 2003; Shukla et al., 2009; Townsend et al., 2009; Urnov et al., 2005). These ZFNs have been used to generate high rates of precise genome modifications either by the use of mutagenic non-homologous end-joining (in which short insertions or deletions are created at the site of the ZFN induced double-strand break) or by the use of gene targeting, including creating genetically modified zebrafish and rats (Doyon et al., 2008; Geurts et al., 2009; Meng et al., 2008).

In human cells, ZFNs have been used to stimulate gene targeting in a variety of different cell lines. The most recent advances demonstrated that ZFNs can stimulate gene targeting in human embryonic stem (ES) cells and induced pluripotent stem (iPS) cells (Hockemeyer et al., 2009; Zou et al., 2009). Moreover, Perez et al. demonstrated that human T-cells modified at the CCR5 gene by mutagenic repair of a ZFN induced double-strand break could survive

when transplanted back into an immunodeficient mouse (Perez et al., 2008). Nonetheless, to date there has been no easy way to model a therapeutic paradigm in which patient derived cells are precisely modified by ZFN mediated gene targeting *ex vivo* and then transplanted back into the animal, as might be done when trying to treat a patient with a genetic disease. To this end, we have created a mouse model of a generic recessive genetic disease in which a mutated GFP gene has been knocked-in to the ubiquitously expressed ROSA26 locus (Friedrich and Soriano, 1991; Zambrowicz et al., 1997). Using this model, we show that gene correction of 0.17-6% of murine ES cells, novel ROSA-3T3 cell lines, primary embryonic fibroblasts, primary adult fibroblasts, and primary astrocytes can be achieved. We also demonstrate that gene corrected cells can be transplanted back into a recipient mouse and the transplanted cells retain their gene corrected phenotype.

Results

Generation of a Mouse Model of a Generic Recessive Genetic Disease

A critical aspect of translating the use of ZFN mediated gene targeting to clinical use is to develop appropriate animal models to evaluate the feasibility of the strategy. While there are many mouse models of human genetic diseases, these models are generally not useful for modeling ZFN mediated gene targeting for two major reasons. The first is that while the mouse and human genomes are very similar, there is enough sequence divergence that ZFNs designed to target a human gene will usually not recognize the mouse target. The second is that most mouse models of human genetic diseases are not exact genetic models of

the human disease, but instead are phenotypic mimics. Below we describe our use of a new mouse model to study ZFN mediated gene targeting. This model provides a platform to directly determine the efficiency of gene targeting in a specific cell type and whether targeted cells can be successfully transplanted back into a host. We created a mouse model of a generic recessive genetic disease by knocking-in a mutated GFP gene into the ROSA26 locus using standard homologous recombination technology in murine ES cells (**Figure 2.1a**). From two of the targeted ES cell clones we generated transgenic mouse lines in which either one (ROSA26$^{GFP*/+}$) or both alleles (ROSA26$^{GFP*/GFP*}$) of the ROSA26 locus contain a knock-in of the mutated GFP gene. The mutation in the GFP gene consists of an 85 nucleotide sequence which includes both an in-frame stop codon and the recognition site for the I-SceI homing endonuclease. This insertion is 12bp downstream from the ZFN target site in the GFP gene that we have previously designed and validated several different pairs of "GFP"-ZFNs to recognize (Pruett-Miller et al., 2008). This animal model mimics the GFP gene-targeting system that we have used to better understand ZFN mediated gene targeting in tissue culture cells and human ES and iPS cells (Porteus, 2006; Porteus and Baltimore, 2003; Pruett-Miller et al., 2008; Zou et al., 2009). In this system, cells containing the integrated GFP* gene are transfected with two plasmids that each express a ZFN, and a third "donor" plasmid that carries sequence information needed to correct the mutation (by serving as a donor template during homologous recombination) (Durai et al., 2005; Porteus and Baltimore, 2003). Because the mutated GFP gene is knocked-in to the

ubiquitously expressed ROSA26 locus, this model can be used to study the efficiency of gene correction by gene targeting in potentially every cell type in the mouse. There are two possible strategies to use gene targeting for gene therapy. The first is to attempt gene targeting *in vivo* and directly correct disease-causing mutations in cells without removing them from the animal beforehand. The second is to purify cells from the animal first, and correct the mutation *ex vivo* prior to transplanting back into the animal. Here we report on our results at purifying primary cells from this generic genetic disease model and using ZFN mediated gene targeting to directly correct the mutation in the GFP gene and then transplanting those cells back into a mouse.

ZFN Mediated Gene Targeting in Murine Embryonic Stem Cells

While we were generating the mouse line from the ROSA26$^{GFP*/+}$ targeted ES cell clones, we performed a series of targeting experiments in the ES cell clones (**Figure 2.2**). To perform these experiments we transfected ES cells with three plasmids—an EGFP donor vector that contains sequence information necessary to correct the mutation in the integrated target but is non-fluorescent because it is missing the first 37 nucleotides of the coding region, and two ZFN expression plasmids where the ZFN is driven by the ubiquitin C promoter. In prior work we have shown that titrating the amount of the donor plasmid and nuclease expression plasmid is important in maximizing the rate of gene targeting (Porteus, 2006; Pruett-Miller et al., 2008). In the gene targeting titration experiment we varied the amount of donor plasmid and ZFN expression plasmids while keeping the total amount of DNA transfected the same. We achieved the

maximal rate of targeting when the transfection mix consisted of ~90% donor plasmid and ~10% of ZFN expression plasmids (**Figure 2.2a**). ROSA26$^{GFP*/+}$ ES cells that were transfected with only the donor plasmid without ZFNs had an absolute targeting rate of less than 0.001% (data not shown). The targeting rate with ZFNs was 0.20%. An important concern involving the use of ZFNs to create DNA double-strand breaks is the potential for either off-target cleavage or off-target plasmid insertion. We have previously reported that the ZFNs used in this study demonstrate low levels of off-target toxicity in the murine embryonic stem cells used in this paper (Pruett-Miller et al., 2008), and no readily visible toxicity was observable in these experiments. In the present study, we did not investigate the potential for off-target plasmid integration. We also tested whether exposing the cells to vinblastine would increase the rate of targeting in murine ES cells as has been demonstrated in other cell types arrested at the S/G2 phase of the cell cycle (Maeder et al., 2008; Potts et al., 2006; Urnov et al., 2005), and found that vinblastine exposure for the first 15 hours after transfection increased the rate of gene targeting in murine ES cells by approximately 4-5 fold to an overall rate of 1.6% (**Figure 2.2b**), at this concentration of vinblastine (100nM) no cytotoxic effects were readily apparent. We also compared two different pairs of ZFNs that target the same sequence in the GFP gene (Table 1) and found that ZFN pair 3/4 was approximately 50% better than pair 1/2 (the pair we have previously published and characterized most extensively (Pruett-Miller et al., 2008)).

To determine if the targeted ES cells retained pluripotency, we evaluated whether the targeted cells could differentiate into all three germ line lineages using a teratoma formation assay. GFP^+ targeted ES cells were purified using fluorescence-activated cells sorting (FACS), and these ES cells were injected subcutaneously into nude mice to create fluorescent teratomas (**Figure 2.3**). Histological examination of the teratomas showed that targeted ES cells formed tissues representative of all three germ line lineages—endoderm, mesoderm, and ectoderm. In summary, ZFNs can stimulate gene targeting in murine ES cells to an absolute rate of 1.6% (a greater than 1000-fold stimulation over the targeting rate without ZFNs) and these targeted cells retain the potential to differentiate into all of the major cell lineages.

Gene Targeting in ROSA-3T3 Cells

We generated immortalized fibroblast cell lines "ROSA-3T3s" from both $ROSA26^{GFP*/+}$ and $ROSA26^{GFP*/GFP*}$ mice using a standard fibroblast immortalization protocol (Todaro and Green, 1963), and performed a gene targeting titration experiment with varying amounts of ZFN expression plasmids and donor plasmid (**Figure 2.4a**). As in the murine ES cells, we observed the maximal rate of targeting when the transfection mix contained ~90% donor plasmid and ~10% ZFN expression plasmids. Gene targeting rates reached a maximum of 1.8% and 6.7% in heterozygous and homozygous lines respectively. Vinblastine exposure did not affect targeting rates in these lines (data not shown).

Gene Targeting in Primary Embryonic and Adult Fibroblasts

We also studied gene targeting rates in fibroblasts derived from both embryonic (E13.5) and adult (3-6 months old) mice. In these experiments we used nucleofection rather than Lipofectamine 2000 as the method of transfection and found that increasing the amount of donor plasmid while keeping the amount of ZFN expression plasmid constant increased the frequency of gene targeting (**Figures 2.4b and 2.4c**). The maximal gene targeting rate in primary adult fibroblasts from ROSA26$^{GFP*/GFP*}$ mice was over 2% (**Figure 2.4b**). In murine embryonic fibroblasts the maximal rate of gene targeting was ~1.8%. Interestingly, the targeting rate in murine embryonic fibroblasts from heterozygous mice was significantly different when low amounts of donor was transfected but this difference disappeared when the highest amount of donor plasmid was transfected (**Figure 2.4c**). In summary, ZFN mediated gene targeting rates of ~2% can be achieved in primary fibroblasts of either adult or embryonic origin.

Gene Targeting in Primary Astrocytes

The final of primary somatic cell type that we examined for ZFN mediated gene targeting was astrocytes, another readily isolatable and robustly cultured cell type. Using nucleofection with a constant amount of ZFN expression plasmid and increasing amounts of donor plasmid, we found that 0.03-0.17% of primary astrocytes could be targeted (**Figure 2.4d**). As in all of the other cell types examined, the rate of targeting increased as the amount of donor plasmid transfected increased. Overall this rate of targeting is lower than what we found in primary fibroblasts. This difference may reflect an intrinsic difference between

astrocytes and fibroblasts but may also reflect an underestimation of the rate of targeting in astrocytes because of a decreased ability to detect targeted cells secondary to the low level of GFP expression from the ROSA26 promoter in astrocytes compared to ROSA-3T3s and primary fibroblasts (**Figure 2.4a,c,d**). As in the immortalized cells and primary fibroblasts, there was no effect of vinblastine exposure on the rate of targeting in primary astrocytes (data not shown).

Transplantation of Gene Corrected Primary Mouse Fibroblasts into an Immunocompetent Recipient

To perform cell based gene correction therapy clinically, gene corrected cells must be able to survive transplantation back into a recipient. A potential treatment for hemophilia is to modify fibroblasts ex vivo to secrete Factor VIII or Factor IX, followed by transplanting the cells back into recipients. This has already been performed with transient but significant clinical benefits observed in mice, rats, and human patients (Palmer et al., 1989; Qiu et al., 1996; Roth et al., 2001). As mentioned above, ZFNs may be toxic to cells. Although no toxic effects were readily apparent upon microscopic examination of treated fibroblasts, we further looked for toxic effects using the previously described toxicity assay (Pruett-Miller et al., 2008). When fibroblasts were subjected to this assay, a small degree of toxicity was measured (**Figure 2.5a**). This was minor in comparison with a previously published pair of ZFNs designed to target the CCR5 locus (Perez et al., 2008), and which exhibit noticeable toxic effects on cells. Both pairs of ZFNs contained the wild-type nuclease domain and were not

modified to prevent homodimerization. To test whether our gene corrected fibroblasts were capable of transplantation, we isolated primary adult fibroblasts and performed another round of gene targeting. On day 6 post-nucleofection, we analyzed a portion of the total cell population and found that 0.21% of the total cells had undergone gene correction. We next took the remaining population, (without selecting for GFP positive cells), embedded the cells in matrigel and transplanted the cells subcutaneously into an immunocompetent, sibling mouse. Two weeks after transplantation we excised the matrigel plug, cultured the isolated cells for 6 days, analyzed the transplanted cells by fluorescence microscopy and found that 0.10% of the isolated cells were GFP positive. The lower frequency of targeted cells may be due to contaminating host cells that were excised with the plug or as a result of cell death after transplantation (**Figure 2.5b).** This study demonstrated that *ex vivo* zinc finger nuclease gene corrected fibroblasts could be successfully transplanted.

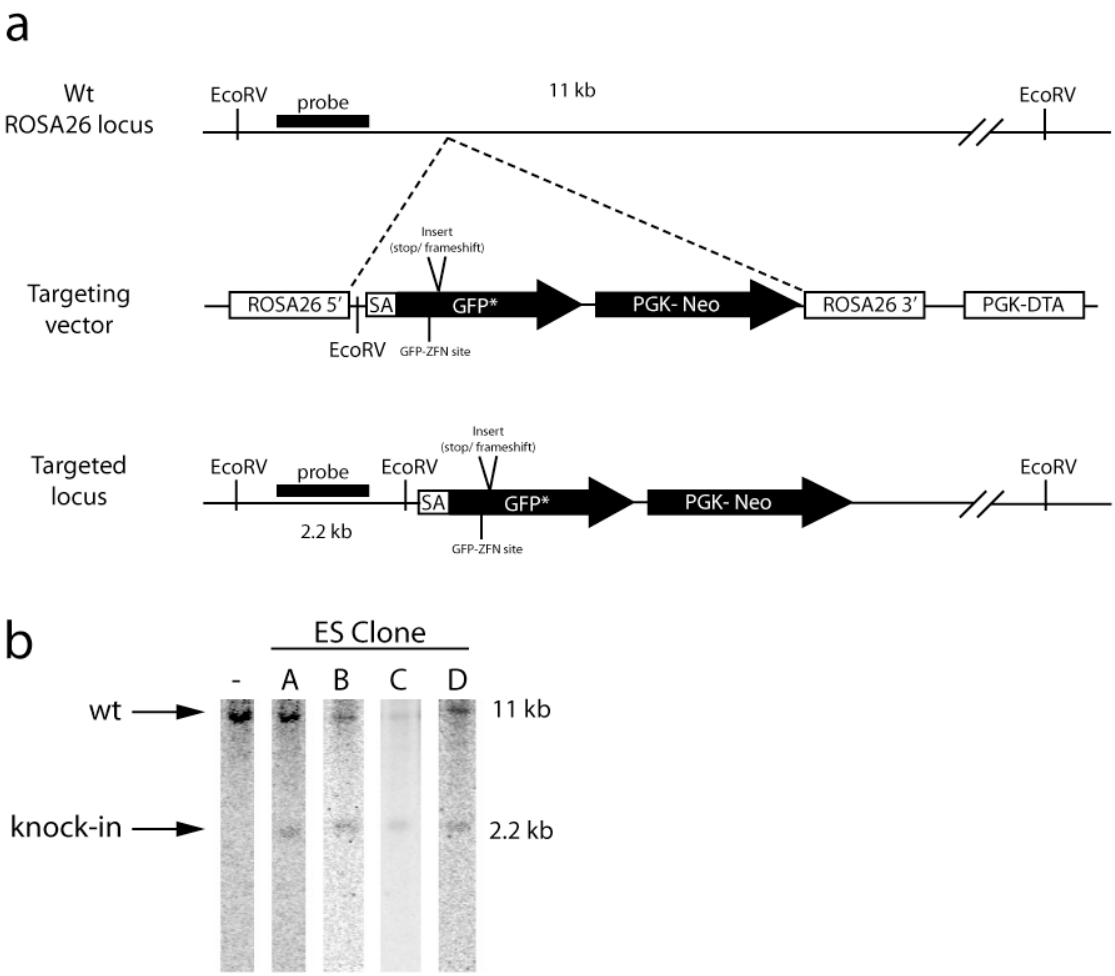


Figure 2.1: Construction of targeting vector and screening of ES clones. (a) Schematic of targeting vector and screening strategy. Correct knock-in of the GFP* reporter cassette to the ROSA26 locus causes the addition of an upstream EcoRV site resulting in a 2.2kb fragment upon digestion. (b) Southern blot analysis of correct knock-in ES clones. Genomic DNA was purified from ES clones, digested with EcoRV and used for Southern analysis. SA, splice acceptor; PGK-Neo, Neomycin resistance cassette; PGK-DTA, diphtheria toxin cassette.

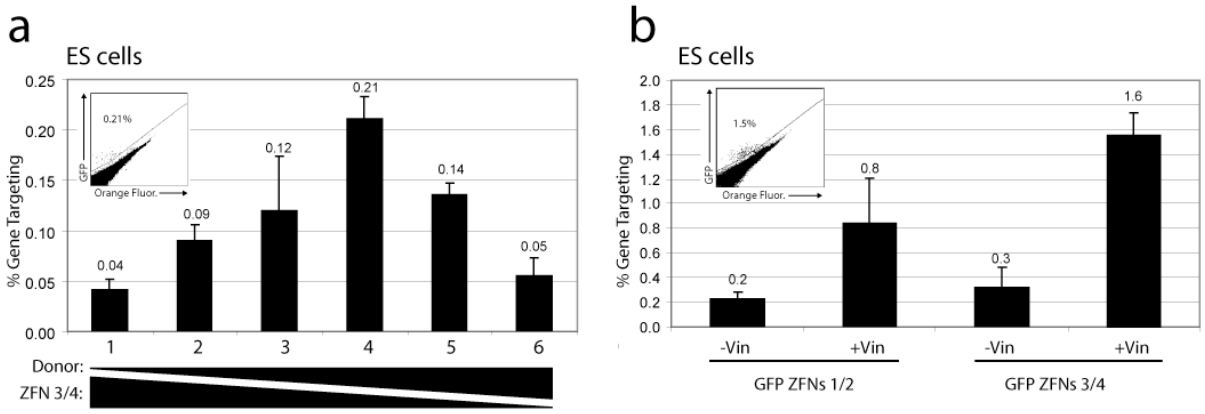
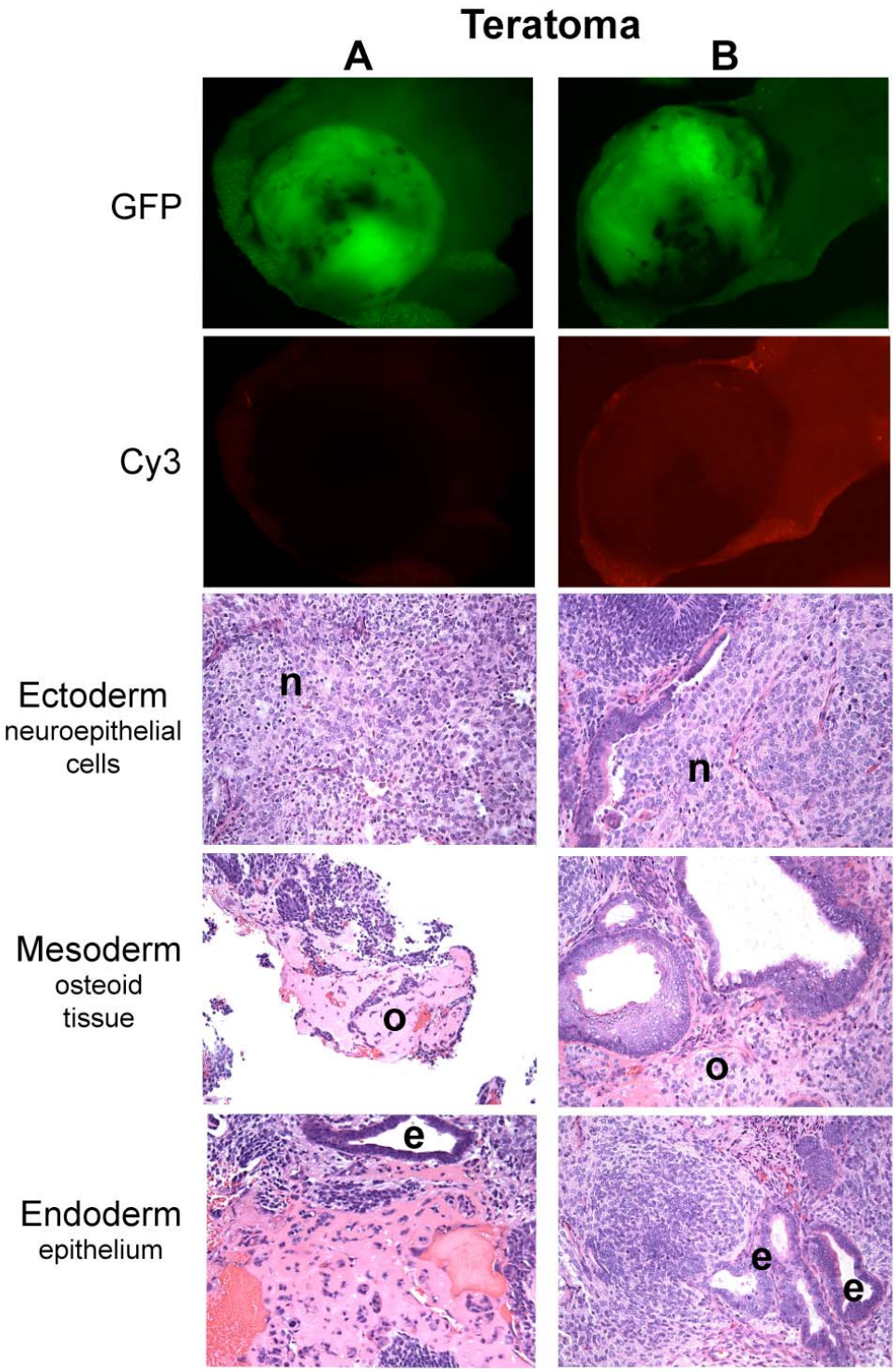


Figure 2.2: Gene targeting in ES cells. (a) Titration of donor plasmid and ZFNs in ES cells. Cells were transfected using Lipofectamine 2000 with different amounts of donor and ZFN plasmids. From left to right, amounts of donor plasmid increased while ZFN amounts decreased. Transfection was performed with Lipofectamine 2000 and the indicated amounts of donor plasmid and ZFNs in each lane, indicated as Donor (ng), ZFN1 (ng)/ ZFN2 (ng). (1) 100, 350/350; (2) 400, 200/200; (3) 600, 100/100; (4) 700, 50/50; (5) 750, 25/25; (6) 775, 13/13. Fifteen hours post transfection, media was changed to normal ESLX. Gene targeting events were analyzed 4 days post transfection using flow cytometry. (b) Gene targeting in ES cells using two sets of GFP-ZFNs, with and without vinblastine treatment. Cells were plated in ESLX with and without vinblastine (100nm). Transfection mix was added, and 15 hours later removed and replated with ESLX. Gene targeting events were analyzed as in (a). In the upper left of each graph is a representative flow cytometry plot after targeting in which GFP fluorescence is measured in the y-axis and background orange fluorescence along the x-axis. The number in the left corner of the flow plot is the percentage of GFP(+) cells.



'e 2.3: Teratoma formation assay. Two teratomas resulting from subcutaneous injection of d cells were harvested from nude mice. (a) Teratomas were initiated by gene targeted ES cells as strated by GFP fluorescence. Teratomas were also photographed under a Cy3 filter to show ound fluorescence. (b) Sections were cut and stained with hematoxylin and eosin to identify tructures indicated in figure. n, neuroepithelial cells; o, osteoid tissue; e, epithelium.

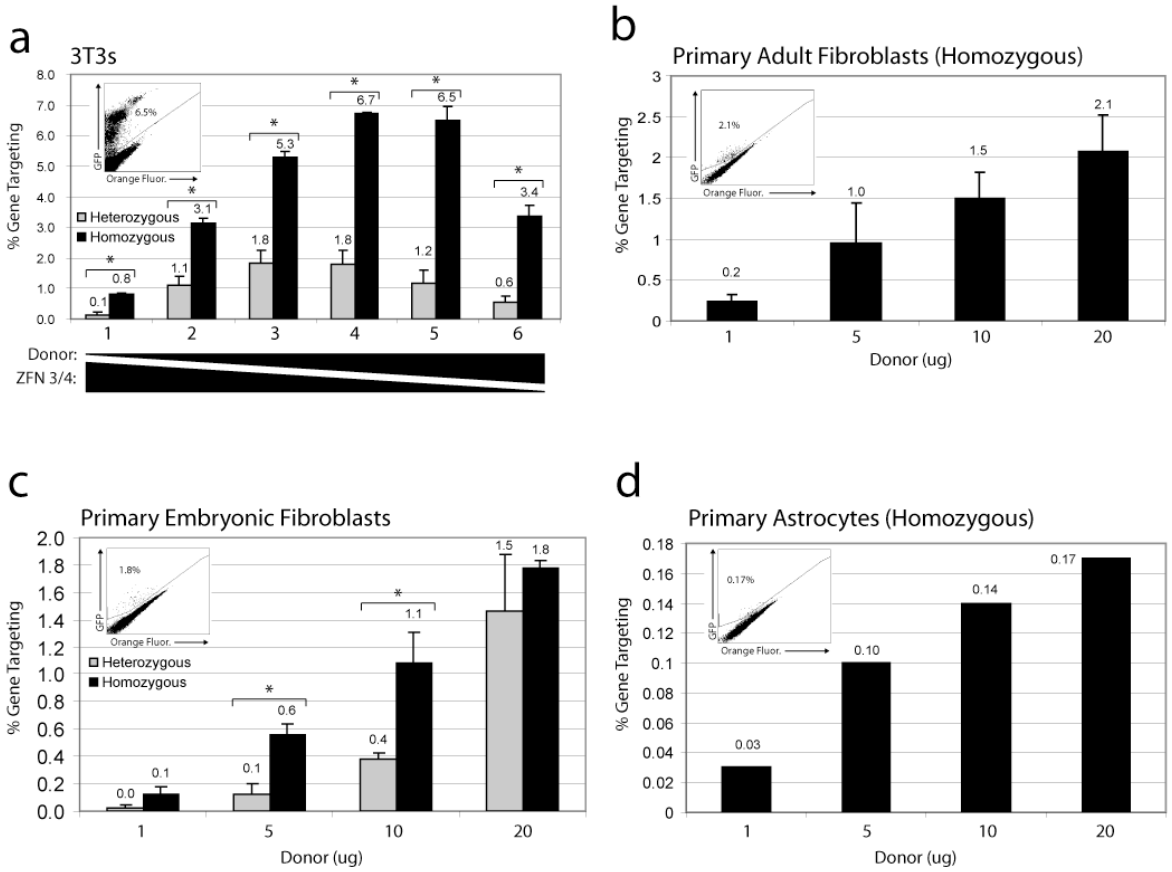


Figure 2.4: ZFN-mediated gene targeting in primary cells. (a) Gene targeting in homozygous and heterozygous ROSA-3T3s: transfections were performed using Lipofectamine 2000 with the indicated amounts of donor plasmid and ZFNs. The next day, media was changed and on day 4 gene targeting events were analyzed. (b,c) Gene targeting in MAF/ MEFs: Transfection of plasmids was performed by nucleofection using 2 ug of each ZFN and the indicated amounts of donor plasmid. Gene targeting events were analyzed 4 days post transfection. (d) Gene targeting in astrocytes: targeting was performed in the same manner as MAFs/ MEFs. In the upper left of each graph is a representative flow cytometry plot after targeting in which GFP fluorescence is measured in the y-axis and background orange fluorescence along the x-axis. The number in the left corner of the flow plot is the percentage of GFP(+) cells. The corrected ROSA-3T3 cells show much higher GFP fluorescence than the primary cells demonstrating that while the ROSA26 locus is ubiquitously expressed, expression levels from the locus vary significantly depending on the cell type. Data are presented as mean ± SEM (*, $P<0.05$).

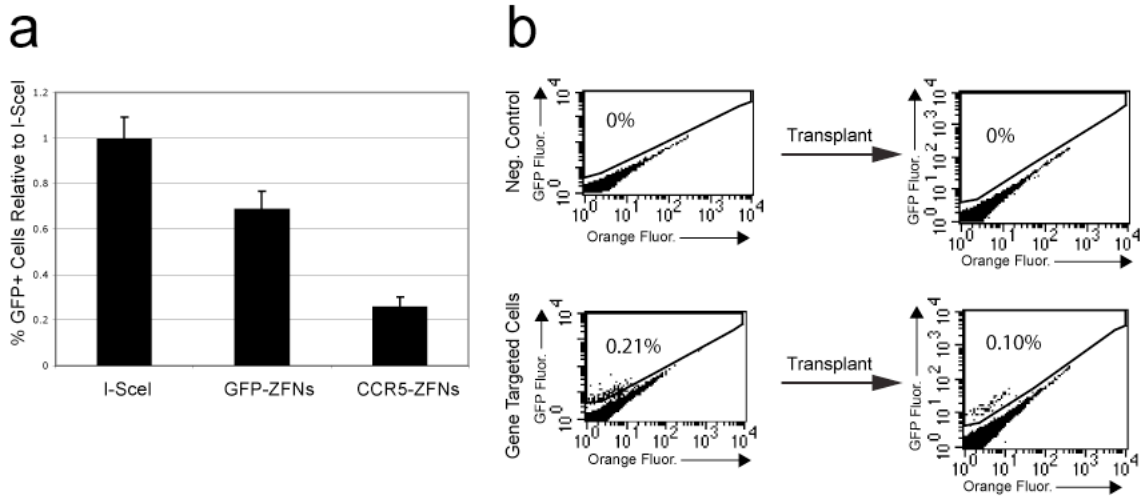


Figure 2.5: Toxicity assay and transplantation of gene targeted adult fibroblasts. (a) Toxicity of GFP-ZFNs in adult fibroblasts measured using a fluorescence reporter assay. Cells were transfected with a tdTomato reporter along with an I-SceI or GFP-ZFN expression plasmids. ZFN toxicity in cells is reported as fluorescence lost (due to cytotoxic effect) compared to cells transfected with the I-SceI expression plasmid. (b) Fibroblasts underwent gene correction by nucleofection of 2 ug of each ZFN expression plasmid and 10 ug of donor plasmid. Gene targeting was measured by flow cytometry immediately before transplantation 6 days after nucleofectionon. Fibroblasts were then injected subcutaneously in a Matrigel matrix. Two weeks after transplantion, the matrigel plug and surrounding skin was excised, cells dissociated, and plated. This was to allow for enrichment from contaminating recipient cells. On day 6, cells were harvest and analyzed using flow cytometry.

Discussion

ZFNs have now been used in a wide variety of situations to create precise genome modifications (Carroll, 2008). The precise genome modifications caused by ZFNs fall into two general classes. The first is to create small insertions/ deletions at a specific locus by the mutagenic repair of a ZFN induced double-strand break. While the exact mutation cannot be controlled using this strategy, the precise location of the mutation is controlled by the specificity of the ZFN induced double-strand break. This strategy is being increasingly used to create knockout cell lines and organisms in which the ability to efficiently create such knockouts was not previously available (Bibikova et al., 2002; Doyon et al., 2008; Geurts et al., 2009; Lloyd et al., 2005; Meng et al., 2008; Morton et al., 2006; Santiago et al., 2008; Shukla et al., 2009; Townsend et al., 2009). In addition, this strategy has been used to create targeted mutations in the CCR5 gene in human T-cells, thereby creating a population of T-cells that are resistant to HIV infection (Perez et al., 2008)—a strategy that has now entered a Phase I clinical trial (Clinicaltrials.gov identifier NCT00842634). Theoretically this approach could also be used to treat dominant genetic diseases by selectively creating deletions in the dominantly acting allele. The second way to create precise genome modifications (gene targeting) is to use the ZFNs to create a gene specific double-strand break and then have the cell repair that break by homologous recombination using an introduced donor sequence as the template for repair. Using this strategy both the specific site of the genome change and the specific sequence change can be controlled. In addition to the double-precision of gene

targeting, it also has the advantage that one can precisely create both small (single nucleotide) or large (the insertion of full a transgene cassette) changes in the genome (Moehle et al., 2007; Porteus, 2006) and reviewed in (Porteus and Carroll, 2005). In addition to being able to create inactivating mutations as the first strategy does, this strategy also allows the controlled integration of a transgene or the direct correction of a disease causing mutation and could theoretically be used to treat recessive genetic diseases. ZFN mediated gene targeting has now been used in a wide variety of different cell types at a large number of different loci. These include the modification of the IL2RG gene in various cell lines (Urnov et al., 2005), the modification of the CCR5 gene in a variety of cell lines including primary human embryonic stem cells and hematopoietic stem cells (Perez et al., 2008), the correction of GFP reporter genes and the modification of the PIG-A gene in human embryonic stem (hES) and human induced pluripotent stem (iPS) cells (Zou et al., 2009), and the modification of the OCT4 (POU5F1) and AAVS1 locus in hES and iPS cells (Hockemeyer et al., 2009).

As part of the effort to translate the use of ZFN mediated gene targeting to clinical use, we have generated a mouse model of a generic recessive disease by knocking-in a mutated GFP gene into the murine ROSA26 locus. Our goal is to generate a model in which one can either test the ability to use ZFN mediated gene targeting to correct mutations directly *in vivo* (similar to work performed by Miller *et al* to assess AAV gene targeting (Miller et al., 2006)), or to mimic a paradigm in which patient cells are purified and then precisely modified *ex vivo*

before transplanting the modified cells back into the patient. Here we report a first step in establishing the *ex vivo* cell modification paradigm by demonstrating that primary embryonic fibroblasts, primary adult fibroblasts, and primary astrocytes can all be isolated from the transgenic mouse line and the mutation in the GFP gene efficiently corrected using ZFN mediated gene targeting *ex vivo*. The rates of correction in primary fibroblasts (~2%) are of a sufficient magnitude to suggest clinical utility in the appropriate situation. Moreover, the gene corrected adult fibroblasts could be successfully transplanted back into an immunocompetent mouse where they both survived and retained their corrected phenotype. Unlike previous studies where modified fibroblasts were transplanted back into a recipient mouse, by using gene correction the corrected gene is being driven by its own promoter and thus will not undergo silencing as occurs with transgenes that are driven by viral elements.

The use of gene targeting by homologous recombination without the use of nucleases in murine ES cells is a well-established procedure. Here we demonstrate that ZFNs can stimulate gene targeting in >1% of cells without selection which is an order of magnitude or more than the rate of targeting in murine ES cells using I-SceI(Donoho et al., 1998) or in human ES or iPS cells(Hockemeyer et al., 2009; Zou et al., 2009). Thus this efficiency is high enough that one might be able to use ZFNs to create ES cells with extremely precise genetic modifications without using selectable markers and/or the Cre/Lox system.

This work also highlights an alternative strategy to creating gene corrected iPS cells. In prior work, investigators have first converted fibroblasts into iPS cells and then used gene targeting by homologous recombination in the iPS cells to correct a mutation (Hanna et al., 2007). The relatively high rates of gene correction stimulated by zinc finger nucleases in primary fibroblasts that we demonstrate here suggests that one could first correct disease causing mutations in patient fibroblasts, and subsequently convert those gene corrected fibroblasts into iPS cells. The advantages and disadvantages of whether to correct disease causing mutations at the primary cell stage or after conversion to iPS cells will likely be assessed in future studies.

This work shows that in primary cells, just as in cell lines, that the best gene targeting rates are obtained when the optimal mixture of ZFN expression plasmids and donor plasmid are introduced. Our finding that optimal targeting rates are achieved when the donor plasmid is at least 10-fold more abundant than the ZFN expression plasmid suggests that keeping the two elements separate so that each can be introduced at their optimal amounts will be important in achieving optimal targeting rates in different cell types.

In summary, we have developed a mouse model of a generic recessive disease in which ZFN mediated gene targeting can be studied in any cell of the mouse. A particular advantage of this model is that correctly targeted cells can easily be quantified and isolated for subsequent experiments using flow cytometry. In this paper, we demonstrated gene targeting and quantified the rates within several cell types from this mouse model and that targeted adult

fibroblasts, a potentially clinically relevant cell type, can be transplanted back into a mouse. Our current work focuses on isolating stem cells (such as hematopoietic stem cells, mesenchymal stem cells, muscle progenitor cells, and adipocyte precursor cells) from the mouse followed by performing gene targeting and subsequently transplanting the targeted stem cells back into a mouse, thus extending our current results of correcting and then transplanting a somatic cell. This study represents an important step in developing a paradigm for gene correction based gene therapy in an animal model prior to human clinical trials.

Materials and Methods

Generation of ROSA26-GFP targeting construct*

We constructed the ROSA26-GFP* targeting construct by first destroying the XhoI and XbaI sites present in the insert of the published GFP* reporter gene. This was done by digesting GFP* with XhoI followed by blunting of ends with Klenow. Next the XbaI site was destroyed in the same manner creating GFP*-B. We then used PCR to fuse a XhoI-XbaI-ClaI linker to GFP* using the follower primers:

GFP*-XhoI-XbaI-ClaI-Forward:

5'GTCTCGAGTCTAGAATCGATATGGTGAGCAAGGGCGAGG-3'

and GFP*-XhoI-Reverse:

5'-GACTCGAGTTACTTGTACAGCTCGTCCATGCCG- 3'.

The vector GFP*-B-Pgk-Neo was created by ligating GFP*-B into the XhoI site of Pgk-Neo-Pgk-PA, giving GFP*-B-Neo. Next a splice acceptor with ClaI ends was generated by PCR of pSAbGeo with primers:

SAClaI-Forward: 5'- GGATCGATATCTGTAGGGCGCAGTAGTCCAG-3'

and SAClaI-Reverse: 5'-GTATCGATACCGTCGATCCCCACTG-3'.

We digested the PCR product with ClaI and ligated it into the ClaI site of GFP*-B-Neo, generating SA-GFP*-B-Neo. Finally this vector was digested with XbaI, and ligated into the XbaI site of the ROSA26-1 targeting vector. All restriction enzymes were ordered from New England Biolabs Inc. The vectors pSAbGeo and pROSA26-1 were generous gifts from Dr. Philippe Soriano (Mount Sinai, New York, NY).

Generation of Reporter Mice

All experiments involving mice were approved by the IACUC at the University of Texas Southwestern Medical Center. Targeting of the GFP* reporter construct to the ROSA26 locus was performed by the UT Southwestern Transgenic core facility. Briefly, 50ug of targeting vector was linearized with KpnI and electroporated into $1x10^7$ 129/SvEvTac (SM-1) ES cells. Selection for targeted clones was performed using G418 at 250ug/ml. Resistant clones were picked and screened using southern analysis (**Figure 2.1**). Correctly targeted ES cells were injected into pseudo pregnant C57/Bl6 females. Chimeric offspring were bred to identify mice transmitting the ROSA26-GFP* allele to their progeny.

Gene targeting in ES cells

ES cells were cultured in ESLX media consisting of ES-DMEM, 20% ES Qualified FBS, 1X non-essential amino acids, 1X nucleosides, 1000 U/ml ESGRO LIF all purchased from Chemicon (Chemicon, Billerica, MA), 2mM L-Glutamine, 1X Pen/ Strep (Invitrogen, Carlsbad, CA), and 0.12mg/ ml sodium

pyruvate and 0.1mM BME (Sigma-Aldrich, St. Louis, MO). Before reaching confluency, cells were harvested and plated 100,000 cells per well in a 24 well gelatinized plate with and without 100nM vinblastine. We then transfected ES cells using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) with the indicated amount of donor and ZFN plasmids. Two sets of ZFNs designed to target GFP were compared. Both sets bind to the same recognition site; however, their amino acid sequences differ. Fifteen hours post-transfection media was changed to ESLX without vinblastine. Gene targeting events were analyzed four days post transfection using a FACScalibur (Becton Dickinson).

Gene targeting and teratoma formation in ES cells

Gene targeting experiments in ES cells were performed by nucleofection using the Mouse ES Cell kit (Lonza, Switzerland, Cat. VPH-1001). Briefly $1.5x10^6$ cells were nucleofected with 1.8 ug donor substrate and 1.6 ug of each GFP-ZFN using program A-30, giving an initial targeting rate of 0.07%. Cells underwent two rounds of sorting for targeted GFP^+ cells. For teratoma analysis, nude mice were sublethally irradiated with 350 rads using a cesium-137 source. The following day mice received subcutaneous injects in the hind flank with $1.3x10^6$ ES cells suspended in PBS. Teratomas were harvested 16-20 days after injections, photographed for GFP and Cy3 fluorescence using a Zeiss Stemi-11 Stereoscope equipped with an epifluorescence illuminator and Optronics Macrofire CCD camera. Teratomas were subsequently fixed in 4% paraformaldehyde for 48hrs and paraffin processed for histopathologic analysis. Resulting H&Es stains were photographed on a Leica DM2000 upright

microscope with standard bright-field optics and Optronics Microfire CCD camera.

Gene targeting in primary cells

MEFs: MEFs were isolated from E13.5 embryos using the WiCell protocol (WiCell Research Institute). Cells were cultured in DMEM, 10% FBS, Pen/Strep, and L-Glut. Before senescence occurred, cells were harvested and nucleofected in triplicate using the Basic Fibroblast kit (Lonza, Switzerland, Cat. VPI-1002) with program U-23. Each nucleofection consisted of $4x10^5$ cells plus 2 ug of each GFP-ZFN and 1, 5, 10 or 20 ug of donor plasmid. We analyzed gene targeting events on day 4 post-transfection.

MAFs: MAFs were isolated from the ears of 3-6 month old mice and cultured in DMEM, 20% FBS, Pen/Strep, L-Glut, Fungizone and 1X non-essential amino acids. $5x10^5$ cells per sample were nucleofected per sample using the Basic Fibroblast kit (Lonza, Switzerland, Cat. VPI-1002) with program U-23 and analysis performed on day 4 post-transfection.

Astrocytes: Astrocytes were isolated from newborn mice and cultured in the same media used for MAFs. Nucleofections were performed on $5x10^5$ cells per sample using the Mouse Astrocyte kit (Lonza, Switzerland, Cat. VPG-1006) with program T-20.

Transplantation of gene targeted primary adult fibroblasts and cell survival assay

Toxicity of GFP-ZFNs was measured as previously described (Pruett-Miller et al., 2008). Briefly, $1x10^6$ primary fibroblasts were nucleofected with 4 ug of a tdTomato expression plasmid along with 4 ug of an I-SceI expression plasmid or

2ug of each GFP-ZFN, and replated in a 12 well plate. On day 2, a portion of the cells was analyzed using flow cytometry and the remainder of cells re-plated in a 12 well plate. On day 6, cells were harvested and analyzed for tdTomato expression using flow cytometry. Toxicity compared to I-SceI was determined by first calculating the change in GFP expression from day 6 to day 2 of the I-SceI transfected cells and then the GFP-ZFN transfected cells ($DGFP_{nuclease}$ = $\%GFP^{+}_{Day6}$ / $\%GFP^{+}_{Day2}$). Toxicity compared to I-SceI was then calculated as Toxicicty = $DGFP_{ZFNs}$ / $DGFP_{I\text{-}SceI}$. Analysis of toxicity for ZFNs targeting the CCR5 gene "CCR5 ZFNs" was performed in parallel using 2 ug of each CCR5 ZFN.

For transplantation experiments, fibroblasts underwent gene targeting by nucleofection as described above, using 2 ug of each ZFN expression plasmid and 10 ug of donor plasmid. Cells were analyzed by flow cytometry 6 days after nucleofection, immediately before transplantation. $9.3x10^5$ fibroblasts, of which 0.21% were GFP+, were injected subcutaneously in a Matrigel (BD Biosciences, San Jose, CA) matrix in the back of a wild type, immunocompetent mouse not containing our *ROSA26-GFP** transgene. The Matrigel plug and surrounding tissue were excised 2 weeks later. Cells were dissociated by incubation in collagenase/displace (4mg/ml) (Roche Diagnostics, Indianapolis, IN) for 1 hour at 37 degrees. One ml MAF media was then added and cells incubated overnight at 37 degrees. The next morning, cells were triturated, filtered with a 70uM cell strainer (BD Biosciences, San Jose, CA) and plated in 24 well plates. On day 6 cells were harvest and analyzed using flow cytometry.

Chapter III- ZINC FINGER NUCLEASE AND TAL EFFECTOR NUCLEASE MEDIATED SAFE HARBOR GENE ADDITION WITHOUT SAFE HARBOR GENE DISRUPTION IN MOUSE PRIMARY FIBROBLASTS

Abstract

Nuclease-mediated safe harbor gene addition strategies are promising as next generation gene therapy technology. Heretofore, “safe harbors” have been defined as loci that can be disrupted without physiologic consequence and which carry no oncogenic potential when disrupted. In this study, we propose a modification of homologous recombination-mediated safe harbor targeting which does not require disruption of the endogenous gene product. In short, DNA which results in the same amino acid sequence as the target locus, but is non-homologous to the target locus by modification of the wobble position within multiple codons, can be targeted in-frame to result in no protein deficiency from the safe harbor. To demonstrate the feasibility of this strategy, we used a GFP reporter assay which we have previously described (Connelly et al., 2010). In this assay, a GFP gene which carries an insertional mutation that renders the protein non-functional was knocked-into the mouse ROSA26 locus. For gene addition, a donor plasmid containing arms of homology to the GFP gene surround the desired “gene of interest” to be added to the genome. Importantly, 5’ to the “gene of interest”, we include a non-homologous sequence of DNA which codes for the completion of the C-terminus of GFP. Either zinc finger nucleases (ZFNs) or TAL effector nucleases (TALENs) specific for the GFP locus were co-transfected with this donor resulting in a gene addition event that restores GFP expression. We designed multiple donor plasmids with these GFP elements and included as our

“gene of interest” the Ubc promoter driving human growth hormone (hGH) cDNA, an array of multiple hGH genes linked by 2A peptides, human platelet derived growth factor (PDGF-B) or ΔNGFR, a surface selectable marker that was targeted in-frame with GFP by a 2A peptide without the use of an exogenous promoter. Targeting frequencies ranged from 0.04-1.9% in primary fibroblasts depending on the donor construct or nucleases used, and targeting events were selectable by sorting for GFP or the surface marker ΔNGFR. Transgene expression for hGH was quantitated by ELISA to be 6.5-19.3 ng per million GFP positive cells per 24 hours. PDGF expressed reached 25.8 ng per million GFP positive cells per 24 hours. We directly compared the ability of ZFNs or TALENs to stimulate targeting at the same site, and found that TALENs markedly improved the efficiency of targeting over ZFNs with a simultaneous decrease in associated cellular toxicity. We also observed that targeting multiple copies of a transgene linked with the 2A peptide increases expression after targeting and that targeted fibroblasts could be re-introduced subcutaneously into either a sibling recipient mouse or an immunosuppressed unrelated strain of mouse for at least 30 days. The impact of the targeting system described here is two-fold. First, gene addition in a murine safe harbor locus can now be studied with virtually any gene of interest in any primary cell type with an easily assayable and quantifiable GFP reporter. Importantly, the restoration of GFP is specific for targeting events only. This is not the case with any other reporter for gene addition described to date. Secondly, the system described here provides proof

of principle for an evolution in safe harbor gene addition technology where the disruption of the target locus gene product is no longer required.

Introduction

Gene therapy, or the replacement of a defective gene with a functional copy, provides a definitive cure for many genetic diseases. Recently, the precision of this strategy has improved tremendously where the placement of the functional gene copy can now be guided by site-specific nuclease mediated homologous recombination. Homologous recombination is one pathway by which cells repair double strand breaks (DSBs), in doing so, the use of a homologous template is required. By providing a homologous template (an exogenous plasmid) with a desired sequence change, a site specific modification can be affected within the genome. The induction of a site-specific DSB at this site significantly increases the efficiency of this process (Kim and Chandrasegaran, 1994; Rouet et al., 1994a).

There are currently two major approaches for the induction of site-specific DSBs. The first, and more broadly studied to date, is the zinc-finger nuclease (ZFN) class of proteins. These proteins are non-natural, chimeric fusions of a zinc-finger DNA binding domain and a non-specific nuclease domain from the FokI endonuclease (Kim et al., 1996). Zinc-finger DNA binding domains consist of multiple fingers, each of which binds 3 nucleotides. The DNA binding specificity stems from the nucleotide recognition sequence from the combination of fingers, and also from the fact that the nuclease domain must dimerize to create a DSB. Thus, a pair of 3-finger ZFNs would have a 9 base-pair recognition

site on each side, for a total of an 18 base-pair recognition sequence, which is theoretically unique in the human genome. The second, and more recently discovered class of proteins are the TAL effector nucleases (TALENs). TALENs are also chimeric nucleases that possess the FokI nuclease domain, but the DNA binding domain is derived from the Transcription activator-like (TAL) effectors, which are virulence factors in the phytopathogenic bacteria *Xanthomonas.* TAL effectors bind DNA through a series of 34 amino acid repeats that differ by only two residues. These two residues are responsible for nucleotide binding specificity, and proteins with 11 to 17 repeats can be designed modularly to specifically bind a DNA sequence of interest (Boch et al., 2009; Moscou and Bogdanove, 2009). Both ZFNs and TALENs can efficiently bind and cleave DNA to create site-specific double strand breaks.

To circumvent the problems of insertional oncogenesis associated with earlier forms of gene therapy, both ZFNs and TALENs have been used for a newer strategy termed “safe harbor” gene addition. Heretofore, “safe harbors” have been defined as loci that can be disrupted without physiologic consequence and which carry no oncogenic potential when disrupted. A few robust examples include the human AAVS1 locus, for which ZFNs have been designed for safe harbor gene addition (DeKelver et al., 2010; Hockemeyer et al., 2009), or the mouse ROSA26 locus which is frequently used for mouse knock-in technology (Zambrowicz et al., 1997). However, this paradigm could be improved by a few modifications. First, little evidence has been provided that insertion of transgenes at these loci are truly “safe” in human patients, and this idea should be

approached with caution. For example, AAVS1 is located within the *PPP1R12C* gene on human chromosome 19. *PPP1R12C* encodes for a phosphatase downstream of the AMPKinase pathway involved with proper completion of mitosis (Banko et al., 2011). A better alternative to disrupting a locus such as this might be to develop a strategy where locus disruption is not a side-effect of gene addition. Secondly, some safe harbors may not provide a) accessible target sites in certain cell types because of chromatin status or b) sufficient protein expression in certain cell types for the same reason, which may limit therapeutic efficacy (van Rensburg et al., 2012). Thus, it might be better to develop a strategy where any locus, and at any site within the locus, can be used as a safe site for integration. This provides inherent flexibility in selecting and targeting the most robustly expressed locus for each cell type. Lastly, in the laboratory setting, a specific reporter for gene addition events is currently missing in the literature. Gene addition strategies described to date typically use donor vectors where the transgene has independent expression capabilities. Thus, any integration event (site-specific or random) expresses the transgene, and this is not initially distinguishable without later confirming by Southern blot or PCR based strategies, typically of clonal populations.

To address these issues, we sought to develop a new paradigm for safe-harbor gene addition. Here, we describe a proof of principle strategy where safe-harbor gene disruption is not a consequence of transgene addition, and so virtually any locus in the genome could be used as a safe harbor. In demonstrating this proof of principle, we have also simultaneously developed a

mouse reporter which allows for a gene addition-specific fluorescent read out. With these capabilities, we can rapidly quantitate gene addition frequencies in primary cells derived from a mouse without the need for single cell clone analysis or other Southern blot and PCR based techniques. Thus, the impact of the targeting system described here is two-fold. First, gene addition in a murine safe harbor locus can now be studied with virtually any gene of interest in any primary cell type with an easily assayable and quantifiable GFP reporter. Importantly, the restoration of GFP is specific for targeting events only and so bulk populations can be accurately studied, as opposed to single cell clonal analysis. This is not the case with any other reporter for gene addition described to date. Secondly, the system described here provides proof of principle for an evolution in safe harbor gene addition technology where the disruption of the target locus gene product is no longer required.

We demonstrate that primary fibroblasts derived from a GFP mouse model of a “safe harbor” locus can undergo gene addition of several different transgenes while maintaining expression of the safe harbor gene product (GFP). This is accomplished by including 5’ to the transgene, a non-homologous sequence of DNA which codes for the completion of the C-terminus of GFP within the safe harbor reporter. The transgenes targeted include the Ubc promoter driving human growth hormone (hGH) cDNA, an array of multiple hGH genes linked by 2A peptides, human platelet derived growth factor (PDGF-B), or ΔNGFR, a surface selectable marker that was targeted in-frame with GFP by a 2A peptide without the use of an exogenous promoter. After quantitation of

targeting by flow cytometry for GFP, we confirmed hGH and PDGF-BB expression by ELISA, and demonstrated that targeted cells could be re-introduced into either a sibling recipient mouse or an unrelated mouse strain under immunosuppressive conditions. We are currently investigating the therapeutic utility of transplanting PDGF-B targeted fibroblasts in a mouse wound healing model.

Results

Targeting Growth Hormone cDNA to a Safe Harbor Without Disruption

A disadvantage to current safe harbor gene addition strategies is that a safe harbor must be identified where targeted insertion and disruption of the locus results in no physiologic detriment. We sought to design a gene addition strategy which preserved the gene product of the safe harbor. For this purpose, we utilized a knock-in mouse model which we have previously described (Connelly et al., 2010), to serve as a reporter for gene addition events. Briefly, a mutated, non-fluorescent GFP gene was inserted in the mouse ROSA26 locus. We chose this model because restoration of the endogenous gene product (GFP) provides a reporter which is entirely specific for a gene addition event, an advantage which hasn't been described in the literature to date. Other safe harbor gene addition models transfect site-specific nucleases along with a donor plasmid containing a full-length transgene, typically with independent expression capabilities. When transfected, targeting can occur or random integration can occur. The efficiency of targeting determines the ratio of targeting to random integration. Because of the independent expression capabilities of most

reporters, random integration cannot be conveniently (by flow cytometry for example) distinguished from targeting events, and Southern blot or other methods are required to confirm targeting. In our model, only gene addition restores the expression of our reporter, and so is a more convenient system to study gene targeting events.

We designed a donor plasmid which contained a 5' region of homology to the target locus, followed by a non-homologous sequence capable of completing the C terminus of GFP, followed by a transgene, and lastly a 3' region of homology to the target locus (**Figure 3.1A**). Critically, the non-homologous sequence capable of completing the C terminus of GFP was modified from the wild-type GFP sequence because this prevents recognition as the 3' homology arm, which would exclude insertion of the transgene (**Figure 3.1B**). We generated two constructs, one with approximately every 3rd nucleotide modified (64.5% identity) and one with approximately every 6th nucleotide modified (83.5% identity). We found that both were sufficiently modified to not be recognized as homology (data not shown) and both were capable of restoring GFP expression. In 293T cells, the resultant GFP from both constructs was expressed well enough to be assayed by flow cytometry, however, in primary fibroblasts derived from our mouse model, the 64.5% construct was too dim to reliably distinguish GFP positive cells, possibly due to codon bias and poor protein expression (data not shown). As a result, we proceeded with constructs containing 83.5% GFP identity for the remaining experiments (**Figure 3.1B**). We observed that a construct consisting of two homology arms and our GFP 83.5% construct followed by the

Ubc promoter driving expression of human growth hormone (hGH) cDNA could be targeted in primary fibroblasts derived from our mouse model at a frequency of 0.27% (**Figure 3.1C left and middle panel**). These cells were sorted by FACS for GFP positive cells (**Figure 3.1C right panel**), and subjected to both Southern blot and PCR analysis to confirm targeting (**Figure 3.1D** and **3.1E**). Expression of hGH was confirmed by ELISA to be 15.2 ng per million GFP positive cells per 24 hours (**Figure 3.1F**). This data confirmed that we could generate a donor construct for gene addition which, through modification of the nucleotide sequence to prevent recognition as homology, could maintain (or in this case, restore) safe harbor gene expression after a gene addition event. Further, we established an easily assayable reporter specific for gene addition through GFP restoration which allows for rapid quantitation of gene addition frequencies.

Transplantation of Targeted, Growth Hormone Expressing Fibroblasts

Re-introduction of gene modified cells into a patient is a paramount goal for gene therapy by targeted gene addition. Thus, we asked whether the growth hormone secreting fibroblasts we generated in Figure 3.1 could be implanted into a recipient mouse. We injected fibroblasts subcutaneously in a Matrigel matrix and then harvested the Matrigel 10 or 30 days after transplantation. We observed that in a sibling mouse, 64% of the cells recovered at 10 days after transplantation were GFP positive, normalized to the pre-transplant population. However, after 30 days, only 5% remained. We hypothesized this decrease may be immune mediated, either because of the human growth hormone peptide, or because our knock-in mouse reporter strain is not an isogenic strain. To test this

theory, we also transplanted fibroblasts into an unrelated strain in the presence or absence of anti-mouse thymocyte serum (ATS) (injected intraperitoneally for 4 days prior to transplantation). As expected, we observed that in the absence of ATS 9.8% of transplanted cells remained after 10 days, and after 30 days, only 0.04%. However, after one treatment with ATS, 82% of cells remained after 10 days and 53% after 30 days (**Figure 3.2**). Thus, we were able to demonstrate successful re-introduction of gene modified cells that persisted in a recipient for at least 30 days.

Targeting Multiple cDNA Copies Increases Transgene Expression

Transgenes which are inserted by random integration sometimes have the advantage of high expression levels because the integration event can occur in an array of multiple copies. We asked whether targeting multiple copies of a transgene could result in a similar phenomenon. To generate a multicistronic vector, we turned to the T2A peptide derived from the insect *Thosea asigna* virus. The T2A peptide mediates a ribosomal skipping mechanism which results in linkage and expression of multiple open reading frames (Szymczak et al., 2004). We generated four constructs, each with increasing numbers of the human growth hormone cDNA termed hGH1x, hGH2x, hGH3x, hGH4x (**Figure 3.3A**). We found that we could successfully achieve gene addition of all four constructs at a frequency of 0.07%, 0.04%, 0.05%, 0.02% respectively (**Figure 3.3B**). Next, we sorted for GFP positive fibroblasts and analyzed hGH expression by ELISA. We found that constructs with more copies of hGH linked with the T2A expressed the trangene at higher levels, at 11 ng, 12 ng, 19 ng, and 16 ng per

million GFP positive cells per 24 hours, respectively (**Figure 3.3C**). Thus, targeting an array of transgenes linked with a 2A peptide can result in higher levels of transgene expression, however the gene addition frequency of the larger constructs was decreased.

TAL Effector Nucleases Outperform Zinc Finger Nucleases

The data described in Figure 3.1, 3.2 and 3.3 were achieved with zinc finger nuclease mediated gene addition. We wanted to ask if designing TAL effector nucleases to our locus could improve our targeting strategy. Using the donor construct described in Figure 3.1A, we determined that the targeting frequency for TALENs was fives times higher than for ZFNs (**Figure 3.4A**). We observed that this increase in targeting was maintained even through a titration of the ZFN pairs (**Figure 3.4B**) or TALEN pairs (**Figure 3.4C**). In fact, TALENs were able to stimulate substantial targeting at very low quantities (0.1 ug), data not previously observed with ZFNs. Next, we evaluated the toxicity profile for the ZFN and TALEN pairs. A tdTomato fluorescent plasmid with or without nucleases was transfected and tdTomato expression was analyzed by flow cytometry at day 2 or day 6 post transfection. Cell survival was calculated as a ratio of day 6:day 2 fluorescence normalized to samples transfected without nuclease. It was observed that cells transfected with the Ubc promoter driving one pair of ZFNs retained 96% cell survival, the CMV promoter driving a second pair of ZFNs had 83% cell survival, while the TALEN pair had 100% cell survival (**Figure 3.4D**). Thus, the TALEN pair demonstrated marked superiority compared to both pairs

of ZFNs in terms of both increased gene addition frequency, even at very low transfection quantities, and decreased associated cellular toxicity.

Next, we asked if a transgene could be inserted in-frame with the target locus, so that the use of an exogenous promoter would not be required. A donor construct was designed with a T2A peptide which would read in-frame with the restored GFP segment, followed by the ΔNGFR surface selectable marker (**Figure 3.4E**). It was demonstrated that TALENs could induce high levels of targeting with this donor at 1.9% percent compared with 0.07% for ZFNs and that the targeted fibroblasts could be rapidly and easily selected for with magnetic bead separation for ΔNGFR (**Figure 3.4F**).

Targeting Platelet Derived Growth Factor B in Fibroblasts

Fibroblasts play a critical role in the wound healing process by participating in granulation tissue formation and secreting extracellular matrix proteins. Platelet derived growth factor B (PDGF-B) is a key mediator in this process. After demonstrating site-specifc gene addition of secreted factors in transplantable fibroblasts, we hypothesized that by targeting PDGF-B in fibroblasts, it would be possible to augment the wound healing process. A donor plasmid was designed containing previously described targeting components and the CMV promoter driving expression of human PDGF-B truncated after the 211th amino acid (**Figure 3.5A**). The portion that is truncated encodes for a cell association domain, and so PDGF-B*211 is more actively secreted than full length PDGF-B (Eming et al., 1999; Ostman et al., 1991). Targeting of PDGF-B in mouse fibroblasts was confirmed at a frequency of 1.42% (**Figure 3.5B**). GFP

positive cells were sorted by FACs and PDGF-BB expression was quantitated by ELISA to be 25.8 ng per million GFP positive cells per 24 hours (**Figure 3.5C**). We are currently investigating the therapeutic utility of transplantation of these cells in a mouse wound healing model.

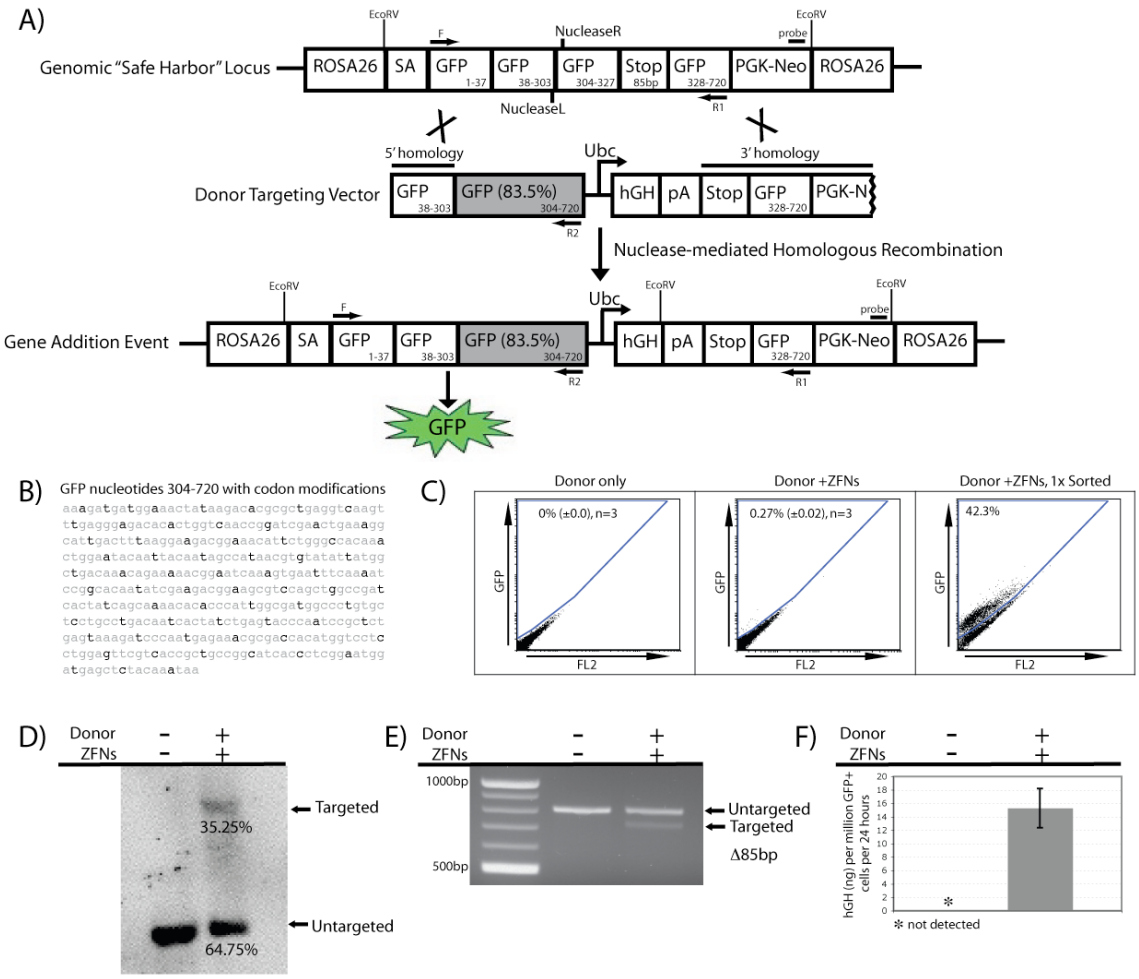


Figure 3.1 We designed a donor plasmid containing regions of homology to the genomic safe harbor locus. When nuclease expression plasmids were co-transfected with the donor, a site-specific gene addition event occurs(A). Critically, we included in our donor a region of DNA which can encode for the c-terminus of GFP, yet is non-homologous for wild-type GFP(B). This allows for the GFP expression to serve as a specific reporter for gene addition while simultaneously allowing transgene insertion. We demonstrated that co-transfection of all 3 plasmids resulted in GFP+ cells and that these could be sorted by flow cytometry (C). Sorted cells were analyzed by DIG-Southern with an EcoRV digest, and gene addition was confirmed (D). PCR of sorted cells also confirmed gene addition (E). ELISA was performed on the sorted population of cells and confirmed growth hormone expression (F). Error bars represent +/- 1 standard deviation.

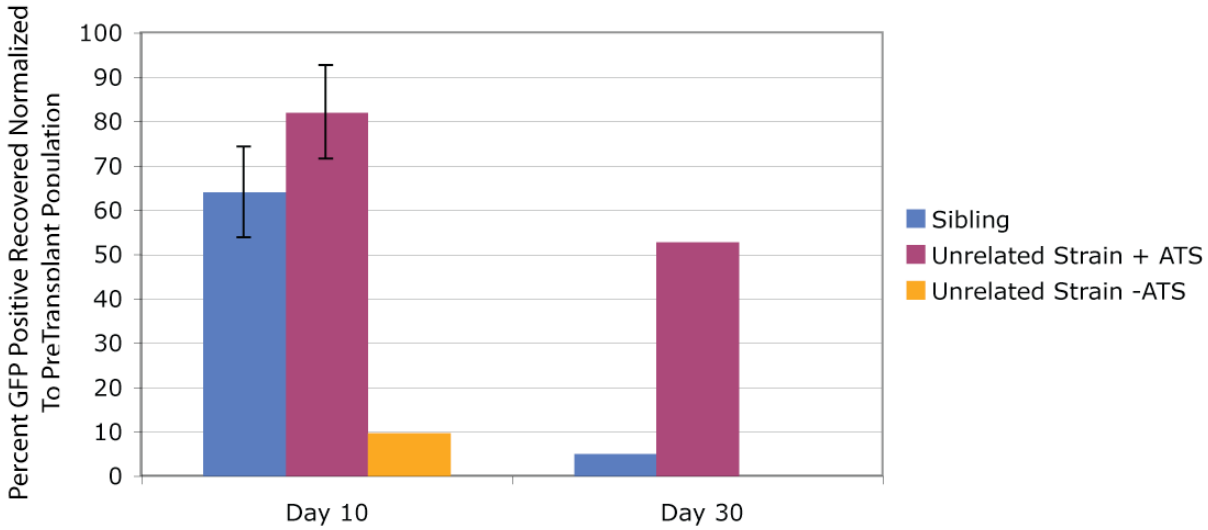


Figure 3.2 We transplanted fibroblasts targeted with the gene addition construct described in Figure 3.1 subcutaneously in Matrigel into either a sibling mouse (blue), an unrelated mouse pre-treated with anti-mouse thymocyte serum (ATS) for immunosuppresion (pink) or an unrelated mouse without ATS treatment (orange). We excised the matrigle plug and observed successful engraftment of the fibroblasts after 10 days in the sibling and unrelated +ATS cohorts. After 30 days however, only the unrelated +ATS cohort had substantial engraftment. Error bars represent +/- 1 standard deviation.

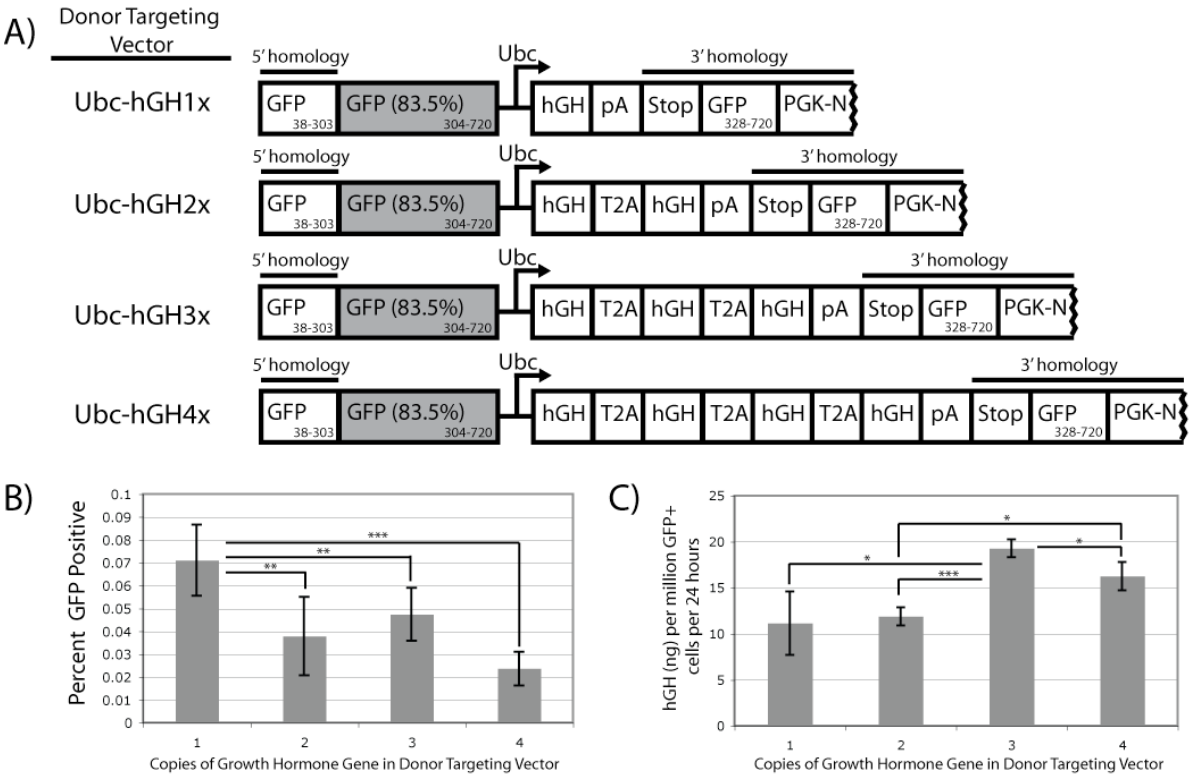


Figure 3.3 We designed four donor constructs, each containing an increasing number of growth hormone cDNA copies linked by a T2A peptide (A). As the size of the donor increased, the targeting efficiency decreased (B). Next, we sorted for GFP+ cells and normalized growth hormone expression (ELISA) to the GFP percentage. We found that increasing the copy number of cDNA can increase expression. However, Ubc-hGH4x did have lower expression than Ubc-hGH3x(C). Error bars represent +/- 1 standard deviation and p values were calculated with a Student's t-test assuming unequal variances. * $p \leq 0.05$ ** $p \leq 0.01$ *** $p \leq 0.001$

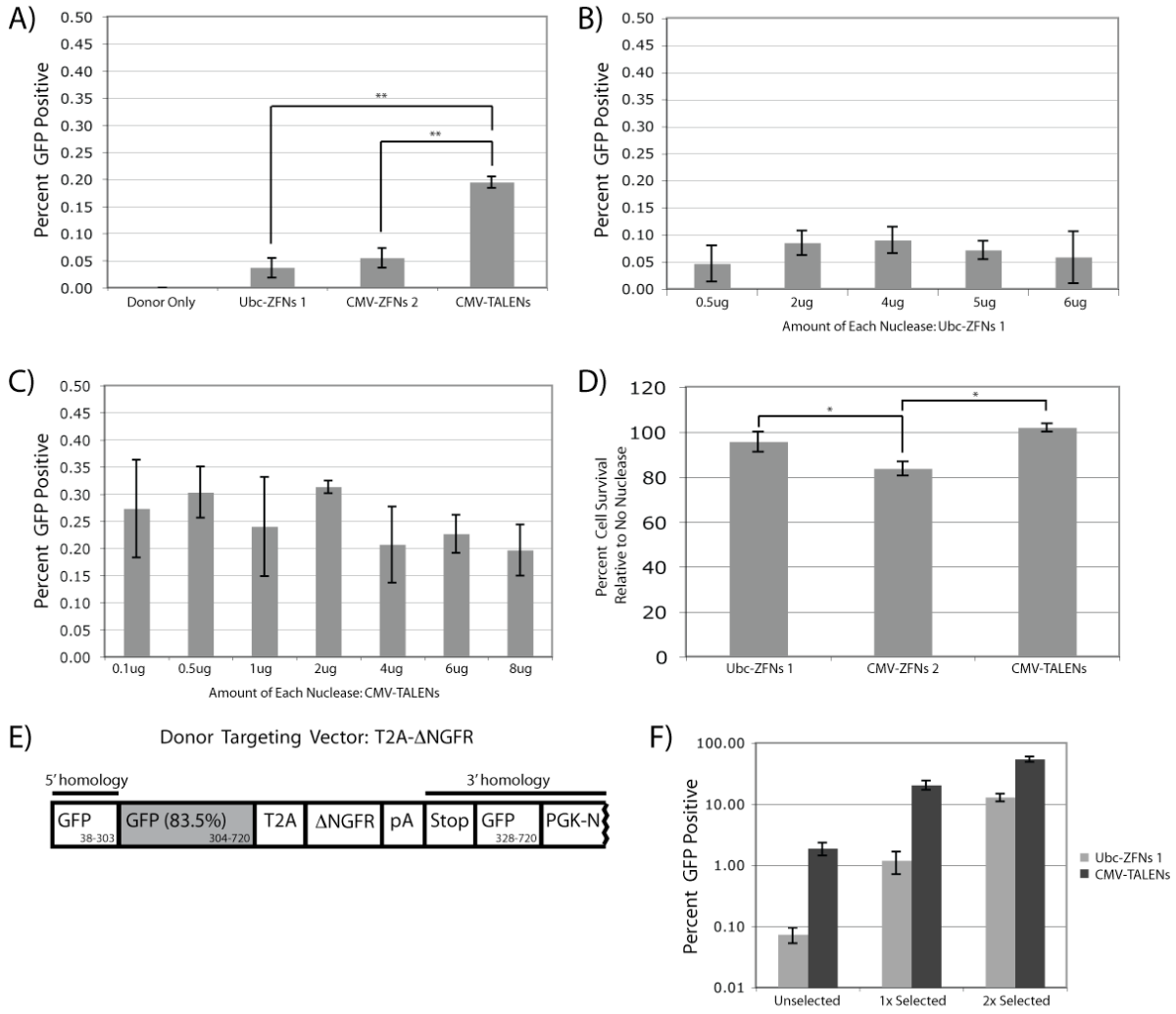


Figure 3.4 We compared the ability of TALENs to stimulate gene addition compared with the ZFNs used in Figure 3.1, 3.2 and 3.3. We found that TALENs outperformed ZFNs in terms of targeting efficiency (A) and also in terms of decreased cellular toxicity (D). We titrated the amount of ZFNs (B) and TALENs (C) and found that TALENs had higher levels of gene addition at all quantities. We then designed a donor construct to test the ability to target a transgene (truncated nerve growth factor receptor) in-frame with the target locus without the use of an exogenous promoter (E). We were able to successfully target and select for the transgene using magnetic beads (F). Error bars represent +/- 1 standard deviation and p values were calculated with a Student's t-test assuming unequal variances. * $p < 0.05$ ** $p \leq 0.001$

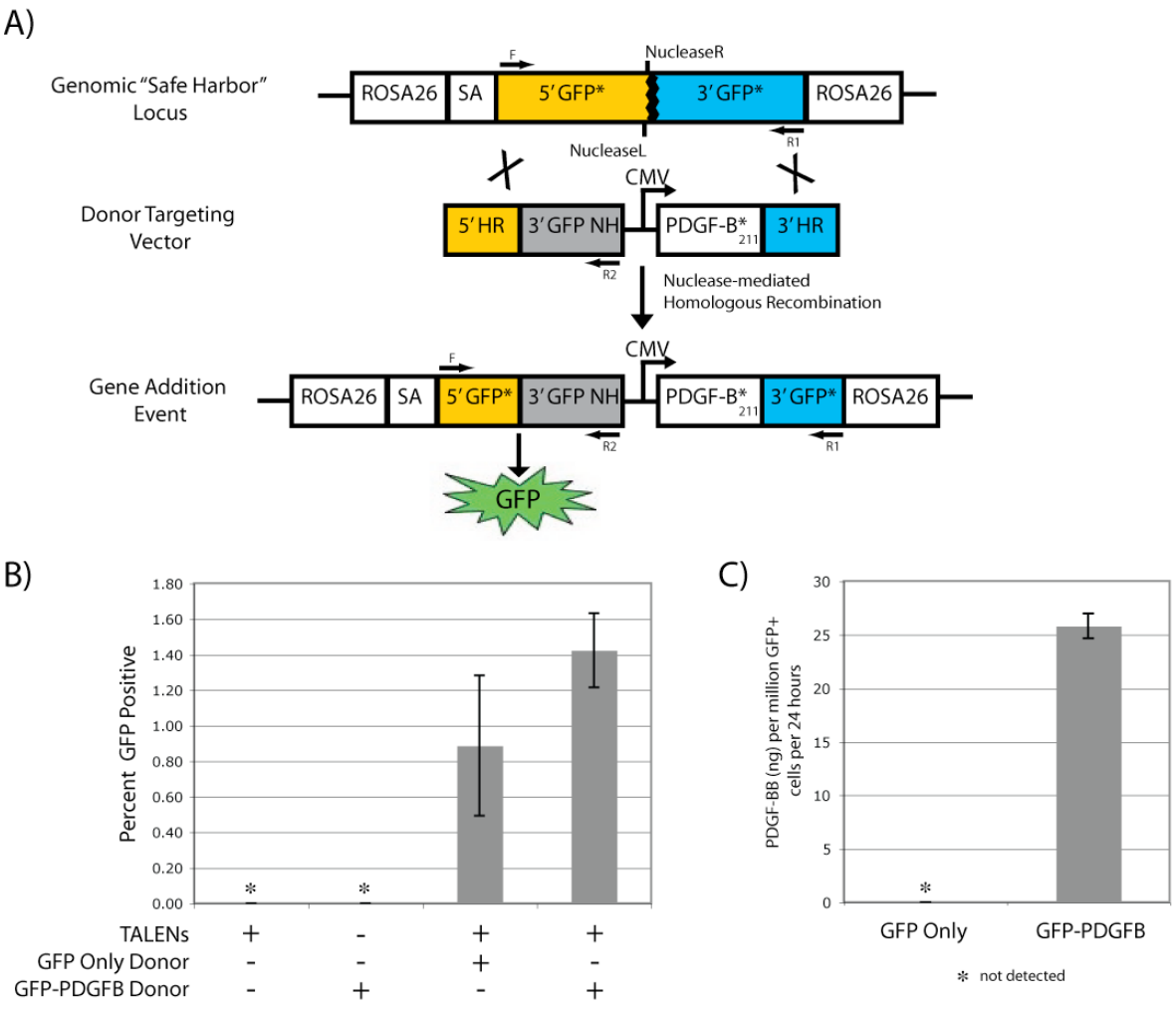


Figure 3.5 We designed a donor plasmid containing human platelet derived growth factor-B (truncated after amino acid 211 to allow for maximum secretion) driven by a CMV promoter (A). We were able to target the transgene, as measured by GFP fluorescence, in primary fibroblasts (B) and could confirm PDGF-BB homodimer expression by ELISA (C).). Error bars represent +/- 1 standard deviation.

Discussion

In Chapter II, we described the paradigm of ex vivo ZFN mediated site-specific gene targeting in mouse adult primary fibroblasts. In Chapter III, we have expanded on this paradigm to demonstrate that fibroblasts can undergo site specific gene addition events to secrete proteins. In the literature, gene addition in fibroblasts has been used for three categories of therapy. First, fibroblasts have been modified in diseases where the fibroblast is directly related to the pathology, such as epidermolysis bullosa (Titeux et al., 2010). Secondly, fibroblasts have been modified to serve as vehicles for systemic protein delivery by secreting ectopic proteins such as Factor VIII and IX for the treatment of Hemophilia A and B (Palmer et al., 1989; Qiu et al., 1996; Roth et al., 2001). Lastly, fibroblasts have been modified to secrete ectopic proteins, such as cytokines, to serve as enhancers of a local biologic process. This has been employed in models of wound healing, in models of tissue ischemia, and even in models of peripheral neuroregeneration through secretion of neurotrophic factors after injury (Breitbart et al., 2001; Mason et al., 2011; Zhang et al., 2011). Though many variations of fibroblast modification exist, none carry the precision of controlled gene addition through nuclease mediated site-specific integration.

Gene addition by conventional strategies can be therapeutically beneficial, but is not ideal for two reasons. First, it is now known that when viral vectors are used to deliver DNA in gene therapy models, infected cells are frequently eliminated by the immune system because of recognition of viral components (Manno et al., 2006; Raper et al., 2003). Secondly, because of the inherent lack

of precision with insertion of a transgene by random integration, gene addition by this strategy can have unpredictable outcomes. For example, in gene therapy clinical trials for severe combined immunodeficiency (SCID), some patients developed leukemia because of insertional activation of a nearby oncogene (Hacein-Bey-Abina et al., 2008). In the present study, we sought to combine the idea of using fibroblasts as a therapeutic protein delivery system with the technology for controlled, site-specific gene addition.

To overcome some of the obstacles associated with current safe harbor gene addition strategies, we sought to develop a strategy where locus disruption is not a consequence of gene addition. We accomplished this by designing a strategy for targeting a non-homologous sequence of DNA which encodes for and completes the C terminal amino acid sequence of the target locus. Altering the nucleotide sequence so that it is not identified as homology is critical because this prevents the homolgous recombination event from excluding the transgene. We demonstrated that altering 16.5% of the nucleotides, or roughly every 6th nucleotide at the wobble position was sufficient to prevent recognition as homology, yet was capable of substantiating expression from the safe harbor. This strategy provides an improvement to current safe harbor targeting strategies because we are no longer limited to loci which can be disrupted without physiologic consequence. In the past, these limitations may have led to selection of safe harbor which do not have the optimal capabilities for targeting or for therapeutic levels of expression. For this reason, we have demonstrated that we can target a locus (with the ΔNGFR transgene), in-frame with the endogenous

gene product and this allows for expression from the endogenous promoter without disrupting the endogenous locus. In some instances, such as within the hematopoietic system, harnessing the robust, tissue-specific expression of endogenous loci through targeted gene addition without safe harbor gene disruption may prove to be a powerful gene therapy strategy.

In conventional transgenesis by random integration, transgenes often integrate in multicopy tandem arrays. This can have consequences ranging from higher levels of transgene expression to silencing of the transgene because of cellular recognition of the array (Henikoff, 1998; Mutskov and Felsenfeld, 2004; Rosser and An, 2010). We hypothesized that targeting multiple copies of a cDNA in our safe harbor locus might result in higher levels of transgene expression, but at a certain copy number threshold, expression might decrease, possibly because of silencing or locus instability. We targeted the human growth hormone cDNA at either 1, 2, 3, or 4 copies and observed hGH expression. We observed that up to 3 copies provided increased expression. Interestingly, 4 copies had a lower level of expression than 3 which might suggest that targeted arrays can be silenced as well.

Most gene targeting literature to date is zinc finger nuclease (ZFN) based. However, the newly discovered TAL effector nucleases (TALENs) have shown promise as a next generation genome engineering tool. A major reason TALENs might be preferable to ZFNs is that they can be rapidly assembled to target virtually any locus with a modular assembly approach. ZFNs have been best assembled with careful selection strategies which are laborious and highly

technical. We sought to directly compare ZFNs and TALENs that target the GFP gene in our safe harbor locus. Our results were striking in the preference of TALENs over ZFNs in terms of targeting efficiency and also cellular toxicity. Our results, combined with the rapid, modular assembly design strategy for TALENs advocates for the continued study of TALENs as next generation site-specific nucleases.

In summary, we have developed a mouse model for site-specific gene addition events. This model simultaneously demonstrates that locus disruption is not a requirement for safe harbor gene addition while providing a gene-addition specific reporter that allows for rapid quantitation of gene addition frequencies. This work provides proof of principle that ectopic proteins can be secreted from fibroblasts and this provides a foundation for further studies. Our current work is focused on the development of gene modified fibroblasts as local delivery vehicles for cytokines such as PDGF-B for the augmentation of biologic processes, especially pertaining to wound healing pathology.

Materials and Methods

Generation of Gene Addition Constructs

The gene addition vector in Figure 3.1 was constructed by synthesizing the GFP nucleotides 38-720 (Genscript). Nucleotides 38-303 consist of the published nucleotides, while 304-720 are modified as described in Figure 3.1B. We then subcloned this construct into a pUB6 expression vector (Life Technologies, Grand Island, NY). Using the same plasmid from which the knock-in mouse was derived (Connelly et al., 2010), the 3' homology region was PCR

amplified with 5'AAGGACGACGGCAACTAC3' and 5'GACGTGCGCTTTTGAAGCGT3' and also subcloned in the pUB6 expression vector. Next, the hGH gene was PCR amplified (SC300088 Origene, Rockville, MD), and subcloned this into the vector along with a PolyA region. For the multicopy hGH constructs, two PCRs were performed for cloning- the first eliminated the stop codon within hGH and the second fused a Furin-SGSG-T2A sequence (5' CGCAAGCGCCGCAGCGGCAGCGGCG AGGGCCGCGGCAGCCTGCTGACCTGCGGCGACGTGGAGGAGAACCCCGG CCCC3') in front of hGH so that when cloned together, the two constructs would be in the same ORF. Serial cloning of these two constructs allowed for generation of multicopy donor vectors. For the ΔNGFR vector, the synthesized construct of GFP 38-720 described above was PCR amplified to eliminate the stop codon and the Furin-SGSG-T2A was fused by PCR to the ΔNGFR construct. Subcloning of these two in-frame resulted in the donor plasmid described in Figure 3.4. Lastly, the CMV-PDGF-B*211 donor was derived by cloning the GFP 38-720 construct and the 3' homology arm into a CMV expression vector. PDGF-B*211 was generated by PCR to truncate full-length PDGF-B (SC111665 Origene, Rockville, MD) after the 211th amino acid and then subcloned into the vector. All restriction enzymes were ordered from New England Biolabs Inc.

Generation of ZFNs and TALENs

The ZFNs described are the same two pairs previously published (Connelly et al., 2010). The TALENs were designed to recognize

TGCCCGAAGGCTACGT on the sense strand and TTGCCGTCGTCCTTGAAG on the anti-sense strand. The spacer between the TALENs is 18 basepairs. Within the TALEN repeats, NN recognizes G, HD recognizes C, NI recognizes A and NG recognizes T. These were cloned into a CMV expression vector along with the wild-type, codon optimized FokI nuclease domain and contain a 3x FLAG tag.

Primary Fibroblasts Culture, Transfection, and Gene Addition Analysis

Primary fibroblasts were isolated from the ears of 3-6 month old mice by 1 hour of digest in collagenase/dispase (4mg/ml) (Roche) and then 1 ml MAF media was then added and cells incubated overnight at 37 degrees. The next morning, cells were triturated, filtered with a 70uM cell strainer (BD Biosciences, San Jose, CA) and then cultured in DMEM, 16% FBS, Pen/Strep, L-Glut, Fungizone and 1X non-essential amino acids. Critically, all cultures were maintained in low oxygen conditions (5%) which drastically improves the survival of the cells and minimizes early senescence. $1x10^6$ cells per sample were nucleofected per sample using the Basic Fibroblast kit (Lonza, Switzerland, Cat. VPI-1002) with program U-23 and analyzed for GFP fluorescence by flow cytometry. Gene addition was confirmed by DIG-Southern (Roche) using an EcoRV digest and a probe designed against the PGK-Neo region at the 3' end of our knock-in locus (described in(Connelly et al., 2010)). PCR for gene addition was performed using the following 3 primers:

F: 5'ATGGTGAGCAAGGGCGAGGA3'

R1: 5'TTACTTGTACAGCTCGTCCATGCCG3'

R2: 5'TTATTTGTAGAGCTCATCCATTCCGAGGG3'

Growth hormone expression was quantitated by ELISA (ELH-GH-001 RayBiotech, Norcross, GA) by culturing 2 x 10^4 fibroblasts in 1 ml of media for 24 hours. PDGF-B expression was detected by the same strategy using an ELISA for the PDGF-BB homodimer (DBB00 R&D systems). ΔNGFR selection was performed by staining with magnetic bead conjugated antibodies (130-092-283 MACS Kit, Miltenyi Biotec). Cells were resuspended in 2.5ml MACs buffer, then incubated in an Easy Sep (18000 Stemcell Technologies) magnet for 10 minutes in a 5 ml tube. Liquid was briskly poured out of the tube, and then the resuspension and magnetic incubation was repeated. After 3-4 days selection was repeated.

Transplantation of Primary Fibroblasts

For transplantation experiments, fibroblasts underwent gene addition by nucleofection as described above. Cells were analyzed by flow cytometry prior to transplantation and were then injected subcutaneously in a Matrigel (BD Biosciences, San Jose, CA) matrix on the dorsum of either a sibling mouse or an anti-thymocyte serum (Fitzgerald industries) treated unrelated mouse. Mice who received ATS treatment were given 120 mg/kg intraperitoneally over the course of 4 days prior to transplantation for a total of 480 mg/kg. Successful lymphocyte knock down was confirmed with CBC analysis using a HemaVet system (Drew Scientific Waterbury, CT). Of note, we found that in our mice the dose needed for this lot of serum was higher than required by previous studies, suggesting that individual lots should be tested on a per strain basis for efficacy. After 10 or 30

days post-transplantation, the Matrigel plug was excised and then processed in the same manner as the initial fibroblast derivation above. Post-transplant fluorescence was quantitated by flow cytometry, and the percent survival was calculated as percent post-transplant GFP positive normalized to pre-transplant GFP positive.

Chapter IV- CONCLUSIONS AND FUTURE DIRECTIONS

Targeted genome engineering of patient derived cells holds promise for the treatment of diseases by providing more effective, more safe and more specific therapies than those which are currently available. However, at present, there are many rate-limiting steps involved in the clinical transition of genome engineering. First, with the partial exception of CD34+ human hematopoietic stem cells, isolation, culture and successful delivery of DNA to different cell types is still not routine, facile and standardized. Secondly, study of the interaction of targeting components (including nucleases and donor constructs) in primary cells and within a host is only beginning in its infancy. Many studies need to be conducted to determine the most efficacious means for targeting in primary cell types. This includes evaluating the most efficacious, and least toxic means for creating site-specific double strand breaks, as well as evaluating the most effective donor plasmids for targeting efficiency, transgene expression and selectability. Lastly, very little has been accomplished in evaluating the safety and efficacy of transplantation of gene modified cells in vivo, and this will be a critical component to translation of genome engineering into a therapeutic reality.

In theory, animal modeling should provide extensive means to study the aforementioned issues, and this is why we have developed the mouse model for genome engineering described in Chapters II and III.

Gene Correction in a Mouse Model of a Generic Recessive Genetic Disease

The ability to precisely correct endogenous genes holds tremendous potential for the clinical treatment of genetic disease. To aid in the advancement of this technology, we have generated a mouse model of a generic recessive genetic disease where correction of mutated GFP gene serves as a reporter that mimics correction of an endogenous locus. First, we demonstrated zinc finger nuclease mediated gene correction in the embryonic stem cells used to derive the mouse model and that these cells could maintain pluripotency after gene correction. Next, we isolated a variety of cells from the mouse including murine embryonic fibroblasts, primary astrocytes and primary adult fibroblasts and were able to demonstrate gene correction of the GFP gene in all cell types. Critically, the ability to achieve maximal gene correction was dependent upon optimizing the ratio of zinc finger nuclease to donor plasmid which fell into the range of 1:5 to 1:10. Lastly, we demonstrated that *ex vivo* gene corrected fibroblasts could be transplanted back into a mouse where they retained the corrected phenotype. We initially published that gene corrected cells could be maintained in a mouse for 2 weeks. In later, unpublished work, we demonstrated retention for 4 months. In summary, potentially therapeutically relevant rates of zinc finger nuclease mediated gene targeting can be achieved in a variety of primary cells, and this work continues to support the possibility of using zinc finger nuclease technology for therapeutic purposes.

The are two subsequent areas for study in the immediate future concerning this mouse model. The first is that the mouse should next be used to study gene correction events in perhaps more clinically relevant cell types,

especially stem cell populations including hematopoietic stem cells, mesenchymal stem cells, and muscle progenitor cells. A major aspect of this work is discovering the most optimal delivery conditions for targeting components into these cell types and this work is currently on going in the Porteus lab. The second area of study involves further investigating the use of fibroblasts as a clinically relevant cell type, and some of the work described in Chapter III seeks to address this issue.

Gene Addition in Primary Fibroblasts Without Safe Harbor Gene Disruption

Chapter III provides data in support of two key improvements in the field of targeted gene addition. The first is that we have developed a mouse model with a reporter that is entirely specific for gene addition events. This has not been described in the literature to date yet is a powerful tool for rapidly quantitating precise gene addition frequencies. The second is that we have demonstrated a strategy where site-specific gene addition can be achieved in a safe harbor locus without disrupting the gene product of the safe harbor. Heretofore, safe harbors have been limited to loci which could be disrupted without physiologic consequence. With this new strategy, virtually any locus in the genome could be used as an integration site for a transgene in a disease-specific and cell type-specific manner. Here, we have demonstrated with these two concepts the site-specific integration of several different transgenes including the Ubc promoter driving human growth hormone (hGH) cDNA, an array of multiple hGH genes linked by 2A peptides, human platelet derived growth factor (PDGF-B) or ΔNGFR, a surface selectable marker that was targeted in-frame with GFP by a

2A peptide without the use of an exogenous promoter, at frequencies ranging up to nearly 2% depending on the donor construct. We demonstrated that these gene modified fibroblasts could be enriched and then transplanted into either a sibling recipient or an unrelated strain of mouse under immunomodulatory conditions and that the fibroblasts persisted for at least 30 days. We also directly compared the ability of two classes of site-specific nucleases to stimulate gene addition at this locus. We found that the TALENs dramatically outperformed the ZFNs in terms of targeting efficiency and also associated cellular toxicity.

The value of the gene addition specific reporter is best demonstrated by the fact that it was developed out of necessity. Most safe-harbor gene addition events quantitated by reporters employ the use of a reporter with independent expression capabilities. Thus, the read-out for the reporter encompasses both site-specific integration events as well as random integration events. Depending on the ratio of these two, the ability to detect, quantitate, and select for low level targeting events can be difficult. In fact, the work described in Chapter III initially began by this same strategy. The experimental design for this project began with insertion of a transgene followed by a T2A peptide and a fluorescent reporter gene. We found that after selecting for the fluorescent reporter that we could not reproducibly demonstrate targeting in the bulk population by Southern blot analysis. Because we were able to achieve gene correction frequencies of nearly 2% as described in Chapter II, we made the assumption that frequencies would be similar for our gene addition experiments. As we later discovered after the development of the gene-addition specific reporter, this was not the case for

some of the larger donor plasmids. Gene addition frequency was a log lower than observed for gene correction. From this data, we concluded that when enriching for our fluorescent reporter, we were in fact enriching predominantly for random integration events and not targeting events. Thus, the development of the gene addition-specific reporter was absolutely essential to the progression of further work in this model, and truly highlights the value of being able to precisely quantitate low level gene addition events for optimization. An important lesson learned.

The initial hypothesis for this work was that we may be able to deliver therapeutic proteins systemically by using transplantation of gene modified fibroblasts as vehicles. Though we were able to demonstrate successful targeting, enrichment and transgene expression in these cell types, we were not able to demonstrate systemically detectable protein after transplantation into mice. We hypothesized that for these types of transgenes, fibroblasts may not be capable of sufficient expression or have sufficient access to the systemic vasculature to be able to affect a systemic change. Instead, we have turned to a more physiologically relevant model for clinical application of gene modified fibroblasts by applying fibroblasts in a setting for local protein delivery instead of systemic protein delivery. Fibroblasts are potent mediators in the wound healing process, and so we have created fibroblasts that also secrete a potent cytokine, PDGF-B, involved in would healing. We have demonstrated successful targeting and transgene expression of the PDGF-B gene and are currently investigating its ability to affect surrounding cells, in both an *in vitro* setting and also in an *in vivo*

setting by transplantation of fibroblasts into a mouse model for wound healing. The potential applications for gene modified fibroblasts in a local delivery setting are numerous and include not only different applications of the technology (such as wound healing, nerve regeneration, prevention of tissue ischemia, and tendon repair) but also the ability to modularly change the proteins to be delivered (in addition to PDGF-B, other stimulatory neoangiogenic, neurotrophic cytokines could be employed.)

In conclusion, the work described in this thesis seeks to thoroughly investigate the potential application for a mouse model for genome engineering. What began as a mouse model for gene correction could later be described as a specific mouse model for gene addition as well. Because gene correction and gene addition are the two most predominate strategies employed in the field of genome engineering, investigation of this model provides valuable insight into these techniques. I would like to thank Dr. Porteus and Dr. Huang for guiding me in the pursuit of this effort. I am grateful for the invaluable lessons I have learned, which I certainly will apply on a daily basis during the remainder of my scientific career.

Acknowledgements

We would like to thank Dr. Robert Hammer and Robin Nguyen at the UT Southwestern Transgenic Core for creation of the ROSA26 knock-in ES cells, derivation of the transgenic mouse and ES cell culture protocols. We also thank Dr. James Richardson and John Shelton at the JAR Molecular Pathology Core for help with photography, performing histological analysis of the teratomas and for preliminary work with our wound healing model. We thank Matthew Foglia for his careful reading and editing of the Chapter II manuscript. We thank the UT Southwestern Flow Cytometry Core, especially Christina Nguyen for cell sorting. Finally we thank Dr. Philippe Soriano for the pROSA26-1 and pSAbGeo vectors. The work in the Porteus lab is supported by the Amon Carter Foundation, the National Institutes of Health (NIH) grants PN2EY018244 (a Nanomedicine Development Center grant as part of the Director's initiative), R01 HL079295, and K08 H1070268, a career development award from the Burroughs Welcome Fund and state of Texas funding through the University of Texas Southwestern Medical Center. JB was supported by the UTSW Medical Scientist Training Program as well as the PhRMA Foundation Paul Calabresi Medical Student Fellowship.

REFERENCES

Abuljadayel, I.S., Quereshi, H., Ahsan, T., Rizvi, S., Ahmed, T., Khan, S.M., Akhtar, J., and Dhoot, G. (2006). Infusion of autologous retrodifferentiated stem cells into patients with beta-thalassemia. ScientificWorldJournal *6*, 1278-1297.

Aiuti, A., Brigida, I., Ferrua, F., Cappelli, B., Chiesa, R., Marktel, S., and Roncarolo, M.G. (2009a). Hematopoietic stem cell gene therapy for adenosine deaminase deficient-SCID. Immunol Res *44*, 150-159.

Aiuti, A., Cattaneo, F., Galimberti, S., Benninghoff, U., Cassani, B., Callegaro, L., Scaramuzza, S., Andolfi, G., Mirolo, M., Brigida, I.*, et al.* (2009b). Gene therapy for immunodeficiency due to adenosine deaminase deficiency. N Engl J Med *360*, 447-458.

Allers, K., Hutter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E., and Schneider, T. (2011). Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. Blood *117*, 2791-2799.

Alwin, S., Gere, M.B., Guhl, E., Effertz, K., Barbas, C.F., 3rd, Segal, D.J., Weitzman, M.D., and Cathomen, T. (2005). Custom zinc-finger nucleases for use in human cells. Mol Ther *12*, 610-617.

Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.C., Stricher, F., Petit, A.S., Patin, A., Guillier, S.*, et al.* (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. J Mol Biol *355*, 443-458.

Arnould, S., Delenda, C., Grizot, S., Desseaux, C., Paques, F., Silva, G.H., and Smith, J. (2011). The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. Protein Eng Des Sel *24*, 27-31.

Arnould, S., Perez, C., Cabaniols, J.P., Smith, J., Gouble, A., Grizot, S., Epinat, J.C., Duclert, A., Duchateau, P., and Paques, F. (2007). Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. J Mol Biol *371*, 49-65.

Ashworth, J., Havranek, J.J., Duarte, C.M., Sussman, D., Monnat, R.J., Jr., Stoddard, B.L., and Baker, D. (2006). Computational redesign of endonuclease DNA binding and cleavage specificity. Nature *441*, 656-659.

Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N.*, et al.* (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. N Engl J Med *358*, 2231-2239.

Bank, A., Dorazio, R., and Leboulch, P. (2005). A phase I/II clinical trial of beta-globin gene therapy for beta-thalassemia. Ann N Y Acad Sci *1054*, 308-316.

Banko, M.R., Allen, J.J., Schaffer, B.E., Wilker, E.W., Tsou, P., White, J.L., Villen, J., Wang, B., Kim, S.R., Sakamoto, K.*, et al.* (2011). Chemical genetic screen for AMPKalpha2 substrates uncovers a network of proteins involved in mitosis. Mol Cell *44*, 878-892.

Beerli, R.R., Dreier, B., and Barbas, C.F., 3rd (2000). Positive and negative regulation of endogenous genes by designed transcription factors. Proc Natl Acad Sci U S A *97*, 1495-1500.

Benabdallah, B.F., Allard, E., Yao, S., Friedman, G., Gregory, P.D., Eliopoulos, N., Fradette, J., Spees, J.L., Haddad, E., Holmes, M.C.*, et al.* (2010). Targeted gene addition to human mesenchymal stromal cells as a cell-based plasma-soluble protein delivery platform. Cytotherapy *12*, 394-399.

Beumer, K., Bhattacharyya, G., Bibikova, M., Trautman, J.K., and Carroll, D. (2006). Efficient gene targeting in Drosophila with zinc-finger nucleases. Genetics *172*, 2391-2403.

Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., and Carroll, D. (2008). Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. Proc Natl Acad Sci U S A *105*, 19821-19826.

Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D. (2003). Enhancing gene targeting with designed zinc finger nucleases. Science *300*, 764.

Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G., and Chandrasegaran, S. (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol *21*, 289-297.

Bibikova, M., Golic, M., Golic, K.G., and Carroll, D. (2002). Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics *161*, 1169-1175.

Bitinaite, J., Wah, D.A., Aggarwal, A.K., and Schildkraut, I. (1998). FokI dimerization is required for DNA cleavage. Proc Natl Acad Sci U S A *95*, 10570-10575.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science *326*, 1509-1512.

Bozas, A., Beumer, K.J., Trautman, J.K., and Carroll, D. (2009). Genetic analysis of zinc-finger nuclease-induced gene targeting in Drosophila. Genetics *182*, 641-651.

Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P.P., Diez, I.A., Dewey, R.A., Bohm, M., Nowrouzi, A., Ball, C.R., Glimm, H.*, et al.* (2010). Stem-cell gene therapy for the Wiskott-Aldrich syndrome. N Engl J Med *363*, 1918-1927.

Breitbart, A.S., Grande, D.A., Laser, J., Barcia, M., Porti, D., Malhotra, S., Kogon, A., Grant, R.T., and Mason, J.M. (2001). Treatment of ischemic wounds using cultured dermal fibroblasts transduced retrovirally with PDGF-B and VEGF121 genes. Ann Plast Surg *46*, 555-561; discussion 561-552.

Brenneman, M., Gimble, F.S., and Wilson, J.H. (1996). Stimulation of intrachromosomal homologous recombination in human cells by electroporation with site-specific endonucleases. Proc Natl Acad Sci U S A *93*, 3608-3612.

Brunet, E., Simsek, D., Tomishima, M., DeKelver, R., Choi, V.M., Gregory, P., Urnov, F., Weinstock, D.M., and Jasin, M. (2009). Chromosomal translocations induced at specified loci in human stem cells. Proc Natl Acad Sci U S A *106*, 10620-10625.

Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L., and Cui, X. (2010). Targeted genome modification in mice using zinc-finger nucleases. Genetics *186*, 451-459.

Carroll, D. (2008). Progress and prospects: zinc-finger nucleases as gene therapy agents. Gene Ther *15*, 1463-1468.

Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., Veres, G., Schmidt, M., Kutschera, I., Vidaud, M., Abel, U., Dal-Cortivo, L., Caccavelli, L.*, et al.* (2009). Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. Science *326*, 818-823.

Cathomen, T., and Joung, J.K. (2008). Zinc-finger nucleases: the next generation emerges. Mol Ther *16*, 1200-1207.

Cavazzana-Calvo, M., Hacein-Bey, S., de Saint Basile, G., Gross, F., Yvon, E., Nusbaum, P., Selz, F., Hue, C., Certain, S., Casanova, J.L.*, et al.* (2000). Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science *288*, 669-672.

Cavazzana-Calvo, M., Lagresle, C., Hacein-Bey-Abina, S., and Fischer, A. (2005). Gene therapy for severe combined immunodeficiency. Annu Rev Med *56*, 585-602.

Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res *39*, e82.

Choulika, A., Perrin, A., Dujon, B., and Nicolas, J.F. (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol *15*, 1968-1973.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. Genetics *186*, 757-761.

Connelly, J.P., Barker, J.C., Pruett-Miller, S., and Porteus, M.H. (2010). Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. Mol Ther *18*, 1103-1110.

Cui, X., Ji, D., Fisher, D.A., Wu, Y., Briner, D.M., and Weinstein, E.J. (2011). Targeted integration in rat and mouse embryos with zinc-finger nucleases. Nat Biotechnol *29*, 64-67.

DeKelver, R.C., Choi, V.M., Moehle, E.A., Paschon, D.E., Hockemeyer, D., Meijsing, S.H., Sancak, Y., Cui, X., Steine, E.J., Miller, J.C*., et al.* (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. Genome Res *20*, 1133-1142.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature *330*, 576-578.

Donoho, G., Jasin, M., and Berg, P. (1998). Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol *18*, 4070-4078.

Doyon, J.B., Zeitler, B., Cheng, J., Cheng, A.T., Cherone, J.M., Santiago, Y., Lee, A.H., Vo, T.D., Doyon, Y., Miller, J.C*., et al.* (2011a). Rapid and efficient clathrin-mediated endocytosis revealed in genome-edited mammalian cells. Nat Cell Biol *13*, 331-337.

Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J*., et al.* (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol *26*, 702-708.

Doyon, Y., Vo, T.D., Mendel, M.C., Greenberg, S.G., Wang, J., Xia, D.F., Miller, J.C., Urnov, F.D., Gregory, P.D., and Holmes, M.C. (2011b). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods *8*, 74-79.

Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D., and Barbas, C.F., 3rd (2001). Development of zinc finger domains for recognition of the 5'-ANN-3' family of

DNA sequences and their use in the construction of artificial transcription factors. J Biol Chem *276*, 29466-29478.

Dreier, B., Fuller, R.P., Segal, D.J., Lund, C.V., Blancafort, P., Huber, A., Koksch, B., and Barbas, C.F., 3rd (2005). Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. J Biol Chem *280*, 35588-35597.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., and Chandrasegaran, S. (2005). Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res *33*, 5978-5990.

Eming, S.A., Yarmush, M.L., Krueger, G.G., and Morgan, J.R. (1999). Regulation of the spatial organization of mesenchymal connective tissue: effects of cell-associated versus released isoforms of platelet-derived growth factor. Am J Pathol *154*, 281-289.

Epinat, J.C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Paques, F., and Lacroix, E. (2003). A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res *31*, 2952-2962.

Flisikowska, T., Thorey, I.S., Offner, S., Ros, F., Lifke, V., Zeitler, B., Rottmann, O., Vincent, A., Zhang, L., Jenkins, S.*, et al.* (2011). Efficient immunoglobulin gene disruption and targeted replacement in rabbit using zinc finger nucleases. PLoS One *6*, e21045.

Foley, J.E., Yeh, J.R., Maeder, M.L., Reyon, D., Sander, J.D., Peterson, R.T., and Joung, J.K. (2009). Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). PLoS One *4*, e4348.

Friedrich, G., and Soriano, P. (1991). Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev *5*, 1513-1523.

Gabriel, R., Lombardo, A., Arens, A., Miller, J.C., Genovese, P., Kaeppel, C., Nowrouzi, A., Bartholomae, C.C., Wang, J., Friedman, G.*, et al.* (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. Nat Biotechnol *29*, 816-823.

Gaspar, H.B., Bjorkegren, E., Parsley, K., Gilmour, K.C., King, D., Sinclair, J., Zhang, F., Giannakopoulos, A., Adams, S., Fairbanks, L.D.*, et al.* (2006). Successful reconstitution of immunity in ADA-SCID by stem cell gene therapy following cessation of PEG-ADA and use of mild preconditioning. Mol Ther *14*, 505-513.

Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X.*, et al.* (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. Science *325*, 433.

Grizot, S., Epinat, J.C., Thomas, S., Duclert, A., Rolland, S., Paques, F., and Duchateau, P. (2010). Generation of redesigned homing endonucleases comprising DNA-binding domains derived from two different scaffolds. Nucleic Acids Res *38*, 2006-2018.

Grizot, S., Smith, J., Daboussi, F., Prieto, J., Redondo, P., Merino, N., Villate, M., Thomas, S., Lemaire, L., Montoya, G.*, et al.* (2009). Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. Nucleic Acids Res *37*, 5405-5419.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K.*, et al.* (2008). Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J Clin Invest *118*, 3132-3142.

Hacein-Bey-Abina, S., Le Deist, F., Carlier, F., Bouneaud, C., Hue, C., De Villartay, J.P., Thrasher, A.J., Wulffraat, N., Sorensen, R., Dupuis-Girod, S.*, et al.* (2002). Sustained correction of X-linked severe combined immunodeficiency by ex vivo gene therapy. N Engl J Med *346*, 1185-1193.

Hacein-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J.L., Fraser, C.C., Cavazzana-Calvo, M.*, et al.* (2003a). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N Engl J Med *348*, 255-256.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E.*, et al.* (2003b). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science *302*, 415-419.

Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M.*, et al.* (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science *318*, 1920-1923.

Hasty, P., Rivera-Perez, J., Chang, C., and Bradley, A. (1991). Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. Mol Cell Biol *11*, 4509-4517.

Henikoff, S. (1998). Conspiracy of silence among repeated transgenes. Bioessays *20*, 532-535.

Hinnen, A., Hicks, J.B., and Fink, G.R. (1978). Transformation of yeast. Proc Natl Acad Sci U S A *75*, 1929-1933.

Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKelver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B.*, et al.* (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol *27*, 851-857.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C.*, et al.* (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol *29*, 731-734.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C.*, et al.* (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol *28*, 839-847.

Hutter, G., Nowak, D., Mossner, M., Ganepola, S., Mussig, A., Allers, K., Schneider, T., Hofmann, J., Kucherer, C., Blau, O.*, et al.* (2009). Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. N Engl J Med *360*, 692-698.

Kang, E.M., Choi, U., Theobald, N., Linton, G., Long Priel, D.A., Kuhns, D., and Malech, H.L. (2010). Retrovirus gene therapy for X-linked chronic granulomatous disease can achieve stable long-term correction of oxidase activity in peripheral blood neutrophils. Blood *115*, 783-791.

Khan, I.F., Hirata, R.K., and Russell, D.W. (2011). AAV-mediated gene targeting methods for human cells. Nat Protoc *6*, 482-501.

Khan, I.F., Hirata, R.K., Wang, P.R., Li, Y., Kho, J., Nelson, A., Huo, Y., Zavaljevski, M., Ware, C., and Russell, D.W. Engineering of human pluripotent stem cells by AAV-mediated gene targeting. Mol Ther *18*, 1192-1199.

Kim, H.J., Lee, H.J., Kim, H., Cho, S.W., and Kim, J.S. (2009). Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. Genome Res *19*, 1279-1288.

Kim, Y.G., Cha, J., and Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A *93*, 1156-1160.

Kim, Y.G., and Chandrasegaran, S. (1994). Chimeric restriction endonuclease. Proc Natl Acad Sci U S A *91*, 883-887.

Lee, H.J., Kim, E., and Kim, J.S. (2010). Targeted chromosomal deletions in human cells using zinc finger nucleases. Genome Res *20*, 81-89.

Li, L., Wu, L.P., and Chandrasegaran, S. (1992). Functional domains in Fok I restriction endonuclease. Proc Natl Acad Sci U S A *89*, 4275-4279.

Lloyd, A., Plaisier, C.L., Carroll, D., and Drews, G.N. (2005). Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proc Natl Acad Sci U S A *102*, 2232-2237.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D.*, et al.* (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol *25*, 1298-1306.

Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A.*, et al.* (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell *31*, 294-301.

Maguire, A.M., High, K.A., Auricchio, A., Wright, J.F., Pierce, E.A., Testa, F., Mingozzi, F., Bennicelli, J.L., Ying, G.S., Rossi, S.*, et al.* (2009). Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. Lancet *374*, 1597-1605.

Mahfouz, M.M., Li, L., Shamimuzzaman, M., Wibowo, A., Fang, X., and Zhu, J.K. (2011). De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. Proc Natl Acad Sci U S A *108*, 2623-2628.

Malech, H.L., Maples, P.B., Whiting-Theobald, N., Linton, G.F., Sekhsaria, S., Vowells, S.J., Li, F., Miller, J.A., DeCarlo, E., Holland, S.M.*, et al.* (1997). Prolonged production of NADPH oxidase-corrected granulocytes after gene therapy of chronic granulomatous disease. Proc Natl Acad Sci U S A *94*, 12133-12138.

Manno, C.S., Pierce, G.F., Arruda, V.R., Glader, B., Ragni, M., Rasko, J.J., Ozelo, M.C., Hoots, K., Blatt, P., Konkle, B.*, et al.* (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. Nat Med *12*, 342-347.

Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T., and Serikawa, T. (2010). Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. PLoS One *5*, e8870.

Mason, M.R., Tannemaat, M.R., Malessy, M.J., and Verhaagen, J. (2011). Gene therapy for the peripheral nervous system: a strategy to repair the injured nerve? Curr Gene Ther *11*, 75-89.

Meng, X., Noyes, M.B., Zhu, L.J., Lawson, N.D., and Wolfe, S.A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. Nat Biotechnol *26*, 695-701.

Meyer, M., de Angelis, M.H., Wurst, W., and Kuhn, R. (2010). Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. Proc Natl Acad Sci U S A *107*, 15022-15026.

Miller, D.G., Wang, P.R., Petek, L.M., Hirata, R.K., Sands, M.S., and Russell, D.W. (2006). Gene targeting in vivo by adeno-associated virus vectors. Nat Biotechnol *24*, 1022-1026.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A.*, et al.* (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol *25*, 778-785.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J.*, et al.* (2011). A TALE nuclease architecture for efficient genome editing. Nat Biotechnol *29*, 143-148.

Mittelman, D., Moye, C., Morton, J., Sykoudis, K., Lin, Y., Carroll, D., and Wilson, J.H. (2009). Zinc-finger directed double-strand breaks within CAG repeat tracts promote repeat instability in human cells. Proc Natl Acad Sci U S A *106*, 9607-9612.

Moehle, E.A., Rock, J.M., Lee, Y.L., Jouvenot, Y., DeKelver, R.C., Gregory, P.D., Urnov, F.D., and Holmes, M.C. (2007). Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A *104*, 3055-3060.

Morton, J., Davis, M.W., Jorgensen, E.M., and Carroll, D. (2006). Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. Proc Natl Acad Sci U S A *103*, 16370-16375.

Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. Science *326*, 1501.

Munoz, I.G., Prieto, J., Subramanian, S., Coloma, J., Redondo, P., Villate, M., Merino, N., Marenchino, M., D'Abramo, M., Gervasio, F.L.*, et al.* (2011). Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. Nucleic Acids Res *39*, 729-743.

Mutskov, V., and Felsenfeld, G. (2004). Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. EMBO J *23*, 138-149.

Ochiai, H., Fujita, K., Suzuki, K., Nishikawa, M., Shibata, T., Sakamoto, N., and Yamamoto, T. (2010). Targeted mutagenesis in the sea urchin embryo using zinc-finger nucleases. Genes Cells *15*, 875-885.

Ostman, A., Andersson, M., Betsholtz, C., Westermark, B., and Heldin, C.H. (1991). Identification of a cell retention signal in the B-chain of platelet-derived growth factor and in the long splice version of the A-chain. Cell Regul *2*, 503-512.

Ott, M.G., Schmidt, M., Schwarzwaelder, K., Stein, S., Siler, U., Koehl, U., Glimm, H., Kuhlcke, K., Schilz, A., Kunkel, H.*, et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. Nat Med *12*, 401-409.

Palmer, T.D., Thompson, A.R., and Miller, A.D. (1989). Production of human factor IX in animals by genetically modified skin fibroblasts: potential therapy for hemophilia B. Blood *73*, 438-445.

Papapetrou, E.P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L.M., Kadota, K., Roth, S.L., Giardina, P., Viale, A.*, et al.* (2011). Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat Biotechnol *29*, 73-78.

Paques, F., and Duchateau, P. (2007). Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. Curr Gene Ther *7*, 49-66.

Pattanayak, V., Ramirez, C.L., Joung, J.K., and Liu, D.R. (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. Nat Methods *8*, 765-770.

Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L.*, et al.* (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol *26*, 808-816.

Porteus, M.H. (2006). Mammalian gene targeting with designed zinc finger nucleases. Mol Ther *13*, 438-446.

Porteus, M.H., and Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. Science *300*, 763.

Porteus, M.H., and Carroll, D. (2005). Gene targeting using zinc finger nucleases. Nat Biotechnol *23*, 967-973.

Potts, P.R., Porteus, M.H., and Yu, H. (2006). Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. EMBO J *25*, 3377-3388.

Pruett-Miller, S.M., Connelly, J.P., Maeder, M.L., Joung, J.K., and Porteus, M.H. (2008). Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther *16*, 707-717.

Pruett-Miller, S.M., Reading, D.W., Porter, S.N., and Porteus, M.H. (2009). Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. PLoS Genet *5*, e1000376.

Qiu, X., Lu, D., Zhou, J., Wang, J., Yang, J., Meng, P., and Hsueh, J.L. (1996). Implantation of autologous skin fibroblast genetically modified to secrete clotting factor IX partially corrects the hemorrhagic tendencies in two hemophilia B patients. Chin Med J (Engl) *109*, 832-839.

Raper, S.E., Chirmule, N., Lee, F.S., Wivel, N.A., Bagg, A., Gao, G.P., Wilson, J.M., and Batshaw, M.L. (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Mol Genet Metab *80*, 148-158.

Redondo, P., Prieto, J., Munoz, I.G., Alibes, A., Stricher, F., Serrano, L., Cabaniols, J.P., Daboussi, F., Arnould, S., Perez, C.*, et al.* (2008). Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. Nature *456*, 107-111.

Rosser, J.M., and An, W. (2010). Repeat-induced gene silencing of L1 transgenes is correlated with differential promoter methylation. Gene *456*, 15-23.

Roth, D.A., Tawa, N.E., Jr., O'Brien, J.M., Treco, D.A., and Selden, R.F. (2001). Nonviral transfer of the gene encoding coagulation factor VIII in patients with severe hemophilia A. N Engl J Med *344*, 1735-1742.

Rouet, P., Smih, F., and Jasin, M. (1994a). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A *91*, 6064-6068.

Rouet, P., Smih, F., and Jasin, M. (1994b). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol *14*, 8096-8106.

Sander, J.D., Dahlborg, E.J., Goodwin, M.J., Cade, L., Zhang, F., Cifuentes, D., Curtin, S.J., Blackburn, J.S., Thibodeau-Beganny, S., Qi, Y.*, et al.* (2011). Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods *8*, 67-69.

Santiago, Y., Chan, E., Liu, P.Q., Orlando, S., Zhang, L., Urnov, F.D., Holmes, M.C., Guschin, D., Waite, A., Miller, J.C.*, et al.* (2008). Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A *105*, 5809-5814.

Sargent, R.G., Brenneman, M.A., and Wilson, J.H. (1997). Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination. Mol Cell Biol *17*, 267-277.

Shukla, V.K., Doyon, Y., Miller, J.C., DeKelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X.*, et al.* (2009). Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature *459*, 437-441.

Smiell, J.M., Wieman, T.J., Steed, D.L., Perry, B.H., Sampson, A.R., and Schwab, B.H. (1999). Efficacy and safety of becaplermin (recombinant human platelet-derived growth factor-BB) in patients with nonhealing, lower extremity diabetic ulcers: a combined analysis of four randomized studies. Wound Repair Regen *7*, 335-346.

Smih, F., Rouet, P., Romanienko, P.J., and Jasin, M. (1995). Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res *23*, 5012-5019.

Smith, A.J., and Berg, P. (1984). Homologous recombination between defective neo genes in mouse 3T6 cells. Cold Spring Harb Symp Quant Biol *49*, 171-181.

Smith, J., Bibikova, M., Whitby, F.G., Reddy, A.R., Chandrasegaran, S., and Carroll, D. (2000). Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic Acids Res *28*, 3361-3369.

Smith, J., Grizot, S., Arnould, S., Duclert, A., Epinat, J.C., Chames, P., Prieto, J., Redondo, P., Blanco, F.J., Bravo, J.*, et al.* (2006). A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. Nucleic Acids Res *34*, e149.

Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S. (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. Nature *317*, 230-234.

Smithies, O., Koralewski, M.A., Song, K.Y., and Kucherlapati, R.S. (1984). Homologous recombination with DNA introduced into mammalian cells. Cold Spring Harb Symp Quant Biol *49*, 161-170.

Sollu, C., Pars, K., Cornu, T.I., Thibodeau-Beganny, S., Maeder, M.L., Joung, J.K., Heilbronn, R., and Cathomen, T. (2010). Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. Nucleic Acids Res *38*, 8269-8276.

Sun, N., Liang, J., Abil, Z., and Zhao, H. (2012). Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst *8*, 1255-1263.

Szczepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D.J., and Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol *25*, 786-793.

Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell *33*, 25-35.

Szymczak, A.L., Workman, C.J., Wang, Y., Vignali, K.M., Dilioglou, S., Vanin, E.F., and Vignali, D.A. (2004). Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. Nat Biotechnol *22*, 589-594.

Takasu, Y., Kobayashi, I., Beumer, K., Uchino, K., Sezutsu, H., Sajwan, S., Carroll, D., Tamura, T., and Zurovec, M. (2010). Targeted mutagenesis in the silkworm Bombyx mori using zinc finger nuclease mRNA injection. Insect Biochem Mol Biol *40*, 759-765.

Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., and Melton, D.W. (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. Cell *56*, 313-321.

Titeux, M., Pendaries, V., Zanta-Boussif, M.A., Decha, A., Pironon, N., Tonasso, L., Mejia, J.E., Brice, A., Danos, O., and Hovnanian, A. (2010). SIN retroviral vectors expressing COL7A1 under human promoters for ex vivo gene therapy of recessive dystrophic epidermolysis bullosa. Mol Ther *18*, 1509-1518.

Todaro, G.J., and Green, H. (1963). Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. J Cell Biol *17*, 299-313.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Fu, F., Maeder, M.L., Joung, J.K., and Voytas, D.F. (2009). High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature *459*, 442-445.

Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature *435*, 646-651.

van Rensburg, R., Beyer, I., Yao, X.Y., Wang, H., Denisenko, O., Li, Z.Y., Russell, D.W., Miller, D.G., Gregory, P., Holmes, M*., et al.* (2012). Chromatin structure of two genomic sites for targeted transgene integration in induced pluripotent stem cells and hematopoietic stem cells. Gene Ther.

Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. (1997). Structure of the multimodular endonuclease FokI bound to DNA. Nature *388*, 97-100.

Wright, D.A., Townsend, J.A., Winfrey, R.J., Jr., Irwin, P.A., Rajagopal, J., Lonosky, P.M., Hall, B.D., Jondle, M.D., and Voytas, D.F. (2005). High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J *44*, 693-705.

Wu, J., Kandavelou, K., and Chandrasegaran, S. (2007). Custom-designed zinc finger nucleases: what is next? Cell Mol Life Sci *64*, 2933-2944.

Wyman, C., and Kanaar, R. (2006). DNA double-strand break repair: all's well that ends well. Annu Rev Genet *40*, 363-383.

Young, J.J., Cherone, J.M., Doyon, Y., Ankoudinova, I., Faraji, F.M., Lee, A.H., Ngo, C., Guschin, D.Y., Paschon, D.E., Miller, J.C.*, et al.* (2011). Efficient targeted gene disruption in the soma and germ line of the frog Xenopus tropicalis using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A *108*, 7052-7057.

Yuan, J., Wang, J., Crain, K., Fearns, C., Kim, K.A., Hua, K.L., Gregory, P.D., Holmes, M.C., and Torbett, B.E. (2012). Zinc-finger nuclease editing of human cxcr4 promotes HIV-1 CD4(+) T cell resistance and enrichment. Mol Ther *20*, 849-859.

Zambrowicz, B.P., Imamoto, A., Fiering, S., Herzenberg, L.A., Kerr, W.G., and Soriano, P. (1997). Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. Proc Natl Acad Sci U S A *94*, 3789-3794.

Zhang, Z., Slobodianski, A., Ito, W.D., Arnold, A., Nehlsen, J., Weng, S., Lund, N., Liu, J., Egana, J.T., Lohmeyer, J.A.*, et al.* (2011). Enhanced collateral growth by double transplantation of gene-nucleofected fibroblasts in ischemic hindlimb of rats. PLoS One *6*, e19192.

Zou, J., Maeder, M.L., Mali, P., Pruett-Miller, S.M., Thibodeau-Beganny, S., Chou, B.K., Chen, G., Ye, Z., Park, I.H., Daley, G.Q.*, et al.* (2009). Gene targeting of a disease-related gene in human induced pluripotent stem and embryonic stem cells. Cell Stem Cell *5*, 97-110.

Zou, J., Sweeney, C.L., Chou, B.K., Choi, U., Pan, J., Wang, H., Dowey, S.N., Cheng, L., and Malech, H.L. (2011). Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease-mediated safe harbor targeting. Blood *117*, 5561-5572.