Genetic Correction of Duchenne Muscular Dystrophy using Engineered Nucleases

by

David Gerard Ousterout Jr.

Department of Biomedical Engineering
Duke University

Date:_____

Approved:

_____
Charles A. Gersbach, Supervisor

_____
Kam W. Leong

_____
Nenad Bursac

_____
Ashutosh Chilkoti

_____
Xiao Xiao

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Biomedical Engineering in the Graduate School
of Duke University

2014

ABSTRACT

Genetic Correction of Duchenne Muscular Dystrophy using Engineered Nucleases

by

David Gerard Ousterout Jr.

Department of Biomedical Engineering
Duke University

Date:_____
Approved:

_____
Charles A. Gersbach, Supervisor

_____
Kam W. Leong

_____
Nenad Bursac

_____
Ashutosh Chilkoti

_____
Xiao Xiao

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Biomedical Engineering in the Graduate School of
Duke University

2014

Abstract

Duchenne muscular dystrophy (DMD) is a severe hereditary disorder caused by a loss of dystrophin, an essential musculoskeletal protein. Decades of promising research have yielded only modest gains in survival and quality of life for these patients and there have been no approved gene therapies for DMD to date. There are two significant hurdles to creating effective gene therapies for DMD; it is difficult to deliver a replacement dystrophin gene due to its large size and current strategies to restore the native dystrophin gene likely require life-long administration of a gene-modifying drug. This thesis presents a novel method to address these challenges through restoring dystrophin expression by genetically correcting the native dystrophin gene using engineered nucleases that target one or more exons in a mutational hotspot in exons 45-55 of the dystrophin gene. Importantly, this hotspot mutational region collectively represents approximately 62% of all DMD mutations. In this work, we utilize various engineered nuclease platforms to create genetic modifications that can correct a variety of DMD patient mutations.

Initially, we demonstrate that genome editing can efficiently correct the dystrophin reading frame and restore protein expression by introducing micro-frameshifts in exon 51, which is adjacent to a hotspot mutational region in the dystrophin gene. Transcription activator-like effector nucleases (TALENs) were

engineered to mediate highly efficient gene editing after introducing a single TALEN pair targeted to exon 51 of the dystrophin gene. This led to restoration of dystrophin protein expression in cells from DMD patients, including skeletal myoblasts and dermal fibroblasts that were reprogrammed to the myogenic lineage by MyoD. We show that our engineered TALENs have minimal cytotoxicity and exome sequencing of cells with targeted modifications of the dystrophin locus showed no TALEN-mediated off-target changes to the protein coding regions of the genome, as predicted by *in silico* target site analysis.

In an alternative approach, we capitalized on the recent advances in genome editing to generate permanent exclusion of exons by using zinc-finger nucleases (ZFNs) to selectively remove sequences important in specific exon recognition. This strategy has the advantage of creating predictable frame restoration and protein expression, although it relies on simultaneous nuclease activity to generate genomic deletions. ZFNs were designed to remove essential splicing sequences in exon 51 of the dystrophin gene and thereby exclude exon 51 from the resulting dystrophin transcript, a method that can potentially restore the dystrophin reading frame in up to 13% of DMD patients. Nucleases were assembled by extended modular assembly and context-dependent assembly methods and screened for activity in human cells. Selected ZFNs had moderate observable cytotoxicity and one ZFN showed off-target activity at two chromosomal loci. Two active ZFN pairs flanking the exon 51 splice acceptor site were

transfected into DMD patient cells and a clonal population was isolated with this region deleted from the genome. Deletion of the genomic sequence containing the splice acceptor resulted in the loss of exon 51 from the dystrophin mRNA transcript and restoration of dystrophin expression *in vitro*. Furthermore, transplantation of corrected cells into the hind limb of immunodeficient mice resulted in efficient human dystrophin expression localized to the sarcolemma.

Finally, we exploited the increased versatility, efficiency, and multiplexing capabilities of the CRISPR/Cas9 system to enable a variety of otherwise challenging gene correction strategies for DMD. Single or multiplexed sgRNAs were designed to restore the dystrophin reading frame by targeting the mutational hotspot at exons 45-55 and introducing either intraexonic small insertions and deletions, or large deletions of one or more exons. Significantly, we generated a large deletion of 336 kb across the entire exon 45-55 region that is applicable to correction of approximately 62% of DMD patient mutations. We show that, for selected sgRNAs, CRISPR/Cas9 gene editing displays minimal cytotoxicity and limited aberrant mutagenesis at off-target chromosomal loci. Following treatment with Cas9 nuclease and one or more sgRNAs, dystrophin expression was restored in Duchenne patient muscle cells *in vitro*. Human dystrophin was detected *in vivo* following transplantation of genetically corrected patient cells into immunodeficient mice.

In summary, the objective of this work was to develop methods to genetically correct the native dystrophin as a potential therapy for DMD. These studies integrate the rapid advances in gene editing technologies to create targeted frameshifts that restore the dystrophin gene around patient mutations in non-essential coding regions. Collectively, this thesis presents several gene editing methods that can correct patient mutations by modification of specific exons or by deletion of one or more exons that results in restoration of the dystrophin reading frame. Importantly, the gene correction methods described here are compatible with leading cell-based therapies and *in vivo* gene delivery strategies for DMD, providing an avenue towards a cure for this devastating disease.

Dedication

This work is dedicated to my parents; they have provided me with every opportunity to begin this path and their limitless positivity, love, and support has seen me through it.

Contents

List of Tables

List of Figures

Acknowledgements

I have many people to thank along the way, but no one more important than my advisor Charlie Gersbach. Charlie – thank you for your unwavering support and for helping me keep the big picture in mind when the little things seemed overwhelming. I'm still impressed that you have kept your door open through year after year of endless questions, concerns, and random thoughts. Thank you to my committee members as well for their valuable guidance and contributions as I completed this work.

I have made so many amazing friendships throughout graduate school; hopefully, I can do a little justice here to the some of the many people who made getting through the demands of graduate school possible. One of the best experiences of my life has been coming to Duke to start in a new lab, with all of the fun and challenges that came with it. To Ami – I certainly could not have survived the past five years without coffee time and our friendship. Not bad for a couple of people who, at first, never thought that they would get along, like ever. I would also like to thank Pablo for his invaluable mentorship and friendship that proved to be the difference in all of my experiments failing forever and getting a few working from time to time. Pratiksha, thank you for reading this giant block of text and for all of the laughs, blue and green drinks, memes, youtubes, and advice – "Are you done yet?". Tyler, thank you for the many memories over the years, including our overly critical "debates", pressure pump

dances around Ami, and awkward references to “Papa Bear” and “Papa O.” Lauren – thank you for the many unsolicited words of encouragement and kindness, for being the co-“lab enforcer”, and just all around hilarious – “Cashed!”. Of course, I would be remiss if I didn’t mention that much of my career here was spent toiling on ways to teach Matt how to calculate dilutions, and along the way I gained a great friend who taught me so much about becoming a better mentor as well. Jacqueline – best of luck with the mess that I leave behind and thank you for being so generous – I am sure you will be tremendously successful! To all of the members of the Gersbach lab, including Josh, Dewran, and Isaac, thank you for making lab a fun place to be every – single – day and for even more fun outside of it.

And to The Gang – Lucas, Jaclyn, Josh and Cindy – thank you for the many jokes about “SCIENCE!”, excursions to Flavin and other exotic locales, and all around good times. Thank you to my parents, this work is dedicated to them because it simply could not exist without their support and love and all of the opportunities they created for me.

Finally, to my dear, sweet wife Jamie, who has put up with the demands that graduate school has placed on my time and sanity. You are my rock; your patience while I pursued my degree is nothing short of heroic and you have always made every day better than the last. Thank you.

Chapter 1. Introduction and Specific Aims

1.1 Rationale and Objectives

Duchenne muscular dystrophy (DMD) is a severe X-linked hereditary disease for which no effective treatments exist. The molecular basis of DMD is a deleterious mutation in the dystrophin gene, discovered over 25 years ago [1, 2], that leads to the complete absence of this essential skeletal muscle protein. DMD presents with progressive muscle wasting that leads to death within the third decade of life due to respiratory or heart failure. The current standard of care for DMD is palliative and has focused on managing respiratory and cardiac failure with steroid and ACE inhibitor therapy [3]. Despite significantly increasing the life span of these patients, most patients do not live beyond early adulthood and are increasingly dying of cardiac complications. There have been various novel approaches to treating DMD that focus on transiently restoring the mutant dystrophin gene [4] or replacing it using gene therapy to deliver a functional dystrophin gene [5] to muscle tissue. However, significant challenges associated with these strategies include insertional mutagenesis, the use of synthetic promoters, and transient gene restoration.

Recent advances have enabled targeted genetic therapies based on engineered enzymes that exploit cellular DNA repair pathways to create site-specific, predefined genetic modifications in complex genomes [6]. These synthetic enzymes are commonly based on meganucleases [7], zinc finger nucleases (ZFNs) [6], transcription activator-like

effector nucleases (TALENs) [8, 9], and more recently the RNA-guided CRISPR/Cas system [10, 11]. These nucleases create site-specific double-strand breaks (DSBs) at predefined sites to stimulate endogenous DNA repair pathways that are used to achieve desired genetic changes.

The *overall objective* of this thesis work is to develop engineered nucleases that can specifically edit and correct the human dystrophin gene. The *central hypothesis* is that genetic correction of the dystrophin gene will restore the dystrophin reading frame and rescue dystrophin expression.

1.2 Specific Aims

Aim 1: Engineer nucleases to introduce targeted micro-frameshifts in exon 51 to restore the dystrophin reading frame.

TALE nucleases were engineered to generate small insertions and deletions in exon 51 of the dystrophin gene to restore aberrant dystrophin reading frames. The *working hypothesis* was that micro-frameshifts in exon 51 that corrected the dystrophin reading frame would result in restored dystrophin protein expression. TALENs targeted to exon 51 were delivered to cell lines from DMD patients with mutations correctable by exon 51 modification. Dystrophin gene correction was assessed at the genomic and protein levels *in vitro*. Nuclease-related toxicity was monitored by measuring cytotoxicity in human cells, exome sequencing of genetically modified cell populations, and *in silico* prediction of potential off-target sites in the human genome.

Aim 2: Engineer nucleases to delete exon 51 from the genome to restore the dystrophin reading frame.

A panel of zinc-finger nucleases was engineered to target genomic loci flanking exon 51 of the dystrophin gene. The *working hypothesis* was that co-expression of two nucleases would result in genetic deletion of essential exon 51 splicing sequences, resulting in the loss of exon 51 from the dystrophin transcript and restoration of dystrophin protein expression. Dozens of zinc-finger nucleases were created and screened for activity against episomal and chromosomal targets. Two selected ZFN pairs flanking the exon 51 splice acceptor were delivered to skeletal myoblasts from DMD patients and a clonal population of genetically corrected cells was isolated. Changes to the dystrophin gene and mRNA transcript were monitored *in vitro*, and restoration of dystrophin protein expression was assessed both *in vitro* and *in vivo*. Off-target activity of selected ZFNs was assessed in human cells by measuring ZFN-specific cytotoxicity and off-target mutagenesis.

Aim 3: Develop broad patient-specific and universal gene correction strategies to correct the dystrophin gene.

CRISPR/Cas9 gene correction can utilize single and multiplex gene editing to introduce small insertions and deletions within exons or to efficiently delete targeted regions of the dystrophin gene. The *working hypothesis* is that targeting the dystrophin gene across the hotspot mutational region in exons 45-55 will enable patient-specific or

universal gene editing approaches that can address up to 62% of DMD patient mutations. The *S. pyogenes* CRISPR/Cas9 system was utilized to introduce targeted micro-frameshifts in specific exons across exons 45-55, to delete the entire exon 51 locus, or to delete the entire region containing exons 45-55. Modification of the dystrophin gene was monitored at the DNA and RNA level *in vitro*, and protein levels *in vitro* and following *in vivo* transplantation of a bulk population of CRISPR/Cas9 treated cells. Off-target effects for selected CRISPR/Cas9 targets were monitored by cytotoxicity assays and direct interrogation of predicted off-target chromosomal loci by PCR.

1.3. Significance

This approach builds on the promise of genome editing heralded by the recent entrance of ZFNs into clinical trials for disruption of the HIV-1 co-receptor CCR5 [12, 13] and disruption of the glucocorticoid receptor in T cells for glioblastoma treatment. Importantly, the genetic correction methods presented in this thesis could be expanded to a range of diseases that are correctable by restoring disrupted reading frames introduced by intragenic insertions, deletions or aberrant stop codons in non-essential coding regions, including Collagen type VII-associated dystrophic epidermolysis bullosa [14], Fukuyama congenital muscular dystrophy [15], and Limb-girdle muscular dystrophy type 2B [15].

We propose that genome editing is a promising method that addresses many of the present challenges for effective long-term correction of the genetic defects causing

DMD. In particular, genome editing provides a solution that is permanent and may require as little as one treatment to achieve therapeutic effect. Moreover, genome editing strategies for DMD are compatible with existing cell-based therapies and gene transfer modalities for skeletal muscle gene therapy, such as adeno-associated virus gene transfer. This thesis analyzes a diverse range of gene editing platforms to implement gene correction strategies that restore the reading frame within the native dystrophin gene and presents genome editing as a unique method to permanently address the genetic basis of DMD.

Chapter 2. Literature Review

2.1. Duchenne muscular dystrophy and the dystrophin gene

Duchenne muscular dystrophy (DMD) is a severe monogenic X-linked hereditary disease that occurs in about 1 in 3500 male births in the United States [16]. The pathogenesis of DMD manifests during the first decade of life and results in progressive weakness and muscle degeneration leading to death. Typically, most DMD patients die in the third decade of life due to respiratory and/or cardiac complications. Dystrophic muscle tissue is characterized by continuous cycles of necrosis and regeneration, resulting in centrally nucleated fibers with increased permeability and abnormally high adipose and fibrotic tissue content [16]. The causative agent of DMD is a defect in the gene encoding dystrophin, an essential musculoskeletal protein that is completely absent in these patients. Dystrophin is a large, rod-shaped protein with four major domains: the N-terminus, rod-domain, cysteine-rich domain, and C-terminus [17]. The C-terminus of dystrophin associates with numerous other proteins to form the dystrophin glycoprotein complex (DGC), a complex that acts as a linker between the cytoskeleton (via actin bound at the N-terminus of dystrophin) and laminin in the extracellular matrix (bound through the DGC). It has been hypothesized that the DGC, and dystrophin specifically, protects the sarcolemma from excessive force during muscle contraction and relaxation [18]. Frame-disrupting or premature stop-codon truncations of dystrophin break this mechanical link. The loss of this link results in sarcolemma

instability and muscle necrosis that eventually results in fibrosis and loss of muscle function [19].

The dystrophin locus is the largest identified gene in the human genome at approximately 2.4Mb. The primary skeletal muscle isoform, Dp427m, results in a 13,993 base pair mRNA encoding 79 exons that produce a 3,685 amino acid protein (~427 kDa in size) with an isoform-specific N-terminus [20]. The majority of dystrophin mutations that cause DMD are deletions of exons that result in reading frame disruption [5]. However, the loss of internal dystrophin exons that retain the proper reading frame cause the less severe Becker muscular dystrophy (BMD). This has led to efforts create internally minimized dystrophin proteins that can be packaged into small viruses, such as the AAV, or to restore the disrupted native dystrophin reading frame in Duchenne patients, thereby producing internally deleted, but functional dystrophin proteins [4]. Deletions that occur in the exon 45-55 region, contained within the rod domain, and leave the rest of the protein intact generally produce highly-functional proteins, and many carriers of this mutation are asymptomatic [4]. An interesting phenomenon observed in BMD patients are divergent observed phenotypic outcomes in Becker muscular dystrophy (BMD) patient siblings with the same underlying mutations in the dystrophin gene [21, 22]. This suggests that there are other unknown factors that may alter disease progression and severity in patients with internal dystrophin deletions. Despite this uncertainty, deletions in the exon 45-55 region that disrupt the dystrophin

protein are an attractive class of DMD mutations to target, since correction would presumably result in functional dystrophin and improve the DMD phenotype [22]. Furthermore, it is estimated that up to 62% of patients with DMD could benefit from correction in exons 45-55 [4]. Notably, exon 51 is frequently adjacent to frame-disrupting deletions in DMD patients, account for nearly 13% of all DMD mutations, and has been targeted in clinical trials for oligonucleotide-based exon skipping with promising early therapeutic results [4, 23, 24]. Thus, exon 51 is of particular interest in designing genetic therapies for DMD.

This thesis takes advantage of the highly defined genetic basis of DMD as it is especially amenable to the gene correction methods presented in this work. In particular, deletion of exon 51 is ideally suited for exploring permanent correction of the native dystrophin gene by genome editing because it is an established therapeutic target in current exon-skipping clinical trials and is likely to produce highly functional proteins. Moreover, deletion of the entire exon 45-55 region could potentially address 62% of patient mutations and generate a dystrophin protein that retains a high level of functionality, based on similar proteins found in asymptomatic carriers or the mild disease Becker muscular dystrophy.

2.2. Cell-based therapies for DMD

Cell-based therapies for DMD aim to introduce functional dystrophin expression in patient muscle by genetically correcting autologous patient cells or by isolating

healthy allogeneic cells that can be used to repopulate and replace dystrophic tissue in a DMD patient. Initially, cell-based therapy focused on transplantation of allogeneic healthy donor muscle progenitors directly into DMD patient muscle by injection and was combined with immunosuppression [25]. Several early clinical trials explored combining different methods of immune suppression in tandem with transplantation of donor allogeneic myoblasts. These early trials had varied success, but at best only a few patients had meaningful donor engraftment and restored dystrophin expression [25]. Some of the challenges associated with these therapies included immune rejection [26], insufficient donor engraftment [27], and poor migration from the injection site [28]. Therefore, while conceptually simple, these methods are technically difficult and require numerous injections directly into target tissue and poor engraftment of donor tissue has precluded meaningful clinical results.

New approaches based on genetic engineering of autologous cells have led to renewed interest in cell-based therapies for DMD. It has been shown previously that autologous patient cells can be genetically modified to stably express full-length dystrophin [29, 30] or minidystrophin [31]. A notable and innovative approach introduced a human artificial chromosome (HAC) containing an entire intact DMD locus that successfully restored dystrophin expression in a progenitor cell population. These cells could home to and repopulate dystrophic tissue in an animal model of DMD [30]. However, isolation of these progenitor cell populations, as well as the subsequent

genetic engineering and identification of appropriately corrected cells, is technically difficult and time consuming. In contrast, genome editing offers a streamlined, highly efficient approach to directly correct the dystrophin gene *in situ*, thereby obviating the need for extensive genetic engineering required for modifying and characterizing autologous cells containing HACs. More recently, site-specific integration of a minidystrophin gene cassette into a nonessential genetic locus was used to restore dystrophin expression in human myoblasts [31]. However, a major concern for this approach is the possibility of off-target integration of the gene cassette that may potentially interfere with oncogenes and tumor suppressors.

Site-specific genome engineering is rapidly becoming a robust, easy-to-use platform for creating desired genetic changes or additions. The correction of the native gene or site-specific replacement of the entire dystrophin gene may offer a superior method to restore the dystrophin gene in autologous patient cell lines. These methods are compatible with a myriad of cell types currently under investigation for cell-based therapies for muscular dystrophies, including induced pluripotent stem cells [29, 32], bone marrow-derived progenitors [33], skeletal muscle progenitors [34], mesoangioblasts [30], CD133+ cells [35], and dermal fibroblasts [36]. Additionally, advances in immortalization of human myogenic cells may greatly simplify clonal derivation of genetically corrected myogenic cells [37, 38]. These cell types can be modified by genome editing tools *in vitro* by transfection or electroporation of plasmid

or mRNA, transduction by integrase-deficient lentivirus [39], or treatment with cell-permeable nucleases [40, 41]. Of these, one notable approach is to create iPSCs from autologous patient cells, correct the dystrophin gene *ex vivo*, differentiate the iPSCs towards the myogenic lineage, and inject the cells to repopulate dystrophic tissue [29-32, 42, 43].

This work will focus on genetic correction of the native gene using non-homologous end-joining DNA repair that, as discussed below in Chapter 2.4, requires only delivery of the gene editing nucleases and is an excellent candidate for highly efficient delivery to and gene modification of cell types that can be utilized to reconstitute dystrophin expression in skeletal muscle

2.3. Gene and molecular therapies for DMD

2.3.1. Gene transfer to skeletal muscle

Traditionally, gene therapy aims to deliver functional replacement proteins either directly or by gene transfer to a target tissue. Gene transfer vectors for skeletal muscle include lentivirus [44, 45], adeno-associated virus [46-56], adenovirus [57], and non-viral methods. One of the leading gene transfer vectors for *in vivo* gene therapy is adeno-associated virus (AAV), which is a small, non-pathogenic dependovirus that typically encodes a single-strand DNA genome with a maximum packaging capacity of ~4.7kb [58]. AAV vectors based on AAV2 pseudotyped with alternative muscle-tropic AAV capsids, such as AAV2/1, AAV2/6, AAV2/7, AAV2/8, AAV2/9, and the more recent

AAV2.5 and AAV/SASTG vectors have been shown to efficiently transduce skeletal muscle by systemic and local delivery [46-56] with sustained gene transfer measuring at least several months. More recently, the first AAV-based gene therapy was approved as a product in Europe and is also administered by intramuscular delivery [59]. However, the large size of dystrophin (>11kb cDNA) complicates gene transfer strategies, in particular with the use of adeno-associated virus, due to its limited packaging capacity of approximately 4.7kb [49]. This led to the development of novel minidystrophin genes [60], which can be packaged into adeno-associated virus vectors to deliver a functional, but truncated, dystrophin gene [5] to muscle tissue [49]. These miniaturized dystrophin proteins ("minidystrophins") may be sufficient to ameliorate the symptoms of DMD [60], though careful selection of appropriate minimal dystrophin proteins is still under investigation [61]. One minidystrophin, termed Δ3990, has shown favorable functional profiles [60, 61] and was investigated in Phase I clinical trials for AAV-based gene transfer [48]. However, AAV gene transfer is generally limited to postmitotic skeletal muscle fibers and thus far no study has demonstrated sustained gene transfer in progenitor cell populations. This is an important aspect of potential DMD therapies to ensure that corrected genomes are not lost during tissue regeneration following normal muscle damage and repair. The non-integrative nature of AAV also means that gene transfer vectors in progenitor cell populations *in vivo* would be quickly diluted out from transduced progenitor populations. Therefore, it is likely that prolonged clinical efficacy

would require intermittent readministration of the gene transfer vector, which presents significant challenges in overcoming primary immunity to the vector after initial treatment. However, AAV-based gene transfer is still an attractive method for long-term expression of a minidystrophin transgene and methods to overcome vector immunity are an area of active investigation [15]. Therefore, AAV remains a valuable tool for systemic delivery of transgenes, such as gene editing enzymes, to skeletal muscle.

Non-viral methods are also an attractive method for efficient, tissue-specific gene transfer, though their efficiency in skeletal muscle expression is limited. Skeletal muscle has been shown to readily take up plasmid DNA [62], which would greatly simplify gene transfer to skeletal muscle at low cost. These methods were expanded upon to increase efficiency by electroporation [63], hydrostatic pressure [64], ultrasound [65], and non-ionic carriers [66]. Nanoparticles can also be programmed to deliver a payload to specific tissues by conjugating targeting peptides to the nanoparticle complex, enabling versatile control over delivery of gene editing therapeutics [67]. Moreover, these complexes can be used to deliver therapeutic peptides or RNA in addition to conventional gene transfer via DNA. Notably, it has been demonstrated that nanoparticles can mediate *in vivo* gene editing by simultaneously delivering both the gene editing agent and a suitable donor template [68]. Thus, non-viral methods may be a promising method to deliver short bursts of gene editing drugs.

There is a significant opportunity to combine these rapid advances in gene transfer to skeletal muscle with genome editing strategies to achieve *in situ* gene correction *in vivo*. This work aims to develop novel genome editing technologies that are compatible with many of the leading gene therapy vectors, including AAV. Importantly, genome editing would also enable permanent gene correction following transient delivery of targeted nucleases, in contrast to the transient delivery of replacement dystrophin genes that relies on exogenous expression to achieve therapeutic effect. Correction of the native genome directly also ensures physiologic expression and expression of all dystrophin isoforms. Therefore, genome editing would be an important advance in treating DMD that is compatible with established gene transfer vectors, such as AAV, integrase-deficient lentivirus, or other non-viral DNA delivery methods.

2.3.2. Oligonucleotide-mediated exon skipping

Restoration of the native dystrophin gene product may lead to improved clinical outcomes by salvaging the maximum amount of protein functionality [69]. This is a relatively new area of the gene and molecular therapy field that has seen new strategies developed to selectively exclude portions of a damaged, out-of-frame gene or to force read-through of aberrant stop codons. The most successful approach to date is based on small oligonucleotides that transiently and selectively remove exons ("exon skipping") from dystrophin mRNA to restore the reading frame of dystrophin [4]. This approach has led to several successful Phase I/II clinical trials [24, 70, 71], including a recent trial

that demonstrated a significant increase in mean walking distance in patients continuously treated for 48 weeks by intravenous injection with the exon skipping compound eteplirsen (previously AVI-4658), which acts to selectively skip exon 51 [23]. However, delivery of exon skipping molecules at sufficient therapeutic concentrations remains a significant challenge for some tissue types, especially cardiac tissue [5]. Additionally, some exons cannot be skipped exclusively, and must be skipped along with other exons (i.e. exon 44), which may further impact the resulting functionality of the truncated protein [69], as well as the therapeutic efficacy given the requirement for sufficient intracellular levels of more than one compound. Finally, small molecules that can force ribosomal stop codon read-through [72] have been investigated as possible agents to restore truncated dystrophin proteins, including PTC124 [73] (rebranded as Ataluren) and gentamicin [74]. However, these compounds have shown little, if any, therapeutic benefit for treating DMD, perhaps due to insufficient induction of full-length dystrophin expression.

Genome editing is a valuable approach to permanently recapitulate all of the above methods to restore dystrophin expression by rewriting the dystrophin gene. Genomic therapies for DMD would be a significant benefit over transient therapies that must be continuous for the life of the patient to maintain therapeutic efficacy. These continuous therapies also require inconvenient weekly injections of large quantities of expensive therapeutic agents. Importantly, genome editing can theoretically accomplish

therapeutic gene correction in a single dose since genetic modifications result in permanent genetic changes that last for the life of a modified cell and its progeny.

2.4. Targeted genome editing

Genome editing with engineered site-specific endonucleases has emerged as a new technology to selectively replace or correct disrupted genes, in contrast to conventional gene therapy methods of gene addition [75]. These engineered nucleases act by creating a targeted double-strand break in the genome that stimulates cellular DNA repair through either homology-directed repair (HDR) or non-homologous end-joining (NHEJ) (**Figure 1**). HDR uses a designed synthetic donor DNA template to guide repair and can be used to create specific sequence changes to genome, including the targeted addition of whole genes. HDR has enabled integration of gene cassettes of up to 8kb in the absence of selection at high frequency (~6%) in human cells [76]. Generally, gene correction strategies have been based solely on HDR, the efficiency of which is dependent on cell-cycle state and delivery of an exogenous DNA template [77-81]. In many cases, antibiotic selection is used in tandem with genome editing for gene correction in cell types with low levels of HDR repair [78-80]. Gene targeting may also be enhanced by using AAV-based delivery of donor templates and gene editing enzymes [82-86]. Moreover, AAV delivery of zinc finger nucleases (ZFNs) with a donor template has been shown to mediate efficient gene targeting *in vivo* [71, 87]. HDR-based genome editing is a valuable strategy to introduce a functional dystrophin gene cassette

at a predefined point of the genome [31], offering controlled expression and minimizing the potential deleterious effects of random integration. Moreover, some patient deletions may necessitate addition of absent essential coding regions to the dystrophin gene to restore function.

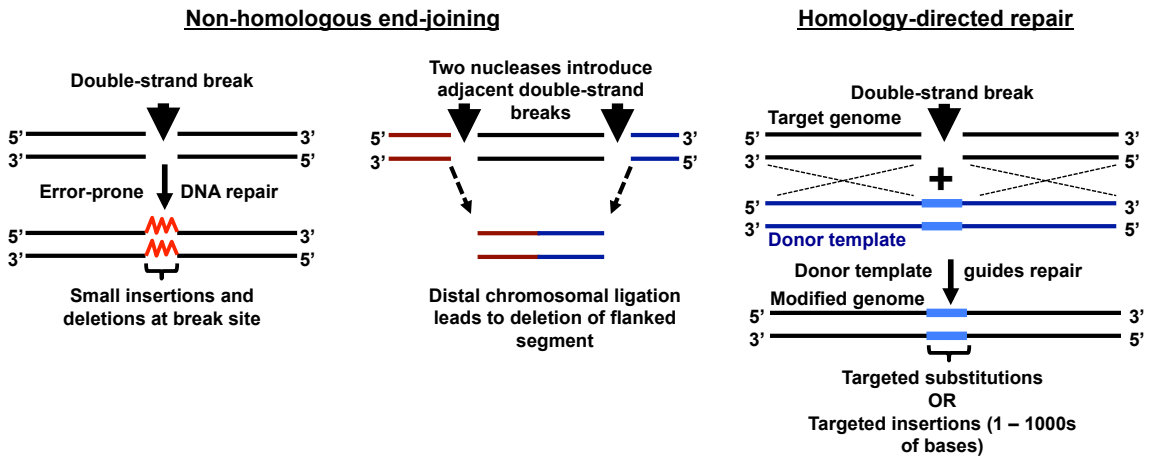


Figure 1: Mechanisms of DNA repair following the creation of a double-strand break by an engineered nuclease.

Although HDR is extremely valuable for restoring the complete coding sequence of the mutant gene, evidence from studies investigating transient exon skipping [88] and premature stop codon read-through [72] demonstrate that restoring partially truncated proteins may also provide therapeutic benefit in a large fraction of DMD patients. The template-independent re-ligation of DNA ends by NHEJ is a stochastic, error-prone repair process that introduces random small insertions and deletions at the DNA breakpoint. Small insertions and deletions resulting from NHEJ repair could permanently restore a disrupted reading frame or remove a premature stop codon, thereby ameliorating the symptoms of some genetic diseases in a similar manner to

current transient methods [89, 90]. Traditionally, gene editing by NHEJ has been used in mammalian cells to disrupt genes [6] or delete chromosomal segments [91, 92], but it has been proposed that genetic mutations created by endonucleases could be utilized to restore an aberrant reading frame [36, 89]. Moreover, the NHEJ gene repair pathway operates in all cell cycle states, while the efficiency of HDR is cell cycle-dependent. Given the high number of patient mutations correctable by frame restoration alone, NHEJ presents an optimal DNA repair pathway to efficiently correct the dystrophin gene. Thus, this thesis explores this opportunity to develop NHEJ-based gene correction strategies to restore the native dystrophin reading frame without the need for exogenous DNA to replace a defective gene.

2.5. Genome editing platforms

2.5.1. Overview

There are numerous platforms for generating site-specific gene modifications in a human genome, but to date the most successful have been based on ZFNs [93], TALENs [75, 94] and more recently, the RNA-guided CRISPR/Cas system [10, 11] (**Figure 2**). These systems are at present the most developed publicly available platforms for robust and efficient targeted gene editing. In particular, the recent development of TALENs and CRISPR/Cas9 has dramatically advanced genome editing due to their high rate of successful and efficient genetic modification [8-11, 75, 95-103]. In addition, there are several others gene editing systems available, including oligonucleotide-mediated exon

skipping [4, 70], meganucleases [104], triplex-forming oligonucleotide (TFO) complexes [105, 106], integrases [107, 108] and programmable recombinases based on zinc finger [109-111] or TALE DNA-binding domains [112]. However, meganucleases have proven to be difficult to engineer due to interdependence of the DNA-binding and cleavage domains, while TFO complexes have demonstrated relatively low levels of gene modification. Programmable recombinases are a promising next-generation gene editing technology, but target site requirements, overall efficiency, and unknown off-target effects are still major challenges to the widespread adoption of this technology.

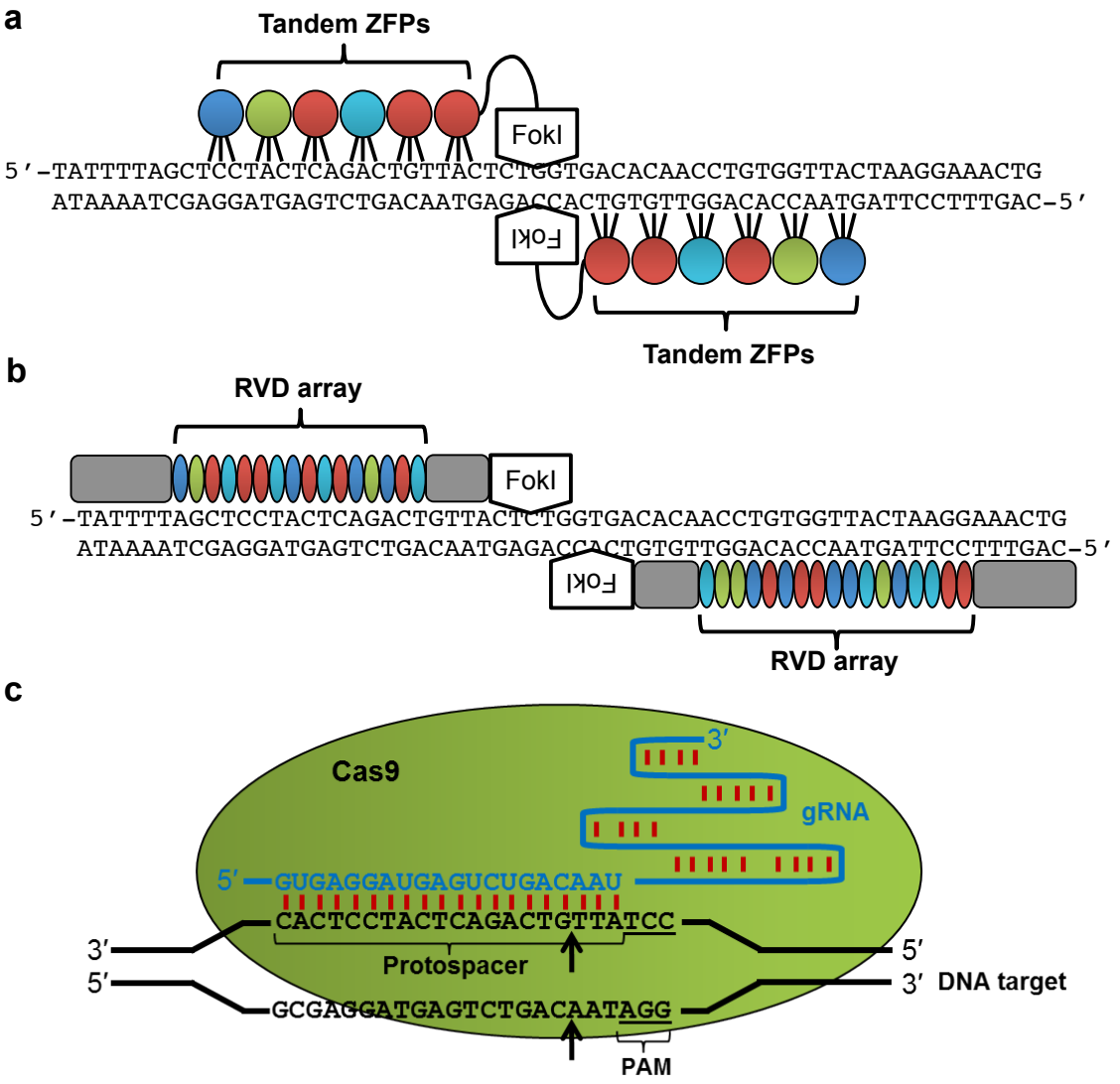


Figure 2: Schematic of nuclease-based genome editing technologies, including (a) ZFNs, (b) TALENs, and (c) CRISPR/Cas9, with the DNA target in black, the gRNA in blue, and the Cas9 nuclease in green.

2.5.2. Chimeric nucleases based on the *Fok*I domain

ZFNs and TALENs are chimeric nucleases that utilize an independent, programmable DNA-binding domain fused to the non-specific catalytic domain of the endonuclease *Fok*I [113]. Site-specific double-strand breaks are created when two

independent nucleases bind to adjacent target DNA sequences, thereby permitting dimerization of *Fok*I and cleavage of the target DNA. Thus, since *Fok*I acts as a dimer, these nucleases are designed in pairs to guide each half of *Fok*I to a desired target site. Several improvements have been made to enhance the specificity of these chimeric nucleases. The first major advance was the identification of mutations that require left and right obligate heterodimerization [114, 115], thereby preventing potential homodimerization of nuclease monomers at unintended target loci. Several groups also created nuclease pairs that simply nick DNA by introducing inactivating mutations to the *Fok*I domain on one of the nucleases in each pair, thereby retaining high recombinogenic activity while drastically reducing error-prone NHEJ repair [116, 117]. Directed evolution of the *Fok*I domain identified highly active variants that significantly increase chimeric nuclease activity in a target site-independent manner [118].

2.5.2.1. Zinc Finger Nucleases (ZFNs)

ZFNs are polydactyl proteins that recognize DNA by linking individual zinc-finger (ZF) motifs, with each motif recognizing 3bp of DNA, in tandem [6] (**Figure 2a**). The zinc-finger DNA binding domains are engineered based on the Cys2-His2 zinc finger domain, the most common DNA-binding motif in the human proteome. The DNA-binding specificity of synthetic zinc finger domains has been extensively altered to recognize almost any DNA target through site-directed mutagenesis and rational design or the selection of large combinatorial libraries [6]. This work led to the establishment of

ZFNs as one of the earliest and most widely used genome editing tools [77, 119-122]. Custom ZFNs can be assembled by a variety of publicly available techniques, including assembly from a predefined library of zinc-finger modules [123-126], selection-based assembly [122, 127], a library of prevalidated zinc-finger arrays [128], and commercial synthesis. Despite the difficulty of engineering new arrays, ZFNs have had success in the clinic [13] and are therefore a promising technology for manipulating the human genome.

2.5.2.2. TALENs

TALENs recognize DNA through a set of linked protein repeats, with the DNA preference of each repeat dictated by a unique repeat variable diresidue (RVD) in each repeat that makes base-specific DNA contacts (**Figure 2b**). The DNA binding domain for TALENs was adapted from the DNA binding domain of a plant pathogen protein and consists of an array of RVD modules, each of which specifically recognizes a single base pair of DNA [97, 98]. RVD modules can be arranged in any order to assemble an array that recognizes a defined sequence, requiring only that each target site be immediately preceded by a 5'-thymine for efficient DNA recognition [113, 129, 130]. The resulting engineered TALE DNA binding domain is then fused to the *Fok*I domain to form a nuclease. Several studies have shown that appropriate truncation of the C-terminus of TALEs enhances nuclease activity when fused to the *Fok*I domain [96, 113, 131]. Custom TALENs can be rapidly created from a relatively small library of plasmids using

publicly available reagents utilizing 'Golden Gate' molecular cloning techniques to assemble new arrays within days [8, 132] or high-throughput methods that utilize solid-phase assembly [9, 133] or ligation-independent cloning techniques [134]. The apparent ease of engineering TALENs to novel targets is attractive for engineering novel molecular therapies for DMD.

2.5.3. CRISPR/Cas9

Recently, a new class of DNA editing enzymes was adapted from an innate bacterial defense system, CRISPR/Cas9, for use in mammalian cells. Unlike ZFNs or TALENs, CRISPR/Cas9 recognizes a target DNA sequence through the RNA-DNA interaction of a guide RNA sequence that tethers to and guides a nuclease protein domain to cleave a predefined DNA sequence (**Figure 2c**). CRISPR/Cas9 has been successfully adapted to work in mammalian cells by co-expression of a Cas9 nuclease that pairs with a guide RNA (gRNA) molecule that targets the Cas9 nuclease to a predefined DNA target matching the 5′ end of the gRNA sequence. The adapted systems consist of a Cas9 nuclease that is co-expressed with a single guide RNA (sgRNA) molecule. The Cas9 nuclease forms a complex with the 3′ end of the sgRNA, and the protein-RNA pair recognizes its genomic target by complementary base pairing between the 5′ end of the sgRNA sequence and a predefined 20 bp DNA sequence, known as the protospacer. By simply exchanging the 20 bp recognition sequence of the expressed sgRNA, the Cas9 nuclease can be directed to new genomic targets. The only

restriction for protospacer targeting is that the sequence must be immediately followed by the protospacer-adjacent motif (PAM), a short sequence recognized by the Cas9 nuclease that is required for DNA cleavage. A unique capability of the CRISPR/Cas9 system is the straightforward ability to simultaneously target multiple distinct genomic loci by co-expressing a single Cas9 protein with two or more sgRNAs [10, 11, 135].

Several orthogonal Cas9 nucleases have been identified from a range of bacterial species, including *S. pyogenes*, *S. thermophilus*, *N. meningitidis* and *T. denticola* [11, 100, 136]. These CRISPR/Cas9 systems operate similarly, but with each utilizing distinct Cas9 with unique PAM recognition sequence and guide RNA molecules [100, 136]. The CRISPR system adapted from *S. pyogenes* has a well-defined the PAM sequence for its Cas9 nuclease (SpCas9) as 5′-NRG-3′, where R is either A or G [137]. The on-target specificity of this gene editing system has been extensively characterized in human cells [137-143], observing positional dependence of mismatches in the protospacer on specificity and unexpected recognition of degenerate PAM sequences. Because nickase activity alone greatly reduces NHEJ mutagenesis [116, 117], target site specificity can be greatly increased by utilizing off-set sgRNAs and CRISPR/Cas9 nickases [142]. In this method, two sgRNAs are designed to bind to and nick opposite strand DNA targets immediately adjacent to each other. The two nicks on opposite strands results in a double-strand break and NHEJ mutagenesis and decreases the likelihood that both sgRNAs will bind opposite strands at off-target loci. Another simple method to increase

sgRNA specificity of the *S. pyogenes* CRISPR/Cas9 system is to truncate the length of the protospacer in sgRNAs [143], presumably reducing the tolerance of mismatches. Collectively, these studies demonstrate that SpCas9 system is an exciting new advance towards specific and highly efficient gene editing applications. However, CRISPR/Cas specificity may yet be further enhanced through novel Cas9 nucleases, longer and more stringent PAM sequences, and/or alterations to sgRNA structure that limit off-target recognition.

2.6. Therapeutic applications of genome editing

Genome editing transitioned from bench to bedside with the entrance of ZFNs into Phase I/II clinical trials for disruption of the HIV-1 co-receptor CCR5 [12, 144]. Additionally, there are ongoing Phase I trials for ZFN-mediated disruption of the glucocorticoid receptor in T cells for glioblastoma treatment, demonstrating the therapeutic potential of this emergent technology. Notably, both of these clinical trials utilize ZFN-mediated genetic disruption by error-prone NHEJ to abolish expression of specific receptors, rather than as a tool for genetic correction. A landmark study demonstrated *in vivo* genome editing to correct mutations causing hemophilia B in a mouse model [81]. The investigators achieved phenotypic correction of hemophilia B in a transgenic mouse model of hemophilia, carrying a disrupted human factor IX gene, after *in vivo* delivery of a ZFN that enabled site-specific integration of a gene targeting construct to replace the factor IX gene *in situ.* This work demonstrated that adeno-

associated virus gene transfer for *in vivo* delivery of ZFNs could be used for both homology-independent (i.e., NHEJ) and homology-directed gene modification. A following study demonstrated that *in vivo* gene editing is efficient in the quiescent livers of adult mice as well [87]. Together, these studies suggest that gene editing in non-dividing tissues, such as skeletal muscle, may be feasible. Other preclinical studies have utilized nucleases to correct several other human genetic mutations associated with sickle cell anemia [78, 79], X-linked SCID [77], and alpha-1-antitrypsin deficiency [80, 145], epidermolysis bullosa [146], xeroderma pigmentosum [147], and mitochondrial DNA disorders [148]. Therefore, we propose that genome editing is a promising technology to address mutations in the human dystrophin gene. Furthermore, this technology has the potential to be delivered *in vivo* to correct dystrophin mutations *in situ* in non-dividing skeletal muscle.

Chapter 3. Reading Frame Correction by Targeted Genome Editing Restores Dystrophin Expression in Cells from Duchenne Muscular Dystrophy Patients

Ousterout DG, Perez-Pinera P, Thakore PI, Kabadi AM, Brown MT, Qin X, Fedrigo O, Mouly V, Tremblay JP, Gersbach CA. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. Molecular Therapy 21, pp. 1718-1726. 2013.

3.1. Synopsis

Genome editing with engineered nucleases has recently emerged as an approach to correct genetic mutations by enhancing homologous recombination with a DNA repair template. However many genetic diseases, such as Duchenne muscular dystrophy, can be treated simply by correcting a disrupted reading frame. We show that

genome editing with TALENs, without a repair template, can efficiently correct the reading frame and restore the expression of a functional dystrophin protein that is mutated in Duchenne muscular dystrophy. TALENs were engineered to mediate highly efficient gene editing at exon 51 of the dystrophin gene. This led to restoration of dystrophin protein expression in cells from Duchenne patients, including skeletal myoblasts and dermal fibroblasts that were reprogrammed to the myogenic lineage by MyoD. Finally, exome sequencing of cells with targeted modifications of the dystrophin locus showed no TALEN-mediated off-target changes to the protein-coding regions of the genome, as predicted by *in silico* target site analysis. This strategy integrates the rapid and robust assembly of active TALENs with an efficient gene editing method for the correction of genetic diseases caused by mutations in non-essential coding regions that cause frameshifts or premature stop codons.

3.2. Introduction

Genome editing with engineered site-specific endonucleases has emerged as a new technology to selectively replace or correct disrupted genes, in contrast to conventional gene therapy methods of gene addition [6, 75]. The recent development of TALENs has dramatically advanced genome editing due to their high rate of successful and efficient genetic modification [8, 9, 75, 95-98, 113, 134, 149-151]. TALENs are engineered fusion proteins of the catalytic domain of the endonuclease *Fok*I and a designed TALE DNA-binding domain that can be targeted to a custom DNA sequence

[113, 149]. The TALE domain consists of an array of repeat variable diresidue (RVD) modules, each of which specifically recognizes a single base pair of DNA [97, 98]. RVD modules can be arranged in any order to assemble an array that recognizes a defined sequence [97, 98]. Site-specific double-strand breaks are created when two independent TALENs bind to adjacent DNA sequences, thereby permitting dimerization of *Fok*I and cleavage of the target DNA [75]. This targeted double-strand break stimulates cellular DNA repair through either homology-directed repair (HDR) or the non-homologous end joining (NHEJ) pathway. HDR uses a donor DNA template to guide repair and can be used to create specific sequence changes to the genome, including the targeted addition of whole genes. In contrast, the template-independent re-ligation of DNA ends by NHEJ is a stochastic, error-prone repair process that introduces random micro-insertions and micro-deletions (indels) at the DNA breakpoint.

Thus far, strategies for the correction of human genes have been based primarily on HDR, the efficiency of which is dependent on cell-cycle state and delivery of an exogenous DNA template [77-81]. In many cases, antibiotic selection is used in tandem with genome editing for gene correction in cell types with low levels of HDR repair [78-80]. Although HDR is extremely valuable for restoring the complete coding sequence of the mutant gene, evidence from studies investigating oligonucleotide-mediated exon skipping and pharmacologic read-through of premature stop codons demonstrates that restoring expression of fully or partially functional truncated proteins can provide

therapeutic benefit for many diseases [4, 5]. Therefore indels resulting from NHEJ-mediated gene repair could restore a disrupted reading frame or remove a premature stop codon and ameliorate the symptoms of these genetic diseases. This approach would result in permanent gene correction, in contrast to pharmacologic approaches that act transiently at the level of mRNA splicing or translation.

NHEJ has been used in human cells to disrupt genes [6, 12] or delete chromosomal segments [91, 92], although it has been proposed that genetic mutations created by endonucleases could be used to restore an aberrant reading frame [89, 90]. In this study, we provide the first example of the restoration of protein expression from an endogenous mutated gene through template-free NHEJ-mediated DNA repair. Duchenne muscular dystrophy (DMD), the most common monogenic hereditary disease, is caused by defects in the gene encoding dystrophin. The majority of dystrophin mutations that cause DMD are deletions of exons that disrupt the reading frame [5]. However, deletion of internal dystrophin exons that retain the proper reading frame causes the less severe Becker muscular dystrophy. This has led to efforts to restore the disrupted dystrophin reading frame in DMD patients by skipping non-essential exons during mRNA splicing, thereby producing internally deleted, but still partially or fully functional, dystrophin proteins [4, 70, 71]. In contrast to a transient method targeting the dystrophin mRNA, the correction of the dystrophin reading frame in the genome by transiently expressed TALENs would lead to permanently restored

dystrophin expression by each modified cell and all of its progeny. Notably, exon 51 is frequently adjacent to frame-disrupting deletions in DMD patients and has been targeted in clinical trials for oligonucleotide-based exon skipping with promising early therapeutic results [4, 70, 71]. An ongoing clinical trial for the exon 51 skipping compound eteplirsen recently reported a significant functional benefit across 48 weeks, with an average of 47% dystrophin positive fibers compared to baseline. Therefore, this class of DMD mutations is ideally suited for permanent correction by NHEJ-based genome editing.

This study investigates template-free gene correction by using TALENs to restore aberrant reading frames through the introduction of indels into exon 51 of the dystrophin gene. Accordingly, we designed and validated an optimized TALEN that targets exon 51. The optimized TALEN was transfected into human DMD cells and shown to mediate efficient gene modification and conversion to the correct reading frame. Furthermore, protein restoration was concomitant with frame restoration and could be detected in a bulk population of TALEN-treated cells. The high specificity of the optimized TALEN was demonstrated by *in silico* analysis, cytotoxicity assays, and exome sequencing of clonally-derived modified cells.

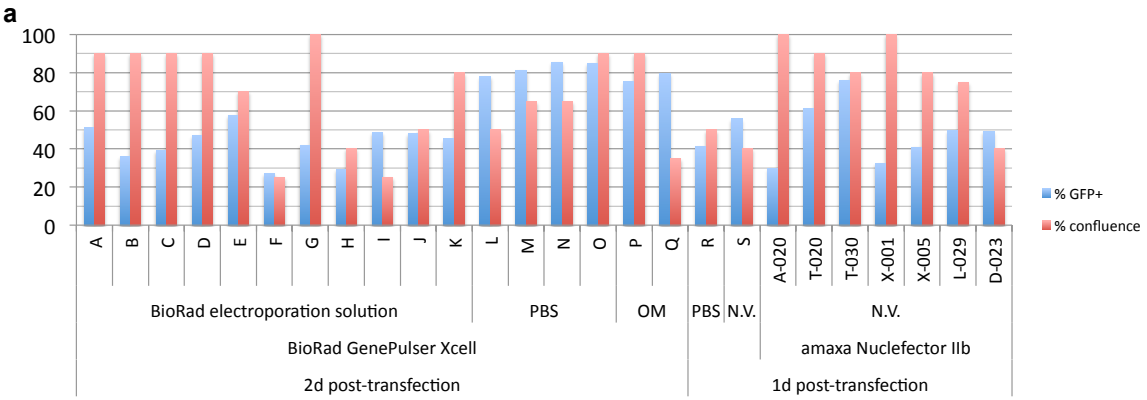
3.3. Materials and Methods

3.3.1. Cell culture and transfection

HEK293T cells were obtained from the American Tissue Collection Center (ATCC) through the Duke Cell Culture Facility and were maintained in DMEM supplemented with 10% bovine calf serum and 1% penicillin/streptomycin. Immortalized myoblasts [152] (one from a wild-type donor, and two Δ48-50 DMD patient derived lines) were maintained in skeletal muscle media (PromoCell) supplemented with 20% bovine calf serum (Sigma), 50 μg/ml fetuin, 10 ng/ml human epidermal growth factor (Sigma), 1 ng/ml human basic fibroblast growth factor (Sigma), 10 μg/ml human insulin (Sigma), 1% GlutaMAX (Invitrogen), and 1% penicillin/streptomycin (Invitrogen). Primary DMD dermal fibroblasts were obtained from the Coriell Cell repository (GM05162A, Δ46-50) and maintained in DMEM supplemented with 10% fetal bovine serum, 1 ng/mL human basic fibroblast growth factor, and 1% penicillin/streptomycin. All cell lines were maintained at 37°C and 5% CO_2. HEK293T cells were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol in 24 well plates. Immortalized myoblasts and primary fibroblasts were transfected by electroporation using the Gene Pulser XCell (BioRad) with PBS as an electroporation buffer using optimized conditions for each line (**Figure 3**). Transfection efficiencies were measured by delivering an EGFP expression plasmid and using flow cytometry. These efficiencies were routinely ≥95% for HEK293T and

≥70% for the primary fibroblasts and immortalized myoblasts. For all experiments, the indicated mass of electroporated plasmid corresponds to the amount used for each TALEN monomer.

a



b

Program	Conditions
A	1000uF, 120V, 1M cells/200uL, 10ug GFP + 10ug empty vector
B	500uF, 120V, 1M cells/200uL, 10ug GFP + 10ug empty vector
C	300uF, 120V, 1M cells/200uL, 10ug GFP + 10ug empty vector
D	500uF, 100V, 1M cells/200uL, 10ug GFP + 10ug empty vector
E	500uF, 150V, 1M cells/200uL, 10ug GFP + 10ug empty vector
F	500uF, 200V, 1M cells/200uL, 10ug GFP + 10ug empty vector
G	1000uF, 100V, 1M cells/200uL, 10ug GFP + 10ug empty vector
H	1000uF, 150V, 1M cells/200uL, 10ug GFP + 10ug empty vector
I	1000uF, 100V, 0.2M cells/200uL, 10ug GFP + 10ug empty vector
J	1000uF, 100V, 0.4M cells/200uL, 10ug GFP + 10ug empty vector
K	1000uF, 100V, 1M cells/200uL, 10ug GFP + 10ug empty vector
L	950uF, 190V, 0.5M cells/150uL, 5ug GFP
M	950uF, 190V, 1M cells/200uL, 10ug GFP
N	950uF, 160V, 0.5M cells/150uL, 5ug GFP
O	950uF, 160V, 1M cells/200uL, 10ug GFP
P	950uF, 220V, 2M cells/300uL, 10ug GFP
Q	950uF, 100V, 1M cells/200uL, 10ug GFP
R	950uF, 190V, 1M cells/200ul, 2ug pmaxGFP
S	950uF, 190V, 1M cells/100uL, 2ug pmaxGFP

Figure 3: Optimization of electroporation conditions for myoblasts. (a) DMD myoblast cells (cell line 1) were electroporated using BioRad Gene Pulser Xcell or amaxa Nucleofector IIb devices using the indicated programs. Several different buffers were tested, including BioRad electroporation solution, Sigma phosphate-buffered saline product #D8537 (PBS), Invitrogen OptiMEM I (OM), or amaxa Nucleofector solution V (N.V.). Conditions using the GenePulser device used infinite resistance. For nucleofection, 1M cells/100 µL nucleofection solution and 2 µg of GFP vector were used according to the manufacturer's specifications. Electroporation using the GenePulser device with program O in PBS solution was selected as the optimal conditions for electroporating myoblasts. (b) Conditions used to optimize BioRad Gene Pulser Xcell electroporation.

3.3.2. TALE nuclease assembly and off-target site prediction

TALENs targeted to exon 51 of the human dystrophin gene were designed *in silico* using the TALE-NT webserver [8]. TALEN target sites were chosen to include half-site targets approximately 15-19 bp in length, preceded by a 5′-T [113]. Plasmids encoding these TALENs were assembled using the Golden Gate assembly method [8] and standard cloning techniques into a modified pcDNA3.1 (Invitrogen) destination vector containing the Δ152/+63 TALEN architecture [113] derived from the pTAL3 expression vector provided in the Golden Gate kit from Addgene. The *Fok*I endonuclease domains were codon optimized and contained the ELD/KKR obligate heterodimer [115] and *Sharkey* mutations [118] as described previously [153]. Potential off-target sites for TALEN pair TN3/8 in the human genome were predicted *in silico* using the Paired Target Finder tool on the TALE-NT 2.0 webserver [154]. All predicted off-target sites were scanned using the following parameters: recommended score cutoff (3.0), spacers of range 12-23 bp, and upstream base set to "T only". Valid likely potential off-target sites were only considered as those with up to 4 mismatches per TALEN half-site binding sequence (maximum of 8 mismatches per TALEN pair target site).

3.3.3. Cel-I quantification of endogenous gene modification

TALEN-induced lesions at the endogenous target site were quantified using the Surveyor nuclease assay, which can detect mutations characteristic of nuclease-mediated

NHEJ. After electroporation, cells were incubated for 3 or 10 days at 37°C and genomic DNA was extracted using the DNeasy Blood and Tissue kit (QIAGEN). The target locus was amplified by 30 cycles of PCR with the AccuPrime High Fidelity PCR kit (Invitrogen) using primers 5′-GAGTTTGGCTCAAATTGTTACTCTT-3′ and 5′-GGGAAATGGTCTAGGAGAGTAAAGT-3′. The resulting PCR products were randomly melted and reannealed in a PCR machine with the program: 95°C for 240 s, followed by 85°C for 60 s, 75°C for 60s, 65°C for 60s, 55°C for 60 s, 45°C for 60 s, 35°C for 60 s, and 25°C for 60s with a -0.3°C/s rate between steps. Following reannealing, 8 μl of PCR product was mixed with 1 μl of Surveyor Nuclease S and 1 μl of Enhancer S (Transgenomic) and incubated at 42°C for 1 hour. After incubation, 6 μl of digestion product was loaded onto a 10% TBE polyacrylamide gel and run at 200V for 30 min. The gels were stained with ethidium bromide and quantified using ImageLab (Bio-Rad) by densitometry as previously described [153].

3.3.4. Cytotoxicity assay

To quantitatively assess potential TALEN cytotoxicity, HEK293T cells were transfected with 10 ng of a GFP reporter and 100 ng of each nuclease using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). The percentage of GFP positive cells was assessed at 2 and 5 days by flow cytometry. The survival rate was calculated as the decrease in GFP positive cells from days 2 to 5 and

normalized to cells transfected with an empty nuclease expression vector as described [155].

3.3.5. Clone isolation procedure

Immortalized DMD myoblasts were electroporated with 10 μg of each TALEN plasmid (20 μg total). After 7 days, isogenic clones were isolated by clonal dilution in hypoxic conditions (5% O_2) to accelerate myoblast growth. Genomic DNA was extracted from clones using the QuickExtract Kit (Epicentre) and the target locus amplified by PCR using the Cel-I primers and conditions above. The resulting PCR products were either mixed with equal amounts of PCR product from untreated cells and analyzed by the Surveyor assay (**Figure 7a**) and/or directly submitted for conventional Sanger sequencing (**Figure 7c**) to identify modified clones.

3.3.6. Viral transduction and forced MyoD overexpression in primary fibroblasts

300,000 fibroblasts were plated transduced in 10 cm plates with a lentiviral vector encoding a full-length human MyoD cDNA under the control of a dox-inducible promoter and a constitutive puromycin resistance cassette. Two days post-transduction, fibroblasts were selected for 6 days in 1 μg/mL puromycin (Sigma) to enrich for transduced cells. Fibroblasts were then plated at a density of 200,000 cells in 10 cm dishes and MyoD expression was induced by adding 3 μg/mL doxycycline (Fisher Scientific) to the media, which was exchanged every two days.

3.3.7. Western blot analysis

To assess dystrophin expression, immortalized myoblasts were differentiated into myofibers by replacing the growth medium with DMEM supplemented with 1% insulin-transferrin-selenium (Invitrogen) and 1% antibiotic/antimycotic (Invitrogen) for 4-7 days. Fibroblasts were transdifferentiated into myoblasts by inducing MyoD overexpression and incubating the cells in DMEM supplemented with 1% insulin-transferrin-selenium (Invitrogen), 1% antibiotic/antimycotic (Invitrogen) and 3 μg/mL doxycycline for 15 days. TALEN expression was assessed at 3 days after transfecting HEK293T cells. Cells were collected and lysed in RIPA buffer (Sigma) supplemented with a protease inhibitor cocktail (Sigma) and the total protein amount was quantified using the bicinchoninic acid assay according to the manufacturer's instructions (Pierce). Samples were then mixed with NuPAGE loading buffer (Invitrogen) and 5% β-mercaptoethanol and heated to 85°C for 10 minutes. Twenty-five micrograms of protein were separated on 4-12% NuPAGE Bis-Tris gels (Invitrogen) with MES buffer (Invitrogen). Proteins were transferred to nitrocellulose membranes for 1-2 hours in transfer buffer containing 10-20% methanol and 0.01% SDS. The blot was then blocked for 1 hour with 5% milk-TBST at room temperature. Blots were probed with the following primary antibodies: NCL-Dys2 (1:25, Leica), MANDYS8 (1:100, Sigma), GAPDH (1:5000, Cell Signaling), anti-FLAG-HRP (1:2000, Cell Signaling), or anti-myogenin F5D (1:200, Santa Cruz). Dystrophin expression was detected using

MANDYS8 in DMD myoblast line 2 and the DMD fibroblast line or NCL-Dys2 in DMD myoblast line 1. TALEN expression was detected using anti-FLAG. Blots were then incubated with mouse or rabbit horseradish peroxidase-conjugated secondary antibodies (Santa Cruz) and visualized using the ChemiDoc chemilumescent system (BioRad) and Western-C ECL substrate (BioRad).

3.3.8. Immunofluorescence

Fibroblasts were plated on cover slips in 24 well plates at a density of 30,000 cells/well and MyoD expression was induced for 15 days as described above. Cells were then fixed in 4% paraformaldehyde and blocked for 1 hour at room temperature with PBS containing 5% BSA, 2% goat serum and 0.2% Triton X-100. Cells were then stained overnight at 4°C with MF20 (1:200, Developmental Studies Hybridoma Bank) primary antibody and then for 1 hour at room temperature with anti-mouse AlexaFluor 488 (Molecular Probes) secondary antibody. Cover slips were mounted with ProLong Gold antifade (Molecular Probes).

3.3.9. Exome sequencing and analysis

We analyzed the exomes of four clonally derived DMD myoblast lines carrying known TALEN-mediated deletions in exon 51 of the dystrophin gene, as well as the parent line for these cells. Genomic DNA was isolated using the DNeasy Blood and Tissue Kit (QIAGEN) and 3 μg of DNA were submitted to the Duke Institute for Genome Sciences and Policy's Genome Sequencing & Analysis Core. Illumina-

compatible libraries were made and enriched for exonic regions using the SureSelect Human All Exon V4 Kit (Agilent). Five total libraries were prepared from the four treatment samples and one parental line reference sample. The libraries were indexed and sequenced on one lane of Illumina HiSeq2000 (100-bp paired-end sequencing). Bioinformatics analyses were performed by Duke Genome Sequencing & Analysis Core. The analysis pipeline includes the initial QC to remove sequencing adaptors and low quality bases to facilitate mapping. Sequence depth of targeted regions was calculated as >97% at 10x coverage, >91% at 20x coverage, and >82% for 30x coverage (**Appendix A**). Each sequencing reaction generated >64 million reads with >93% of reads above a quality score of 30 and an overall mean quality score of >36.4. High quality reads were mapped to the human reference genome (hg19) using bwa 0.5.9. An exome capture pipeline developed at the Duke Sequencing Core was used to assess the exome capture efficiency. Picard v1.74 is used for removing PCR duplicates. The GATK (v1.6-13) toolkit is used for variant calling, read realignment around INDELs, quality score recalibration and QC filtering. The filtering step discards the variants with 1) low coverage (coverage (<30x), 2) strand-bias, 3) low SNP quality score (< 50) and 4) low allelic frequency (<0.5). Each candidate point mutation or INDEL were reviewed manually by IGV to identify false negative artifacts due to insufficient coverage of the parental line. Identical point mutations and INDELs that occurred in more than two of the four clones were verified as artifacts due to coverage of the reference parent cell line and were discarded.

Common point mutations and INDELs were removed by comparing to human dbSNP135. The remaining point mutations and INDELs were annotated using Annovar (May 25, 2012 version) and classified using a perl script written by the Duke Sequencing Core. The non-exonic point mutations were not considered. All point mutations and INDELs were individually visualized and validated on IGV. The flanking 100bp of each validated mutation was screened for any potential sequence similarity to the TN3/8 target site using the Paired Target Finder tool on the TALE-NT 2.0 webserver [154] using the parameters: recommended score cutoff (3.0), spacers of range 1-30bp, and upstream base set to "T only".

3.4. Results

3.4.1. Design and validation of TALENs targeted to the dystrophin gene

To evaluate TALEN-mediated genetic correction by NHEJ, several TALENs were designed to target exon 51 in the dystrophin gene. TALEN target sites were chosen immediately upstream of the two possible out-of-frame stop codons (**Figure 4**), such that insertions or deletions could restore the dystrophin reading frame in either disrupted frame. Variable lengths of spacers between TALEN monomers and TALEN RVD array lengths were tested to optimize nuclease activity (**Appendix A**), as done previously [113]. Western blots confirmed full-length and robust expression of the TALENs following transfection of TALEN-encoding plasmids into HEK293T cells (**Figure 5a**). All combinations of left and right TALENs were then transfected into HEK293T cells and

the genomic DNA was assessed for modification by the Surveyor assay, which can detect the frequency of allelic modifications with a dynamic range of ~1-50%. Several TALENs with spacers of 14-19 bp were highly active with gene editing efficiencies exceeding modification of 10% of total alleles (**Figure 5b-d**), consistent with previous observations [9, 96, 113]. The gene editing frequencies were stable from day 3 to day 10 (**Figure 5b, e**), confirming that these TALENs are well tolerated in human cells [9, 96, 113]. Furthermore, the engineered TALENs showed minimal cytotoxicity in human cells similar to the well-characterized non-cytotoxic homing endonuclease I-SceI (**Figure 5f, Figure 6**) [96, 155]. TN3/8 was the most highly active and well-tolerated TALEN pair and therefore was used for subsequent experiments.

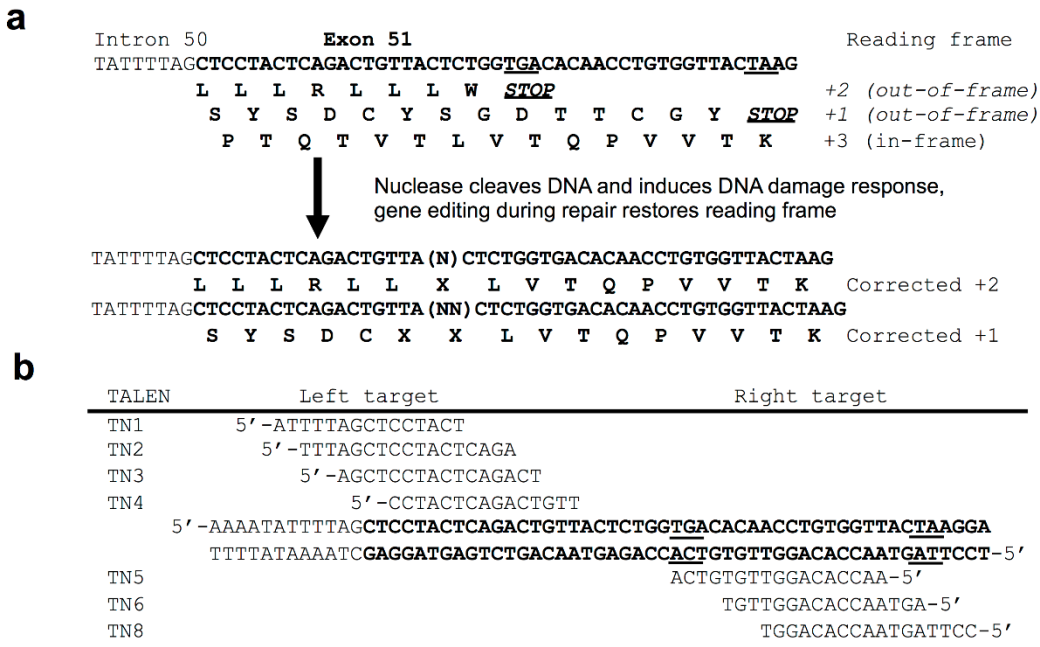
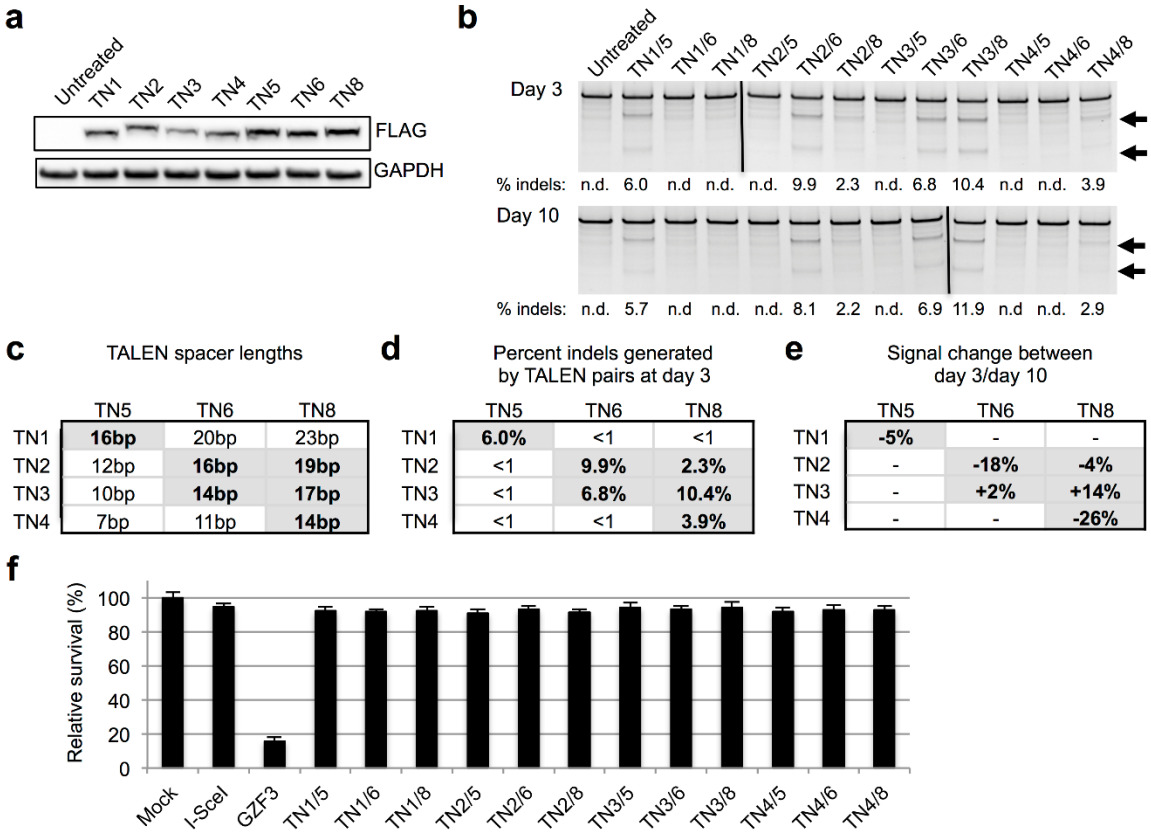


Figure 4: Design of TALENs targeted to exon 51 of the human dystrophin gene. (a) The possible reading frames of human dystrophin exon 51 and expected amino

acid sequences after genome editing. (b) Combinations of TALEN pairs were designed to target immediately upstream of either out-of-frame stop codon (underline) in exon 51 (bold) of the human dystrophin gene.



c	TN5	TN6	TN8
TN1	**16bp**	20bp	23bp
TN2	12bp	**16bp**	**19bp**
TN3	10bp	**14bp**	**17bp**
TN4	7bp	11bp	**14bp**

d	TN5	TN6	TN8
TN1	**6.0%**	<1	<1
TN2	<1	**9.9%**	**2.3%**
TN3	<1	**6.8%**	**10.4%**
TN4	<1	<1	**3.9%**

e	TN5	TN6	TN8
TN1	**-5%**	-	-
TN2	-	**-18%**	**-4%**
TN3	-	**+2%**	**+14%**
TN4	-	-	**-26%**

Figure 5: Validation and characterization of TALENs. (a) Each TALEN construct was transfected independently into HEK293T cells to confirm full-length expression. All TALENs were the expected size of ~95-110kDa. (b) Combinations of TALENs were co-transfected into HEK293T cells to screen for highly active TALEN pairs. Gene modification frequency was monitored at day 3 and day 10 to assess stable gene modification. Arrows denote expected cleavage band sizes indicative of NHEJ activity. (c) Summary of TALEN spacer lengths. (d) Measured gene modification rates detected by the Surveyor assay from day 3 data in (b). (e) Measured indel signal changes between day 3 and day 10 from the data in (b). (f) Cytotoxicity assay in HEK293T cells for all TALEN combinations. I-SceI is a non-toxic meganuclease and GZF3 is a zinc-finger nuclease known to be cytotoxic to human cells. n.d., not detected.

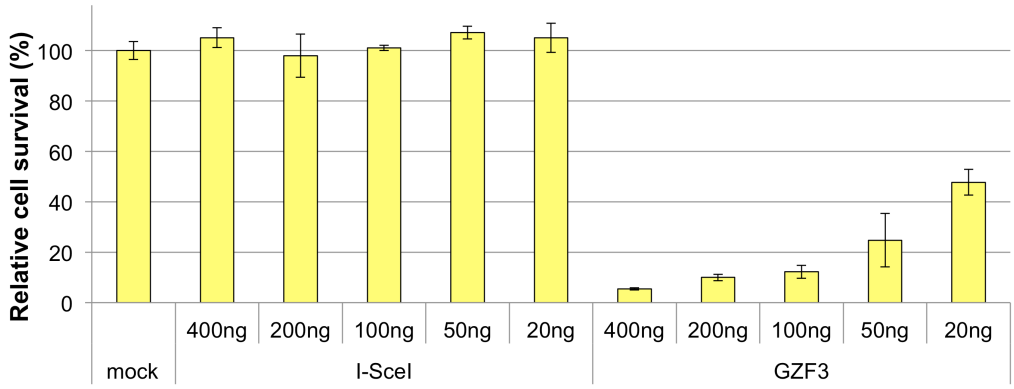


Figure 6: Optimization of cytotoxicity assay using Lipofectamine 2000 in 293T cells. Varying amounts of the non-toxic endonuclease I-SceI and toxic zinc-finger nuclease GZF3 were transfected into 293T cells and assessed for relative survival rates post-transfection. Based on these data, 100 ng of nuclease was used for the cytotoxicity studies.

3.4.2. TN3/8 mediates high efficiency conversion to all three reading frames

NHEJ-based gene modification is expected to create indels of random length and therefore should cause conversion to any of the three reading frames in an exonic sequence. In order to validate the overall gene modification rate and possible reading frames generated following TALEN-induced NHEJ, clonal cell populations were derived from human skeletal myoblasts that had been electroporated with TN3/8-encoding plasmids. These clones were assayed for NHEJ events occurring at the dystrophin exon 51 locus using the Surveyor assay to detect sequence differences relative to untreated cells (**Figure 7a**). Eleven of twenty-eight (39%) clonal cell populations were modified and subsequent sequencing of the alleles from these clones

confirmed indels characteristic of NHEJ (**Figure 7b**). Similar to other studies with TALENs, deletions were heavily favored [156]. The random length of these indels verifies that conversion to any of the three reading frames is possible. The conversion rate to any one of the three reading frames was observed to be roughly proportional to the expected 1/3 of the total NHEJ events (**Figure 7b).** Interestingly, several small deletions were observed that did not alter the original reading frame, demonstrating that this approach could be used to delete aberrant stop-codons (**Figure 7b**).

3.4.3. Reading frame correction leads to restored protein expression

We next assessed whether correction of the dystrophin reading frame by TALEN-mediated NHEJ results in restored dystrophin protein expression. Immortalized human myoblasts derived from DMD patients with a frame-disrupted dystrophin gene caused by deletion of exons 48-50 (Δ48-50) were electroporated with plasmids encoding TN3/8. Clonal cell populations were isolated and screened by PCR amplification of genomic DNA and Sanger sequencing to identify indels characteristic of NHEJ. In this experiment, approximately 5% of clones contained modifications in exon 51, including one clone with an NHEJ event expected to correct the dystrophin reading frame (**Figures 7c, 8**). Following myogenic differentiation, restored dystrophin protein expression was detected by western blot at its predicted size (~412 kDa) only in the corrected clone, and not in clones with non-corrective NHEJ events (**Figure 7d**). These data demonstrate that

NHEJ events that restore the dystrophin reading frame also rescue dystrophin protein expression.

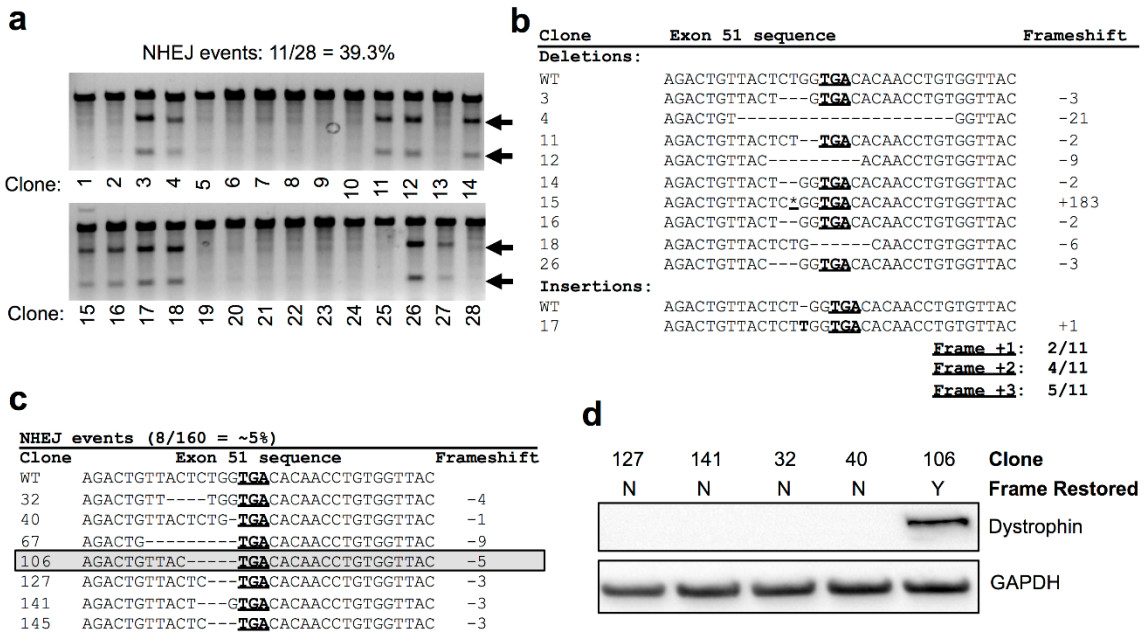


Figure 7: Genetic correction of aberrant dystrophin reading frames by TALEN-mediated genome editing. (a) Isogenic clones were derived from human skeletal myoblasts treated with ten micrograms of each plasmid encoding TN3/8 and screened using the Surveyor assay to detect mutant alleles in reference to the parent (untreated) genomic DNA . Arrows denote expected cleavage band sizes indicative of NHEJ activity. (b) Sanger sequencing of the TALEN target site in exon 51 in mutant clones identified in (a). (c) DMD human myoblast cell line 1 was treated with ten micrograms of each plasmid encoding the TN3/8 TALEN pair and isogenic clones were subsequently derived. Sanger sequencing was used to identify clones with small insertion or deletion mutations at the exon 51 genomic locus characteristic of NHEJ. Clone 106 had a 5 bp deletion expected to restore the reading frame (boxed). All other clones had deletions that were not expected to result in corrective frameshift events. (d) Clonal cell populations with NHEJ events detected at exon 51 were cultured in differentiation conditions for 7 days and analyzed by western blot for dystrophin expression at the expected molecular weight (412 kDa).

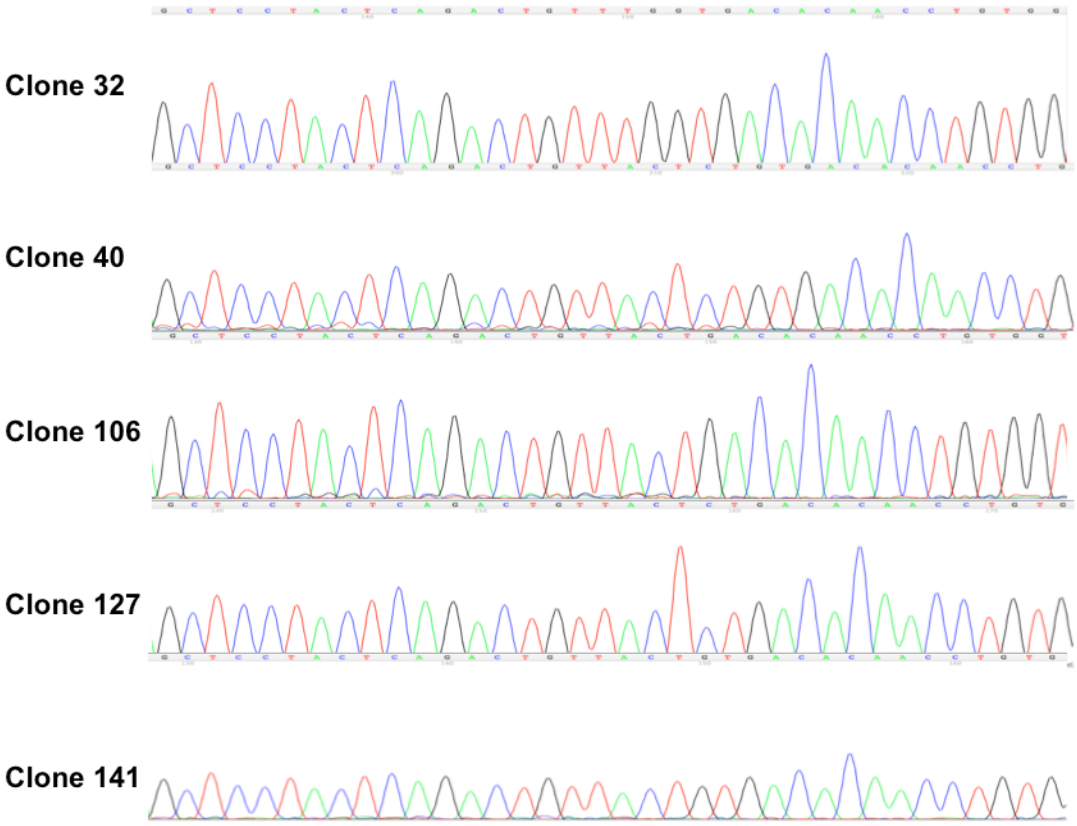


Figure 8: Chromatograms of clones from Figure 7.

3.4.4. TALEN-mediated genetic correction in bulk-treated DMD myoblasts

Efficient *in situ* frame correction in the absence of selection is a powerful use of NHEJ-based gene correction. Accordingly, we investigated the restoration of dystrophin expression in TALEN-treated bulk populations of DMD myoblast lines derived from two different patients containing different deletions of exons 48-50 in the dystrophin gene. As expected, the frequency of gene modification increased with the dose of electroporated TN3/8-encoding plasmids with indels detected in up to 12.7% and 6.8% of alleles, in the two patient lines as measured by the Surveyor assay (**Figures 9a, b**).

Following 7 days of myogenic differentiation induced by serum removal, restored dystrophin expression was detected in the bulk cell populations at the predicted size (~412 kDa) relative to expression from wild-type cells (427 kDa) (**Figures 9c, d**). The increase in dystrophin protein expression with TALEN dose was concomitant with the level NHEJ events detected by the Surveyor assay.

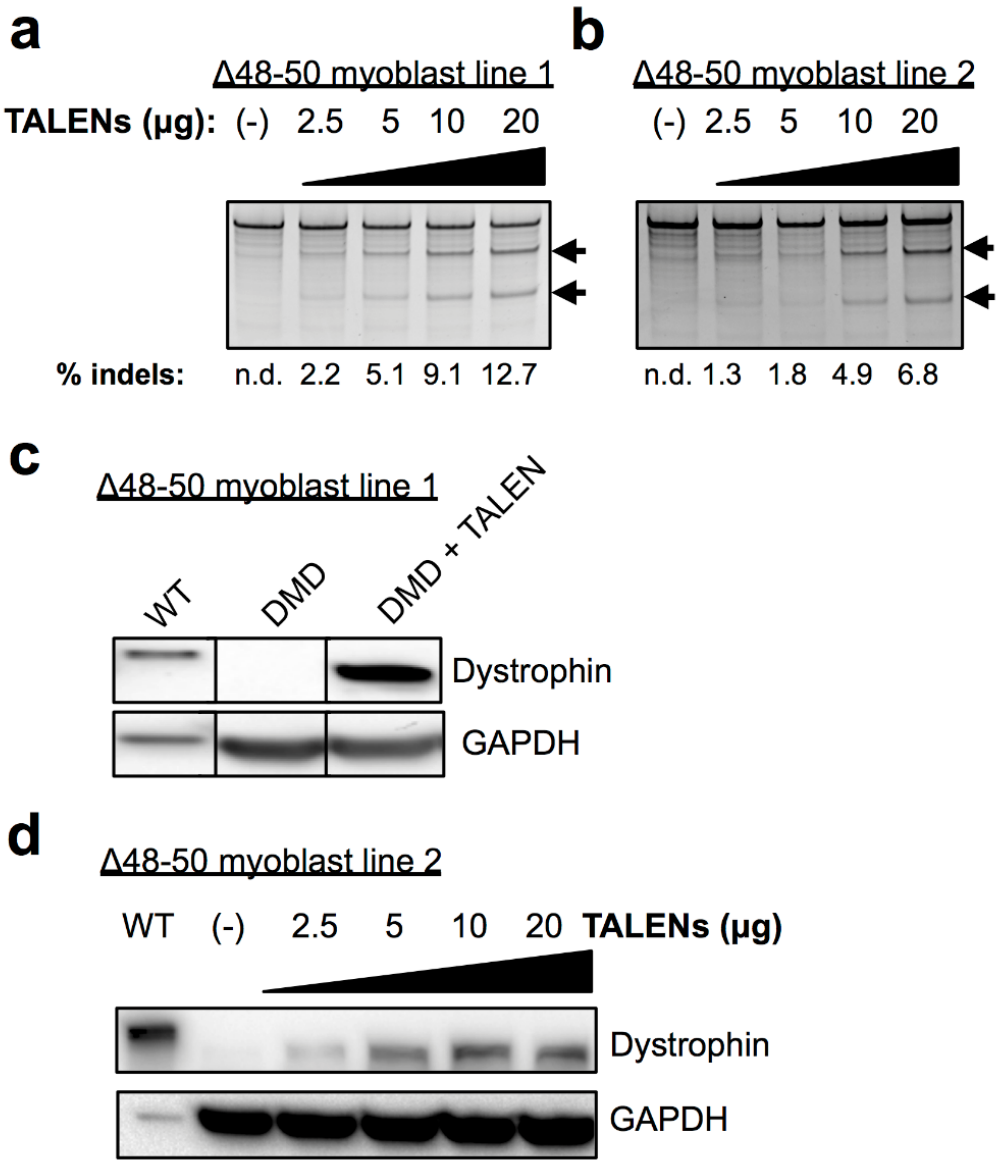


Figure 9: Efficient genetic modification and protein restoration in a bulk population of cells treated with TN3/8. (a,b) Dose-dependent response of NHEJ activity with increasing amounts of TALEN pair TN3/8 measured by the Surveyor assay after transfection of the indicated amount of each TALEN plasmid into two different DMD myoblast lines, each carrying a novel deletion of exons 48-50 (Δ48-50). Arrows denote expected cleavage band sizes indicative of NHEJ activity. (c) DMD myoblast line 1 was treated with five micrograms of each TALEN plasmid and dystrophin expression was assessed after 7 days of differentiation by western blot using the NCL-Dys2 antibody. (d) DMD myoblast 2 was treated with the indicated

amount of each TALEN plasmid and dystrophin expression was assessed after 7 days of differentiation by western blot using the MANDYS8 antibody. Protein from wild-type human myoblasts differentiated in parallel was diluted 1:100 and loaded as a positive control for full-length dystrophin expression (427 kDa) relative to the truncated Δ48-50 product (412 kDa).

3.4.5. Gene restoration in primary DMD dermal fibroblasts

The simplicity of this NHEJ-based approach can enable efficient correction in proliferation-limited primary cell lines that may not be amenable to homologous recombination or selection-based gene correction. For example, DMD patient-derived primary dermal fibroblasts carrying a frame-disrupting deletion of exons 46-50 (Δ46-50) were electroporated with plasmids encoding TN3/8, resulting in high frequency gene modification in a dose-dependent manner (**Figure 10a**). These treated fibroblasts were then transduced with a lentivirus expressing MyoD under an inducible promoter to stimulate transdifferentiation into the myogenic lineage and dystrophin expression [157, 158]. Expression of myogenin (**Figure 10b**) and myosin heavy chain (**Figure 10c**) confirmed efficient transdifferentiation of wild type and DMD patient fibroblasts. Rescued dystrophin expression was detected in TALEN-treated MyoD-induced fibroblasts in a dose-dependent manner at the predicted size of approximately 400 kDa (**Figure 10b**), similar to the results obtained in skeletal myoblasts (**Figure 9c, d**).

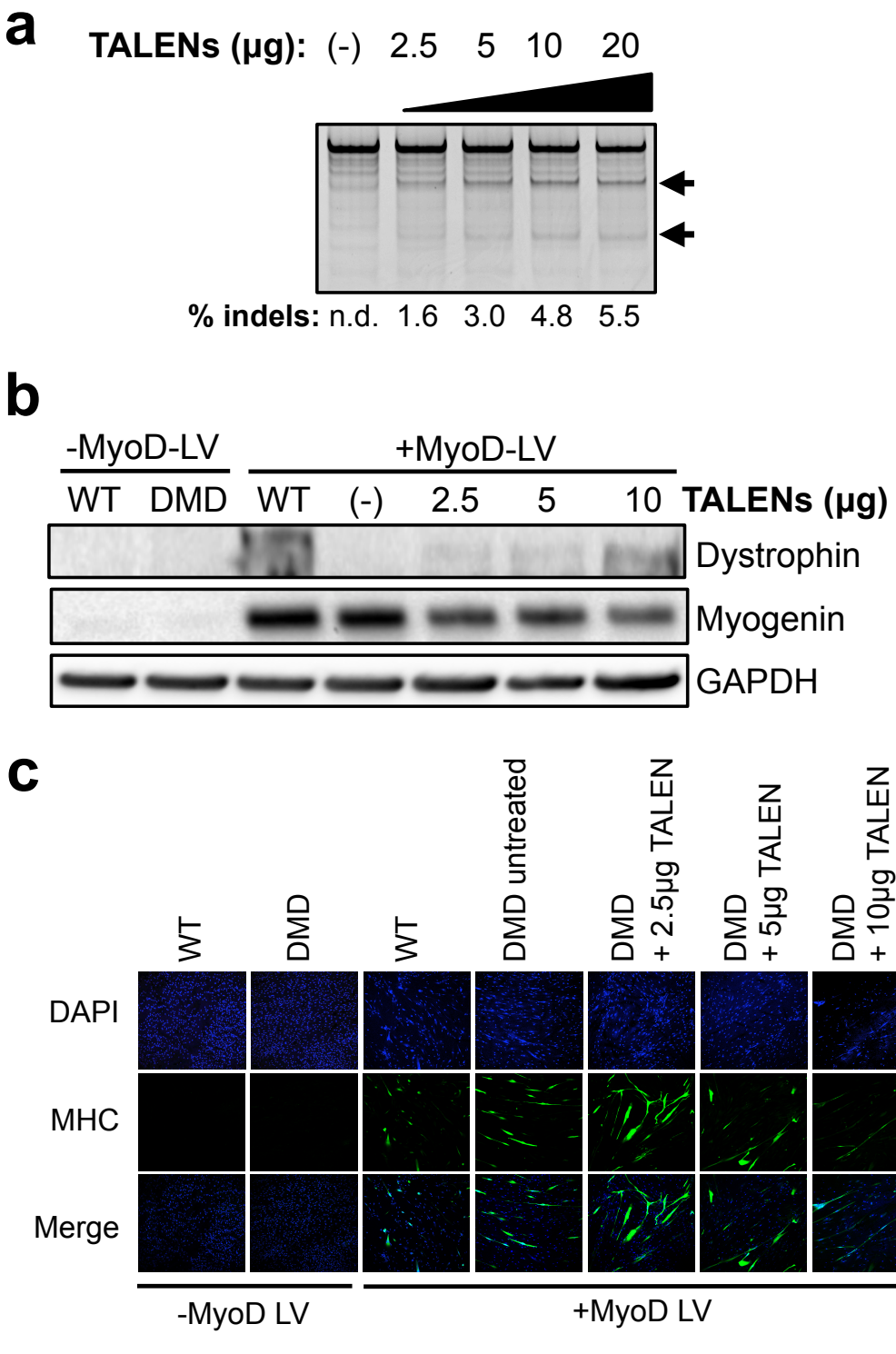


Figure 10: Dystrophin reading frame restoration in primary dermal fibroblasts. **(a) Primary DMD fibroblasts carrying a deletion of exons 46-50 (Δ46-50) were**

electroporated with increasing doses of the indicated amount of each TALEN plasmid and gene modification rates were quantified with the Surveyor assay. Arrows denote expected cleavage band sizes indicative of NHEJ activity. (b) Analysis of myogenin and dystrophin expression (MANDYS8) in wild-type and DMD fibroblasts after treatment with TN3/8 and 15 days of forced MyoD expression. Protein from wild-type dermal fibroblasts is included as a positive control for full-length dystrophin expression (427 kDa) relative to the truncated Δ46-50 product (400 kDa). (c) Immunofluorescence staining to detect myosin heavy-chain (MHC) after MyoD expression by lentiviral gene transfer.

3.4.6. Analysis of off-target effects induced by TN3/8

An important concern for all genome editing strategies is the potential for off-target gene modification events. TN3/8 does not show significant cytotoxicity and is well tolerated by human cells (**Figures 5b, e, f**), suggesting specific gene targeting**.** Potential off-target sites were assessed *in silico* using the TALE-NT 2.0 Paired Target Finder Prediction webserver [154] to scan the human genome for sequences containing up to 4 mismatches per TALEN half-site (up to 8 total mismatches per target site) separated by spacers of any length between 12 and 23 bases. Importantly, this analysis did not produce any potential off-target sites that met these criteria. To further examine unpredicted off-target DNA modifications, we sequenced the whole exomes of clonally derived DMD myoblasts that we had previously confirmed to contain NHEJ events at the on-target exon 51 locus (**Figure 7c, d**). Notably, the only insertion or deletion events characteristic of NHEJ were detected at the on-target exon 51 locus of the dystrophin gene in all four clonal lines analyzed, confirming the specificity of these TALENs (**Table 1, Appendix A**). Consistent with known genomic mutation rates that normally occur during clonal expansion, the exome sequencing revealed several single nucleotide

variants (SNVs) in each clone relative to the parental cell line. Using the TALE-NT 2.0 Paired Target Site Prediction webserver [154], the immediate region around each mutation was scanned for any sequence similarity to the TN3/8 target site to determine if the TALENs could be responsible for the observed SNVs. No target sites with similarity to our TALEN target site with spacers of 1-30 bases were found in the flanking 100 bp of any SNV. Because NHEJ-mediated mutagenesis rarely results in substitutions relative to indels, the detected SNVs are likely to have arisen during clonal expansion as observed in other studies [80, 150]. In summary, there was no apparent off-target activity related to TALEN-mediated, NHEJ-based genetic correction in these clonally derived cells.

Table 1: Summary of clonal sequence variants detected by exome sequencing.

Clone	Mutation Type	Category	AA From	AA To	Gene	Chr	Location	Ref Base	Mutant Base
32	Transition	synonymous SNV	S	S	ANKS1B	12	99201691	C	T
	Transversion	nonsynonymous SNV	F	L	ZNF836	19	52659835	G	C
	Transition	synonymous SNV	P	P	SASH1	6	148664242	T	C
	Transversion	synonymous SNV	L	L	DAXX	6	33287597	T	G
	Transversion	synonymous SNV	L	L	CDH7	18	63525175	T	A
	Deletion	frameshift	-	-	DMD	X	31792285	ACCAG	-
106	Transition	nonsynonymous SNV	E	G	ENG	9	130582267	T	C
	Transition	nonsynonymous SNV	N	D	CCDC36	3	49294344	A	G
	Transition	synonymous SNV	V	V	TARBP1	1	234556520	C	T
	Transition	nonsynonymous SNV	Q	R	UGT3A1	5	35988575	T	C
	Transversion	nonsynonymous SNV	L	I	SOWAHB	4	77817679	G	T
	Transversion	nonsynonymous SNV	Q	P	MEF2A	15	100252738	A	C
	Transversion	nonsynonymous SNV	R	L	RFC1	4	39306505	C	A
	Transition	stopgain SNV	Q	X	ELN	7	73474508	C	T
	Deletion	frameshift	-	-	DMD	X	31792285	ACCAG	-
127	Transition	nonsynonymous SNV	A	V	PLEKHH1	14	68041071	C	T
	Transition	synonymous SNV	P	P	RASAL2	1	178269222	C	T
	Transversion	nonsynonymous SNV	S	C	IGDCC4	15	65676357	G	C
	Transition	synonymous SNV	A	A	LMTK3	19	49001482	G	A
	Transition	stopgain SNV	W	X	PLEKHS1	10	115526378	G	A
	Transition	nonsynonymous SNV	V	I	FAM110C	2	45848	C	T
	Transition	nonsynonymous SNV	G	E	TRAK1	3	42251610	G	A
	Transversion	synonymous SNV	S	S	C15orf39	15	75499997	A	T
	Transition	synonymous SNV	L	L	GPBAR1	2	219127549	C	T
	Deletion	nonframeshift	-	-	DMD	X	31792284	CAC	-
141	Transition	synonymous SNV	L	L	MUC16	19	8999474	T	C
	Transversion	nonsynonymous SNV	F	C	RP1	8	55538286	T	G
	Transversion	nonsynonymous SNV	E	A	PPP1R10	6	30569808	C	G
	Transition	synonymous SNV	T	T	CAMKV	3	49896829	T	C
	Transition	synonymous SNV	F	F	AK2	1	33478842	G	A
	Deletion	nonframeshift	-	-	DMD	X	31792287	CAG	-

3.5. Discussion

NHEJ-exclusive gene correction offers several potential advantages over the HDR pathway. For example, NHEJ does not require a donor template, which may cause nonspecific insertional mutagenesis. In contrast to HDR, NHEJ operates efficiently in all stages of the cell cycle and therefore can be effectively exploited in both cycling and post-mitotic cells, such as muscle fibers. This provides a robust, permanent gene restoration alternative to oligonucleotide-based exon skipping [4] or pharmacologic

forced read-through of stop codons [73] and could theoretically require as few as one drug treatment. NHEJ-based gene correction using TALENs, as well as other engineered nucleases including meganucleases [89] and zinc finger nucleases [159], is also readily combined with other existing *ex vivo* and *in vivo* platforms for cell- and gene-based therapies, in addition to the plasmid electroporation approach described here. For example, delivery of TALENs by mRNA-based gene transfer or as purified cell-permeable proteins [40] could enable a DNA-free genome editing approach that would circumvent any possibility of insertional mutagenesis.

Any of these delivery methods could be utilized with a myriad of cell types currently under investigation for cell-based therapies [36], including induced pluripotent stem cells [29, 32], bone marrow-derived progenitors [33], skeletal muscle progenitors [34], CD133+ cells [35], mesoangioblasts [30], and dermal fibroblasts [36]. Additionally, advances in immortalization of human myogenic cells may greatly simplify clonal derivation of genetically corrected myogenic cells [37]. Significantly, we modified cells *ex vivo* and isolated and expanded clonal populations of immortalized DMD myoblasts that contained a genetically corrected dystrophin gene and were free of nuclease-introduced mutations in protein coding regions of the genome. Alternatively, transient *in vivo* delivery of nucleases by non-viral or non-integrating viral gene transfer [5, 48, 81] or by direct delivery of purified proteins [40] containing cell-penetrating

motifs may enable highly specific correction *in situ* with minimal or no risk of exogenous DNA integration.

Future studies are warranted to investigate the therapeutic efficacy of this approach and similar permanent gene editing strategies to correct endogenous genes. For example, rescuing dystrophin expression to produce these "Becker-like" proteins theoretically introduces novel epitopes in the restored C-terminus. Therefore it will be important to consider potential immune responses following permanent genetic correction of the reading frame [160], though current exon skipping clinical studies suggest a minimal immune response to the restored native gene product [70, 71]. Any reduced function of restored, but truncated, protein products is another potential hurdle to this strategy. In the case of DMD, naturally occurring mutations and their consequences are relatively well understood. It is known that in-frame deletions that occur in the exon 45-55 region contained within the rod domain can produce highly functional dystrophin proteins, and many carriers are asymptomatic or display mild symptoms [4]. Theoretically, greater than 60% of patients can be treated by targeting exons in this region of the dystrophin gene [161]. Collectively, these previous studies indicate that the restored dystrophin proteins created by our approach will be highly functional and alleviate disease symptoms when expressed in skeletal muscle tissue.

Genome editing is a powerful approach for creating custom alterations to the genome, as evidenced by the recent entrance of zinc finger nucleases into clinical trials

for disruption of the HIV-1 co-receptor CCR5 [12] and disruption of the glucocorticoid receptor in T cells for glioblastoma treatment. This study utilizes NHEJ-based genome editing to restore the reading frame of the dystrophin gene in patient cells, in contrast to the conventional use of NHEJ for gene knockout. Given the numerous high-throughput methods to engineer new TALENs [8, 9, 134], as well as their apparent lack of cytotoxicity [9, 96], it should be possible to rapidly extrapolate this NHEJ correction method to other regions of the dystrophin gene as well as other diseases that are caused by a loss of protein function introduced by intragenic insertions, deletions, or aberrant stop codons in non-essential regions, including collagen type VII-associated dystrophic epidermolysis bullosa, Fukuyama congenital muscular dystrophy, and limb-girdle muscular dystrophy type 2B. However, the resulting functionality of these proteins following partial gene correction remains to be determined, particularly for gene deletions and associated phenotypes that are not as well-defined as the region targeted in this study. Therefore HDR-based genome editing for complete restoration of gene deletions is also a valuable approach to pursue in parallel. Nevertheless, NHEJ-based gene correction may provide a versatile therapy for DMD when frame restoration is predicted to permanently correct the native gene and restore protein function.

Chapter 4: Gene Correction of Duchenne Muscular Dystrophy by Genomic Excision of Exon 51 using Zinc-Finger Nucleases

Original article co-authored with Ami M. Kabadi, Pratiksha I. Thakore, Pablo Perez-Pinera, Matthew T. Brown, and Charles A. Gersbach

4.1. Synopsis

Exon skipping using oligonucleotides has been shown to be an exciting method to restore the dystrophin reading frame and restore dystrophin protein production. However, these methods may require repeated administration for the lifetime of the patient and may generate incomplete skipping of the targeted exon. In this study, we apply recent advances in genome editing to permanently exclude exons by using zinc-finger nucleases (ZFNs) to selectively remove sequences important in specific exon recognition. ZFNs were designed to remove essential splicing sequences in exon 51 of the dystrophin gene and thereby exclude exon 51 from the resulting dystrophin transcript, a method that can potentially restore the dystrophin reading frame in up to 13% of DMD patient deletions. Nucleases were assembled by extended modular assembly and context-dependent assembly methods and screened for activity in human cells. Two active ZFN pairs flanking the exon 51 splice acceptor site were transfected into DMD patient cells and a clonal population was isolated with this region deleted

from the genome. Deletion of the genomic sequence containing the splice acceptor resulted in the loss of exon 51 from the dystrophin mRNA transcript and restoration of dystrophin expression *in vitro*. Furthermore, transplantation of corrected cells into the hind limb of immunodeficient mice resulted in efficient human dystrophin expression localized to the sarcolemmal membrane. Finally, the toxicity of selected ZFNs was characterized by measuring cytotoxicity in human cells and by quantifying off-target mutagenesis at predicted chromosomal sites. This study demonstrates a powerful method to correct the dystrophin reading frame by permanently deleting exons to restore dystrophin protein expression.

4.2. Introduction

Designer enzymes have rapidly enabled the precise manipulation of genomic sequences of interest in complex genomes [6, 162]. The rapid development of designer enzymes such as ZFNs [6], TALENs [75], and the more recently described RNA-guided CRISPR/Cas9 system [10] has enabled the possibility of genomic therapy. Nuclease-mediated gene editing strategies facilitate site-specific changes to a target genome by creating double strand breaks that stimulate cellular DNA repair pathways. These pathways result either in error-prone DNA repair through non-homologous end-joining in the absence of a donor DNA molecule or in specific changes guided by homology directed repair when co-delivered with a repair template. Genome editing has been

demonstrated to be a powerful method to study and/or correct monogenic mutations associated with hereditary disease [31, 42, 77, 81, 150, 163, 164].

The severe X-linked hereditary disease Duchenne muscular dystrophy is caused by mutations in the dystrophin gene [2] that prematurely truncate this essential musculoskeletal protein. The loss of functional dystrophin expression causes progressive muscle wasting leading to death by the third decade of life in these patients. Oligonucleotide-based exon skipping is a powerful method to exclude specific exons and has been exploited to restore the dystrophin gene around frame-disrupting deletions adjacent to exon 51, thereby addressing potentially up to 13% of all DMD patient deletions [4, 23, 24]. This transient restoration requires regular administration of the exon skipping drug for the life of the patient. In contrast to transient methods, genome editing creates stable changes to a gene in a modified cell that persists even after cell division. Targeted frameshifts using site-specific nucleases has been demonstrated to be a promising method to correct the dystrophin gene [42, 90, 163]. However, the introduction of random small insertions and deletions in the dystrophin gene result in heterogeneous changes to the final protein product that may impact the predictability and reliability of the resulting protein function. Thus, it would be advantageous to explore a gene correction method that results in an expected protein product with predictable functionality.

Zinc-finger nucleases (ZFNs) are a widely studied tool to create targeted genetic modifications. ZFNs are polydactyl proteins that recognize DNA by linking individual zinc-finger (ZF) motifs, with each motif recognizing 3 bp of DNA, in tandem [6]. Site-specific double-strand breaks are created when two independent ZFN monomers bind to adjacent target DNA sequences, thereby permitting dimerization of *Fok*I and cleavage of the target DNA. Thus, since *Fok*I acts as a dimer, these nucleases are designed in pairs to guide each half of *Fok*I to a desired target site. Several improvements have been made to enhance the specificity of these chimeric nucleases including restriction on spacer length between ZFN monomers [165], obligate heterodimers [114, 115], generation of autonomous ZFN pairs [91], and enhancement of the cleavage activity of *Fok*I [118]. In the past decade, numerous preclinical studies have descried the therapeutic utility of ZFNs in to correct several other human genetic mutations associated with sickle cell anemia [78, 79], X-linked SCID [77], and alpha-1-antitrypsin deficiency [80], and haemophilia [81, 87]. Significantly, ZFNs are now being tested in Phase I/II clinical trials for disruption of the HIV-1 co-receptor CCR5 [12, 144].

Genome editing can be utilized to generate precise genomic deletions at a targeted genomic locus [91, 92]. In this study, we engineered zinc-finger nucleases (ZFNs) to specifically delete exon 51 from the dystrophin gene to generate precise and repeatable frameshifts in the resulting transcript in DMD patient cells by the loss of this exon. The advantage of this method is that the resulting changes to the dystrophin

transcript will generate restored dystrophin proteins with predictable protein sequence. First, we engineered a panel of zinc-finger nuclease proteins using the publicly available extended Modular Assembly (e-MA) [159] or Context-Dependent Assembly (CoDA) [128] methods. Engineered nucleases were screened for activity by reporter assays in human cells and by monitoring gene-editing activities at the intended chromosomal loci. Several ZFN pairs demonstrated measurable activity at their intended chromosomal target, including two ZFN pairs flanking the splice acceptor of exon 51. Active ZFN pairs were observed to have modest levels of cytotoxicity and one ZFN pair had low levels of detectable off-target mutagenesis as detected by the Surveyor assay. Two selected ZFN pairs were transfected into DMD patient cells and a clonal cell line was isolated harboring the intended genetic deletion. After differentiation, we demonstrate that exon 51 is lost from the mRNA transcript and dystrophin protein expression was restored. Furthermore, these cells express human dystrophin properly localized to the sarcolemma membrane following transplantation into the hind limb of immunodeficient mice. Importantly, this study demonstrates a general method to delete sequences from the genome that result in permanent exclusion of a specific exon from the resulting mRNA transcript, thereby predictably restoring expression of the dystrophin protein.

4.3. Materials and Methods

4.3.1. Plasmid constructs

Extended Modular Assembly ZFNs were constructed using standard molecular biology techniques from a library of predefined zinc finger modules with predefined specificity [125] as described [159] in expression vectors containing the wild-type *Fok*I nuclease domain. In some cases, this library was supplemented with additional ZFs targeting TGC or TCT by grafting a recognition helix sequence (**Appendix B**) obtained from ZiFiT[16] onto the Sp1C zinc-finger motif backbone used by the other modular assembly ZFs. Coding regions for Context-Dependent Assembly [128] ZFNs were synthesized by BioBasic, Inc. (Ontario, Canada) and cloned by standard molecular biology techniques. The linker used to join zinc-finger domains to the *Fok*I domain was dependent on the spacer size between the half-sites, with the amino acid sequences HLRGS for 5 base-pair spacers, HTGAAARA for 6 base-pair spacers, and HTGPGAAARA for 7 base-pair spacers [165]. For all ZFN assays at chromosomal loci, *Fok*I domains were modified using both the ELD/KKR obligate heterodimer mutations [115] and the Sharkey mutations [118] as described previously [153]. Sequences for ZFN target sites **Appendix B**.

4.3.2. Cell culture and transfection

HEK293T cells were obtained from the American Tissue Collection Center (ATCC) through the Duke Cell Culture Facility and were maintained in DMEM

supplemented with 10% fetal bovine calf serum and 1% penicillin/streptomycin. Immortalized myoblasts [152] from a DMD patient harboring a deletion of exons 48-50 (Δ48-50) in the dystrophin gene were maintained in skeletal muscle media (PromoCell) supplemented with 20% fetal bovine calf serum (Sigma), 50 μg/ml fetuin, 10 ng/ml human epidermal growth factor (Sigma), 1 ng/ml human basic fibroblast growth factor (Sigma), 10 μg/ml human insulin (Sigma), 1% GlutaMAX (Invitrogen), and 1% penicillin/streptomycin (Invitrogen). All cell lines were maintained at 37°C and 5% CO_2. For screening ZFN activity at chromosomal loci, human HEK293T cells were transfected with Lipofectamine 2000 (Invitrogen) with 400 ng of each expression vector according to the manufacturer's protocol in 24 well plates. Immortalized myoblasts were transfected with 10 micrograms of each expression vector by electroporation using the Gene Pulser XCell (BioRad) with PBS as an electroporation buffer using optimized conditions [42]. Transfection efficiencies were measured by delivering an eGFP expression plasmid (pmaxGFP, Clontech) and using flow cytometry. These efficiencies were routinely ≥95% for HEK293T and ≥70% for the immortalized myoblasts.

4.3.3. Single-strand annealing assay

For this assay, extended modular assembly ZFNs were constructed in vectors utilizing the wild-type *Fok*I domain. Construction of the SSA luciferase reporter plasmid pSSA Rep 3-1 has been described previously [159]. Briefly, ZFN binding sites were introduced into the left and/or right arms of a split firefly luciferase gene by PCR, and

cloned into the *Bgl*II/*Eco*RI sites of the vector. All primers used for SSA construction are listed in the **Appendix B**. Human HEK293T cells were co-transfected with 25 ng of each ZFN monomer expression plasmid and 25 ng of SSA reporter plasmid in 96 well-plates using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells were lysed directly in the plate and 30 microliters of each lysate was transferred to 96 well plates for analysis using the Bright-Glo Luciferase Assay System (Promega E2620) and a luminescence plate reader (1 second integration).

4.3.4. Surveyor assay for endogenous gene modification

Genetic modifications were quantified using the Surveyor nuclease assay [166], which detects mutations characteristic of nuclease-mediated NHEJ. After transfection, cells were incubated for 3 or 10 days at 37°C and genomic DNA was extracted using the DNeasy Blood and Tissue kit (QIAGEN). The target locus was amplified by 35 cycles of PCR with the AccuPrime High Fidelity PCR kit (Invitrogen) using primers specific to each locus (**Appendix B**). The resulting PCR products were randomly melted and reannealed in a thermal cycler with the program: 95°C for 240 s, followed by 85°C for 60 s, 75°C for 60s, 65°C for 60s, 55°C for 60 s, 45°C for 60 s, 35°C for 60 s, and 25°C for 60s with a -0.3°C/s rate between steps. Following reannealing, 8 μl of PCR product was mixed with 1 μl of Surveyor Nuclease S and 1 μl of Enhancer S (Transgenomic) and incubated at 42°C for 1 hour. After incubation, 6 μl of digestion product was loaded onto a 10% TBE polyacrylamide gel and run at 200V for 30 min. The gels were stained with

ethidium bromide and quantified by densitometry using the ImageLab software suite (Bio-Rad) as previously described [166].

4.3.5. PCR-based assay to detect genomic deletions

The exon 51 locus was amplified from genomic DNA by PCR (Invitrogen AccuPrime High Fidelity PCR kit) using Cel-I primers flanking the DZF-1 (CelI-DZF1/2/10-R) and DZF-9 (CelI-DZF9-F) target sites (**Appendix B**). PCR products were separated on TAE-agarose gels and stained with ethidium bromide for analysis.

4.3.6 Clone isolation procedure

Immortalized DMD myoblasts were electroporated with 5 μg of each ZFN plasmid (10 μg total). After 7 days, isogenic clones were isolated by clonal density isolation. Genomic DNA was extracted from clones using the QuickExtract Kit (Epicentre) and the target locus amplified by PCR using primers to detect the expected genomic deletion as above (**Appendix B**). The resulting PCR products were analyzed to identify clones carrying the expected deletion and verified by Sanger sequencing.

4.3.7. mRNA analysis

Immortalized myoblasts were differentiated into myofibers by replacing the growth medium with DMEM supplemented with 1% insulin-transferrin-selenium (Invitrogen #51500056) and 1% penicillin/streptomycin (Invitrogen #15140) for 6 days before the cells were trypsinized and collected. Total RNA was isolated from these cells using the RNeasy Plus Mini Kit (QIAGEN) according to the manufacturer's instructions.

RNA was reverse transcribed to cDNA using the VILO cDNA synthesis kit (Life Technologies #11754) and 1.5 micrograms of RNA for 2 hours at 42°C according to the manufacturer's instructions. The target loci were amplified by 35 cycles of PCR with the AccuPrime High Fidelity PCR kit (Invitrogen) using primers annealing to exons 44 and 52 (**Appendix B**). PCR products were run on TAE-agarose gels and stained with ethidium bromide for analysis.

4.3.8. Western blot analysis

To assess dystrophin protein expression, immortalized myoblasts were differentiated into myofibers as above for 6 days. Cells were trypsinized, collected and lysed in RIPA buffer (Sigma) supplemented with a protease inhibitor cocktail (Sigma) and the total protein amount was quantified using the bicinchoninic acid assay according to the manufacturer's instructions (Pierce). Samples were then mixed with NuPAGE loading buffer (Invitrogen) and 5% β-mercaptoethanol and heated to 85°C for 10 minutes. Twenty-five micrograms of protein were separated on 4-12% NuPAGE Bis-Tris gels (Invitrogen) with MES buffer (Invitrogen). Proteins were transferred to nitrocellulose membranes for 1-2 hours in 1X tris-glycine transfer buffer containing 10% methanol and 0.01% SDS. The blot was then blocked for 1 hour with 5% milk-TBST at room temperature. Blots were probed with the following antibodies in 5% milk-TBST: anti-dystrophin C-terminus (1:25 overnight at 4°C, Leica NCL-DYS2), anti-dystrophin rod domain (1:1000 one hour at room temperature, Sigma MANDYS8), and anti-

GAPDH (1:5000 overnight at 4°C, Cell Signal 2118S). Blots were then incubated with horseradish peroxidase-conjugated secondary antibodies (Santa Cruz) and visualized using the ChemiDoc chemilumescent system (BioRad) and Western-C ECL substrate (BioRad).

4.3.9. Transplantation into immunodeficient mice

All animal experiments were conducted under protocols approved by the Duke Institutional Animal Care & Use Committee. Cells were trypsinized, collected and washed in 1X Hank's Balanced Salt Solution (HBSS, Sigma). Two million cells were pelleted and resuspended in five μL 1X HBSS (Sigma) supplemented with cardiotoxin (Sigma #C9759) immediately prior to injection. These cells were transplanted into the hind limb tibialis anterior (TA) muscle of NOD.SCID.gamma (NSG) mice (Duke CCIF Breeding Core) by intramuscular injection. Four weeks after injection, mice were euthanized and the TA muscles were harvested.

4.3.10. Immunofluorescence staining

Harvested TA muscles were incubated in 30% glycerol overnight at 4°C before mounting and freezing in Optimal Cutting Temperature compound. Serial 10 micron sections were obtained by cryosectioning of the embedded muscle tissue at -20°C. Cryosections were then washed in PBS to remove the OCT compound and subsequently blocked for 30-60 minutes at room temperature in PBS containing 10% heat-inactivated fetal bovine serum for spectrin detection or 5% heat-inactivated fetal bovine serum for

dystrophin detection. Cryosections were incubated overnight at 4°C with the following primary antibodies that are specific to human epitopes only: anti-spectrin (1:20, Leica NCL-SPEC1) or anti-dystrophin (1:10, Leica NCL-DYS3). After primary staining, spectrin or dystrophin expression was detected using a tyramide-based immunofluorescence signal amplification detection kit (Life Technologies, TSA Kit #22, catalog #T-20932,). Briefly, cryosections were incubated with 1:200 goat anti-mouse biotin-XX secondary (Life Technologies #B2763) in blocking buffer for one hour at room temperature. The signal was then amplified using streptavidin-HRP conjugates (1:100, from TSA Kit) in blocking buffer for one hour at room temperature. Finally, cryosections were incubated with tyramide-AlexaFluor488 conjugates (1:100, TSA kit) in manufacturer-provided amplification buffer for 10 minutes at room temperature. Stained cryosections were then mounted in ProLong AntiFade (Life Technologies #P36934) and visualized with conventional fluorescence microscopy.

4.3.11. Cytotoxicity assay

To quantitatively assess nuclease-associated cytotoxicity, HEK293T cells were transfected with 10 ng of a GFP reporter and 100 ng of each ZFN expression vector using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). The percentage of GFP positive cells was assessed at 2 and 5 days by flow cytometry. The survival rate was calculated as the decrease in GFP positive cells from days 2 to 5 and

normalized to cells transfected with an empty nuclease expression vector as described [155].

4.3.12. Off-target analysis using the PROGNOS predictive algorithm

Potential off-target sites were scanned for off-targets using the recommended parameters and the ZFN2.0 algorithm [167]. Briefly, the maximum number of mismatches allowed were considered for the length of the target site according to the software, heterodimeric and homodimeric target sites were allowed, and the top sites were binned for spacer lengths ideal for the zinc-finger and *Fok*I protein linker utilized in each monomer. For the purposes of this study, DZF-1 L6/R6 (linker: HTGAAARA) was assumed to have optimal activity on targets with 6 or 7 base-pair spacers between half-sites [165]. Similarly, DZF-9 (linker: HLRGS) was assumed to have ideal activity on 5 or 6 base-pair spacers. Ten micrograms of each monomer was electroporated into human DMD patient myoblasts as described above and genomic DNA was collected 3 days following transfection. Potential off-target loci were PCR amplified using primers were generated from the PROGNOS output (**Appendix B)** and off-target activity was quantified using the Surveyor assay as described above.

4.4. Results

4.4.1. Design of ZFNs to target exon 51

The generation of ZFN pairs that are highly active at chromosomal loci remains a significant challenge in using this technology. Thus, we sought to create a large panel of

ZFN pairs with targets across exon 51 of the dystrophin gene and its flanking introns with the goal of finding a combination of ZFN pairs to delete the entire exon or sequences important to its proper splicing in the resulting mRNA transcript (Figure 11). Using a publicly available webserver [168], this region was scanned for Context-Dependent Assembly (CoDA) ZFN targets and several potential ZFN pairs were assembled. To supplement the limited number of available CoDA targets at this locus, we engineered several additional ZFN pairs using the previously described extended Modular Assembly (e-MA) method [159]. Together, these ZFN pairs are designed to flank the entire exon or either of the two splice junctions on the 5′ or 3′ end, respectively. Deletion of one or more of these conserved splice junctions was predicted to result in loss of the entire exon from the dystrophin transcript.

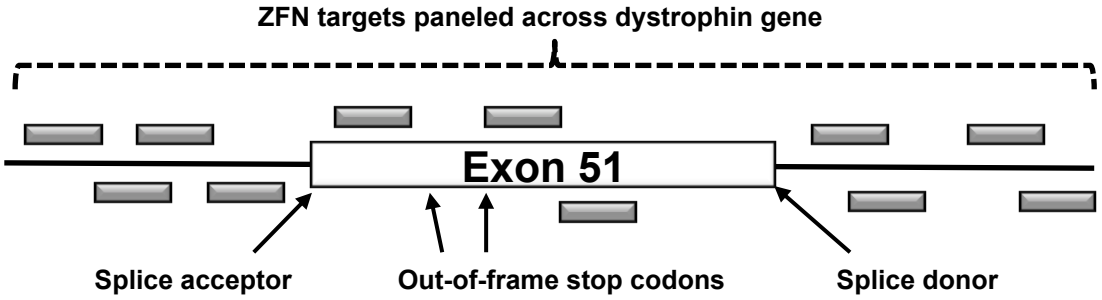


Figure 11: Design of ZFNs targeted to exon 51. ZFN pairs (shown as blocks) were designed as a panel of targets across exon 51 and the flanking introns.

4.4.2. Screening for activity of extended Modular Assembly ZFNs

Extended modular assembly assumes that active ZFNs are likely to have three to six zinc finger motifs in tandem. Accordingly, a panel of e-MA ZFNs was designed

against exon 51 and the flanking introns *in silico* using the publicly available Zinc Finger Tools website [124, 168] to construct ZFNs by e-MA. We generated dozens of engineered proteins consisting of three to six zinc fingers per target site across six target sites. We screened these e-MA ZFNs for activity against their cognate target site to identify optimal zinc finger composition using an episomal luciferase reporter assay for ZFN activity. This assay utilizes a split luciferase gene that has a specific target site separating a flanking region of luciferase homology that will recombine to form active luciferase when the target site is correctly recognized and cleaved by a ZFN pair (**Figure 12a**). Following transfection into human cells, nine candidate e-MA ZFNs with high activity were identified for further analysis (**Figure 12b**). Notably, similar to previous studies for extended modular assembly ZFNs [159], increased activity was observed as additional zinc-finger motifs were added to a ZFN monomer, particularly at four or more zinc-fingers per half site.

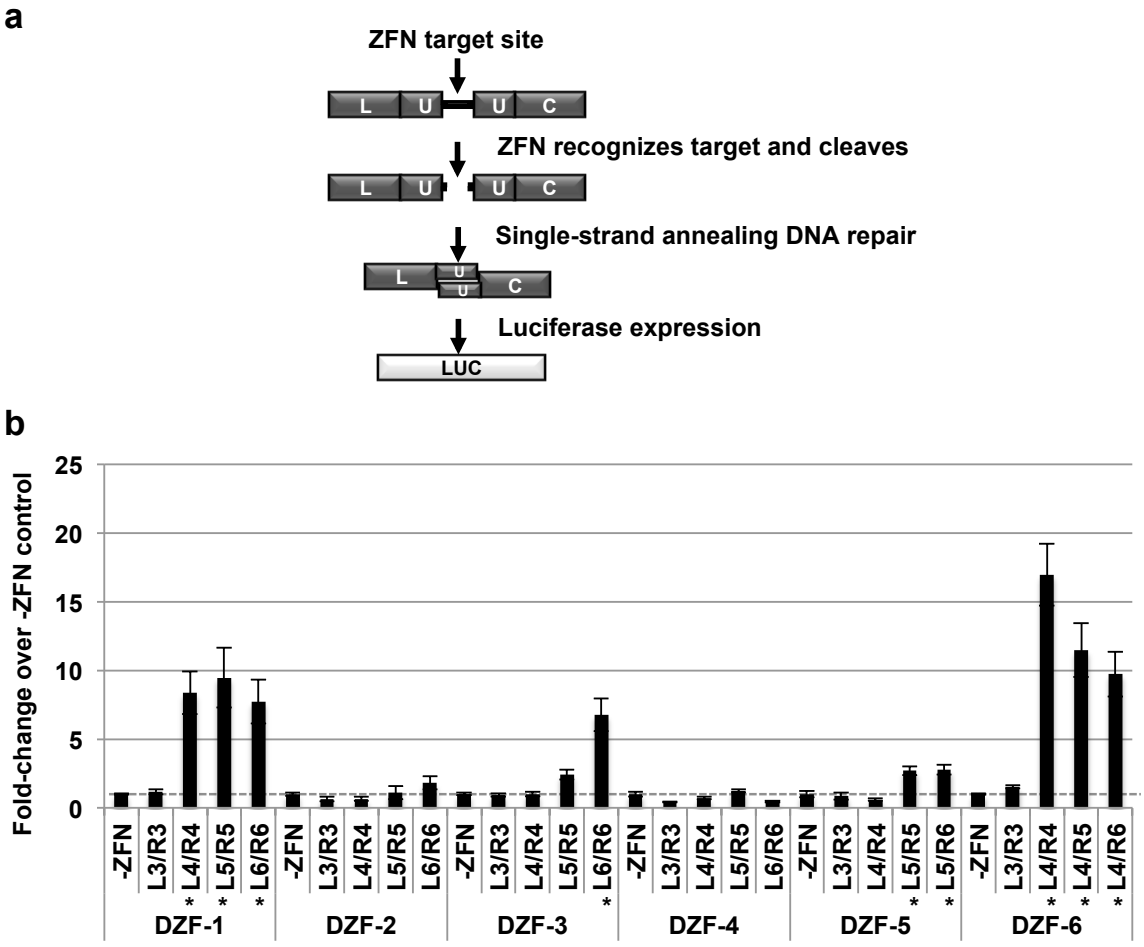


Figure 12: Screening for active extended modular assembly ZFNs using an episomal reporter assay. All ZFNs used the wild-type *Fok*I nuclease domain. (a) Schematic of single-stranded annealing assay to detect ZFN activity. Each target site was cloned between a split luciferase reporter with flanking homology on either side of each target sequence. Luciferase expression occurs when a ZFN pair successfully recognizes and cleaves its cognate site in the reporter, causing single-strand annealing and recombination of an active luciferase gene. (b) Activity of different combinations of extended modular assembly ZFN pairs compared to reporter only control in HEK293T cells. Asterisks indicate ZFN pairs selected for further testing.

4.4.3. Evaluation of ZFN activity at endogenous targets

An alternative method of ZFN design, termed context-dependent assembly (CoDA) [128], creates ZFNs with novel DNA recognition by recombining a library of

previously characterized zinc finger arrays. Initial studies suggested that this method engineers new ZFNs with a 50% success rate for cleaving a chromosomal target. This is a significantly higher rate than previously described for engineering novel modular assembly ZFNs, though extended modular assembly techniques may reach similar success rates [159]. Seven CoDA ZFNs were selected and the gene constructs encoding these ZFNs were assembled. Since CoDA ZFNs have an established high success rate [128], each ZFN pair was immediately tested for activity at chromosomal loci. Nine highly active e-MA ZFNs (**Figure 12b**) and seven designed CoDA ZFNs were electroporated into human myoblasts to test their ability to cleave target chromosomal loci. Using the Surveyor assay, we identified three e-MA ZFNs and three CoDA ZFNs that had activity at the intended chromosomal locus (**Figure 13a**). Gene modification was still detectable for 4 out of 6 ZFN pairs after 10 days, and remained stable (<25% signal change) for all tested e-MA ZFNs (**Figure 13b**). Interestingly, despite efficient gene editing activity at 3 days post-transfection, all three CoDA ZFNs showed a significant or complete loss of signal by day 10 (**Figure 13b**).

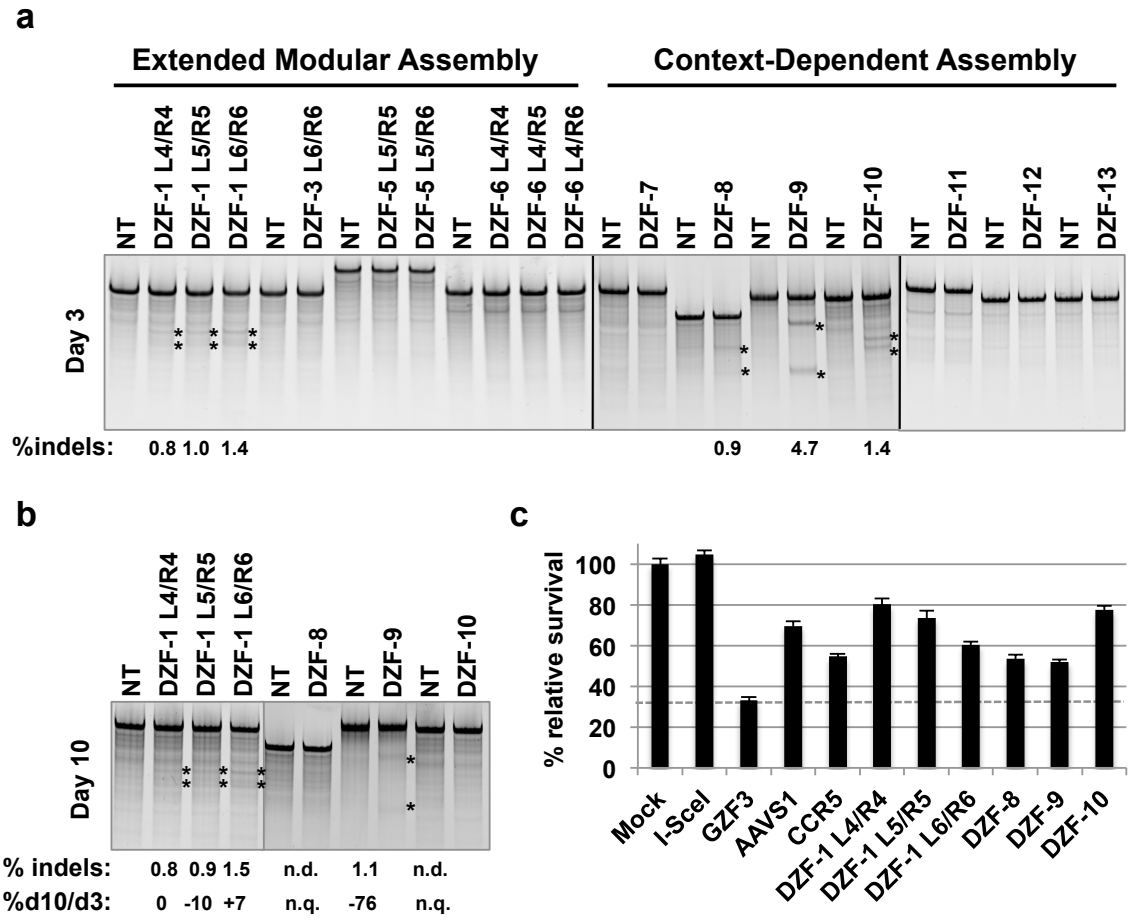


Figure 13: Evaluation and characterization of selected ZFNs in human cells. (a) 400ng of each monomer expression plasmid for all CoDA-ZFNs and selected eMA-ZFNs were transfected into HEK293Ts and endogenous gene editing activity measured at 3 days post-transfection by the Surveyor assay. (b) Activity of ZFN pairs with measurable activity in (a) at 10 days post-transfection. The ratio of gene editing activity at 3 and 10 days was calculated from the data in (a) and (b). n.d.: not detected. n.q.: not quantified. (c) Results of a GFP retention-based cytotoxicity assay in HEK293T cells after transfection with the indicated nucleases and a GFP reporter. Percentage survival was calculated as the ratio of percent GFP cells positive at days 2 or 5 post-transfection and normalized to mock transfection in the absence of nucleases.

4.4.4. Characterization of ZFN cytotoxicity

To further assess the toxicity of designed ZFNs, we transfected human cells with constructs carrying the six ZFNs with detectable chromosomal gene editing activity (**Figure 13a,b**). The cytotoxicity of these ZFNs was evaluated using a flow cytometry-based GFP retention assay that measures the survival of highly transfected cells in a bulk population of transfected cells [42]. All of the ZFNs tested had moderate levels of cytotoxicity compared to I-SceI, a known non-toxic nuclease, and GZF3, a toxic ZFN pair (**Figure 13c**). Interestingly, a modest increase in cytotoxicity was observed as the number of ZF motifs was increased in creating an e-MA ZFN pair targeted to DZF-1 sequence. Despite displaying mildly increased cytotoxicity, gene editing activity appeared stable for all DZF-1 targeting ZFN pairs (**Figures 13a,b**). Overall, the ZFN pairs engineered in this study had measured cytotoxicities comparable to two other well-characterized ZFNs targeting AAVS1 [169] or CCR5 [12] loci.

4.4.5. Restoration of the dystrophin gene by deleting exon 51 from the genome

Co-expression of two nucleases has been demonstrated to mediate deletion of the intervening chromosomal sequence between the two nuclease target sites [91, 92]. This could be exploited to permanently delete an exon at the genetic level, in contrast to current methods that transiently remove the exon at the mRNA level. To apply this to deleting exon 51, we utilized two ZFNs, DZF-1 L6/R6 and DZF-9, that were identified to efficiently cleave chromosomal targets that flank the exon 51 splice acceptor sequence

(**Figures 13a and 14a**). Co-expression of these ZFNs is intended to result in excision of the intervening 2.7 kb segment that is expected to contain sequences necessary to include exon 51 in the dystrophin mRNA transcript (**Figure 14a**). Constructs encoding the DZF-1 L6/R6 and DZF-9 ZFN pairs were electroporated into DMD patient myoblasts. We verified that each ZFN pair was active in human DMD patient cells (**Figure 14b**). The expected genomic deletions were detected by end-point PCR only in cell populations treated with both combinations of ZFNs (**Figure 14c**). After verifying the presence of the expected genomic deletion, isogenic clones of DMD patient cells were derived and screened for this deletion event. One clone of interest was identified and Sanger sequencing analysis confirmed a new junction of intron 50 and intron 51 sequences flanking the target sites of the ZFN pairs, resulting in the loss of the 2.7kb region from the genome (**Figure 14d**). After this 2.7kb sequence is removed, only a partial fragment of exon 51 remains in the genome. Since the deleted segment includes essential splice acceptor sequences, the remaining exon 51 fragment is unlikely to be incorporated into the dystrophin mRNA transcript, resulting in the loss of exon 51 entirely. After differentiating this clonal population into myoblasts, mRNA RT-PCR analysis showed that exon 51 was indeed efficiently removed from the dystrophin transcript (**Figure 14e**). Furthermore, genomic deletion and removal of exon 51 from the dystrophin transcript resulted in restored dystrophin expression in these cells (**Figure 14f**), detected with antibodies targeting upstream (MANDYS8) and downstream (NCL-

Dys2) of the corrected patient mutation. These data demonstrate that gene editing is an effective method to specifically delete exons by removing essential splicing sequences from the genome.

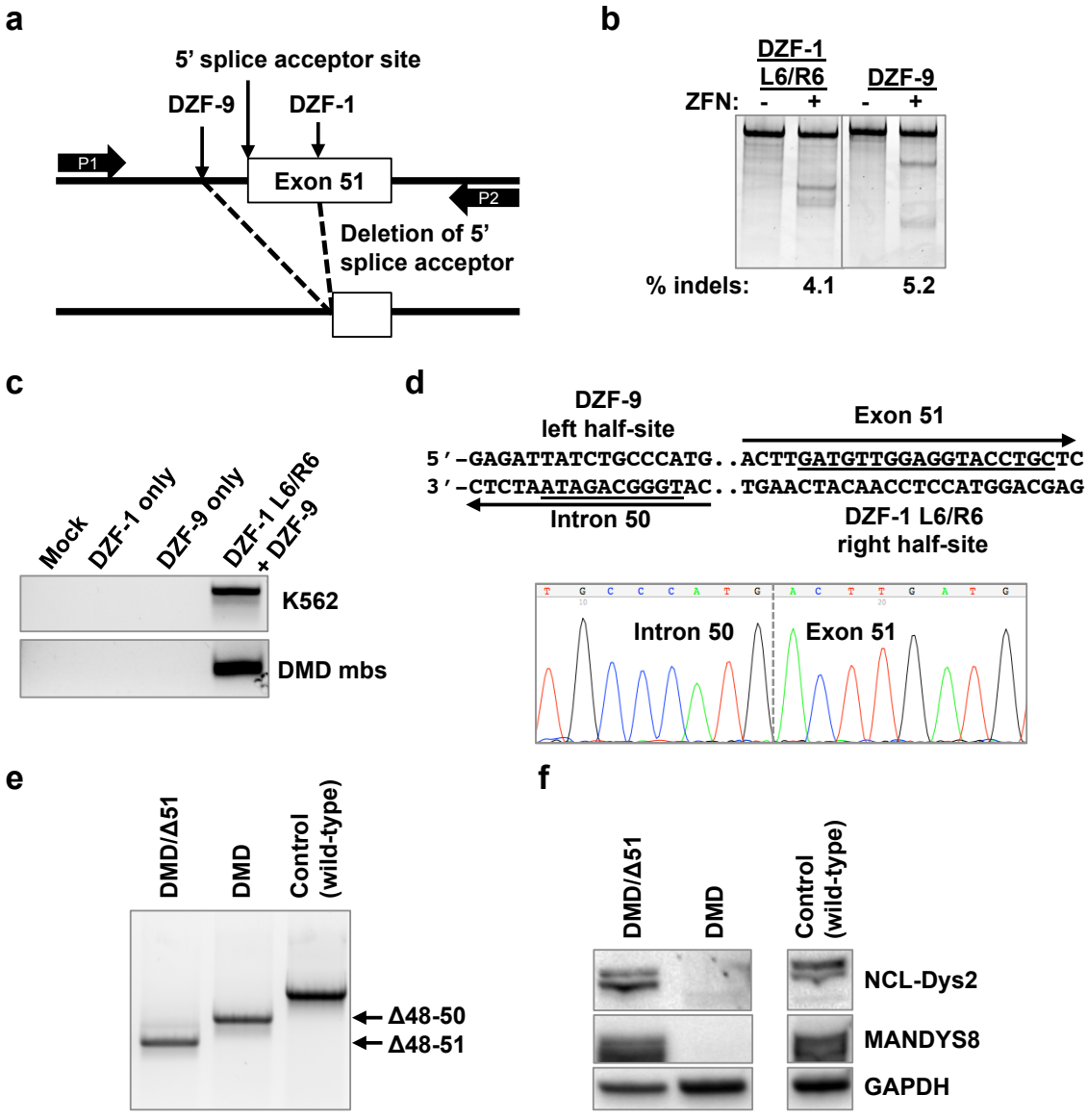


Figure 14: Restoration of the dystrophin reading frame in patient cells. (a) Schematic of strategy to delete exon 51 from the dystrophin gene locus. DZF-1 and DZF-9 flank the 5′ splice acceptor site of exon 51, which is removed after genomic

deletion. P1/P2: Primers used for genomic deletion PCR in (c). (b) Gene modification activities of DZF-1 L6/R6 and DZF-9 as measured by the Surveyor assay after electroporation of 10 micrograms of each monomer expression cassette into DMD patient cells. (c) End-point genomic PCR across deleted locus in human HEK293T or DMD patient cells after treating cells with the indicated pair of nucleases. (d) Sanger sequencing result of PCR product from genomic DNA of a genetically corrected clonal cell population. Underlined sequences show target half-sites for the indicated ZFN target site. (e) End-point RT-PCR analysis of mRNA from control wild-type and untreated or a genetically corrected clonal population (DMD/Δ51) of DMD patient myoblasts after differentiation into myotubes. (f) Dystrophin expression as detected by western blot with antibodies to detect the C-terminus (NCL-DYS2) or rod domain (MANDYS8) in each of the indicated cell populations. Different exposure times for the NCL-Dys2 western images were used to image DMD and DMD/Δ51 or control samples to compensate for overexposure of control protein. The images for MANDYS8 and GAPDH are the same exposure time for all samples.

4.4.6. Human dystrophin expression *in vivo* following transplantation of genetically corrected cells

Transplantation of genetically corrected autologous myoblasts is an attractive method to introduce functional dystrophin expression to skeletal muscle *in vivo* [36]. To demonstrate the feasibility of this approach, we transplanted a clonally derived population of DMD patient cells with a corrected dystrophin gene carrying a deletion of exon 51 (**Figures 14d-f**) and assessed human dystrophin expression *in vivo*. After 4 weeks, muscle fibers positive for human spectrin, which is expressed by both corrected and uncorrected cells, were detected in cryosections of injected muscle tissue (**Figures 15-16**). A significant number of these fibers were also positive for human dystrophin with expression localized to the sarcolemma, demonstrating functional protein correction in these cells (**Figures 15-16**). No fibers positive for human dystrophin were observed in sections from mice injected with the untreated DMD myoblasts (**Figures 15-**

16), indicating that the genetically corrected cells were the source of human dystrophin expression.

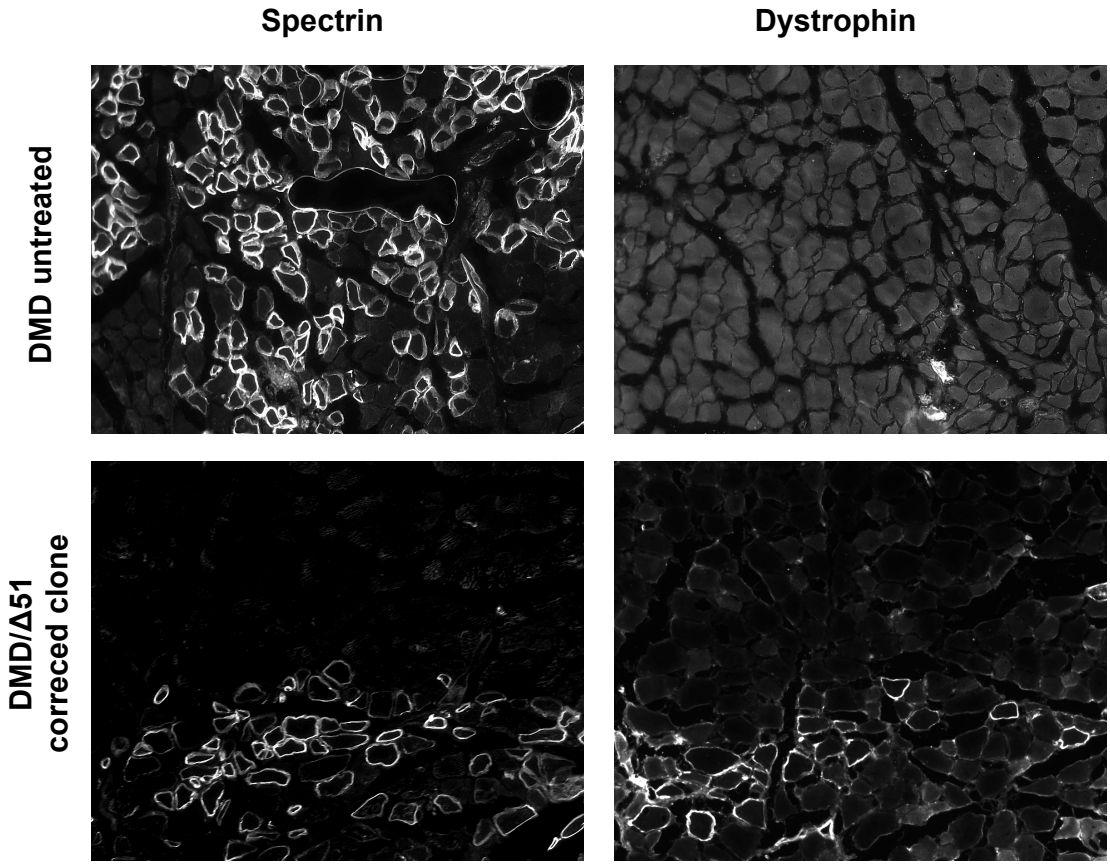


Figure 15: Untreated or genetically corrected (DMD/Δ51) human Δ48-50 DMD myoblasts carrying a background deletion of exons 48-50 were injected into the hind limbs of immunodeficient mice and assessed for human-specific protein expression in muscle fibers after 4 weeks post-transplantation. Cryosections were stained with anti-human spectrin, which is expressed by both uncorrected and corrected myoblasts that have fused into mouse myofibers, or anti-human dystrophin antibodies as indicated.

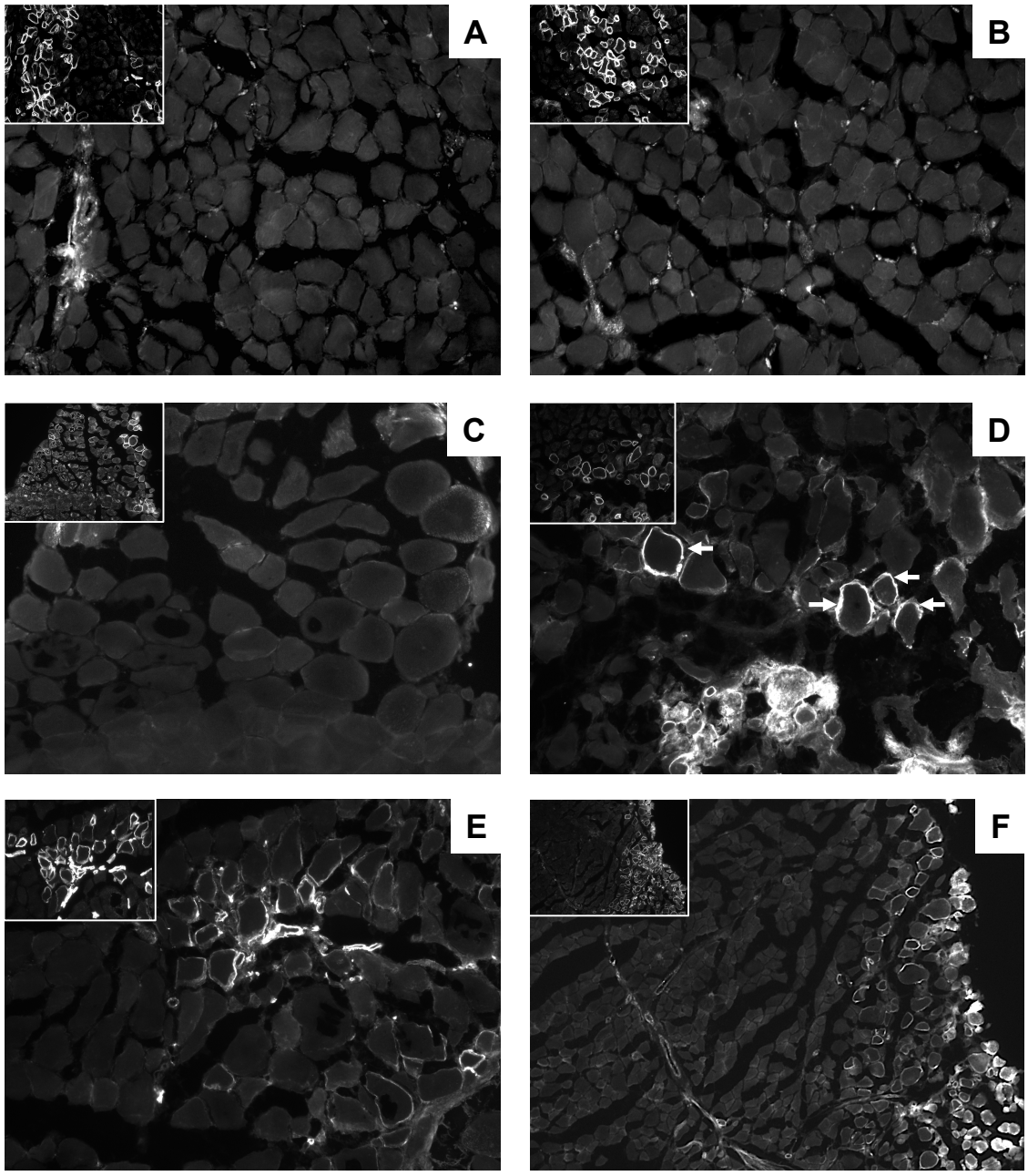


Figure 16: Additional immunofluorescence images probing human dystrophin expression. Serial sections from regions stained with anti-human spectrin are shown inset in top left. (a-c) Sections from muscles injected with untreated human DMD myoblasts. (d-f) Sections from muscles injected with clonally derived DMD myoblasts carrying a deletion of exon 51 to that corrects the dystrophin reading frame. White arrows indicate dystrophin positive fibers.

4.4.7. Analysis of genome integrity after editing the dystrophin gene

Off-target activity of engineered enzymes is a primary concern for gene editing therapies. To predict potential off-target sites, we utilized a publicly available tool, PROGNOS, that compiles and ranks potential off-target sites *in silico* [167]. Using the ZFN2.0 detection algorithm in PROGNOS, we selected the top 10 potential off-target sites in the genome for both DZF-1 L6/R6 and DZF-9 ZFN pairs. Eight of the ten identified off-target sites for each ZFN pair (**Table 2**) were successfully amplified and assessed for activity by the Surveyor assay following transfection of the respective ZFN pairs into human DMD patient cells (**Figure 17**). DZF-1 L6/R6 had no observed off-target activity (0/8 loci, **Figure 17a**), while DZF-9 had measurable activity at 2/8 loci (**Figure 17b**), albeit at lower levels than the on-target locus. Notably, the two *bona fide* off-target loci with observable activity for the DZF-9 ZFN pair had substantial homology to the intended target site (1 mismatch, **Table 2**). While we cannot rule out activity at the other off-target loci that may exist below the sensitivity of the Surveyor assay or off-target activity at other loci that were not assessed here, these data demonstrate the relative specificity of our reagents that is comparable to other studies utilizing ZFNs [87, 167, 170, 171].

Table 2: *In silico* prediction of putative off-target sites for DZF-1 or DZF-9 predicted by the online PROGNOS ZFN v2.0 webtool.

DZF-1 L6/R6									
Top ten with spacers of 6-7 (optimal for this linker)									
OT#	**HS**	**Spacer**	**MM**	**Left Site**	**Right Site**	**Chr**	**Chr Region**	**Region**	**Gene**
1	L/R	6	5/4	CcAACTtGAgATGCCAgC	GgTGTgGGAGGTcgaTGC	chr12	121465217	Intron	OASL
2	L/L	7	6/2	CtAgCaAGAAgTcCCtTC	GATGGCATcTCTAcTTTG	chr2	164675298	Intergenic	-
3	L/L	7	4/3	CAAACTAtAAgTGCCgTg	GATGtCAaTTaTAGTTTG	chr6	3268596	Intergenic	-
4	L/R	6	6/6	attACTtctAATGCCATC	GATaTTGGAGaTAtaTtt	chr11	73441555	Intron	RAB6A
5	L/R	7	6/6	CAAgaacctAATGCCATC	GtcGgaGGAGGTACCccC	chr12	349901	Intron	SLC6A13
6	L/L	7	6/4	CAAAggAGAAATGCCAag	GATGGCtcTTaaAGTTgc	chr14	87103919	Intergenic	-
7	R/R	7	6/6	GCAaactCCTCCAACAag	GATGTTGGAaagACaaaC	chr8	141637370	Intron	EIF2C2
8	R/L	6	6/5	cCAGaTACCaCCttaATC	GcTGGCATgTCTcccTTG	chr1	220550165	Intergenic	-
DZF-9									
Top ten with spacers of 5-6 (optimal for this linker)									
OT#	**HS**	**Spacer**	**MM**	**Left Site**	**Right Site**	**Chr**	**Chr Region**	**Region**	**Gene**
1	R/L	6	1/0	CCCTGCcCC	TGGGCAGAT	chr6	12015411	Intron	HIVEP1
2	L/R	6	1/0	ATCTGCCCA	GGCGCAGtG	chr4	58293132	Intergenic	-
3	R/L	6	1/0	CCCTGCGCC	gGGGCAGAT	chr18	76473917	Intergenic	-
4	L/R	5	1/0	ATCTGCCCA	GGCGCAaGG	chr17	8113802	Intron	AURKB
5	L/R	5	1/0	ATCTGCCCA	GGCcCAGGG	chr7	128034911	Intron	IMPDH1
6	R/L	6	1/0	CCCTGgGCC	TGGGCAGAT	chr10	80912948	Intron	ZMIZ1
7	R/L	5	1/0	CCCTGaGCC	TGGGCAGAT	chr4	1595869	Intergenic	-
8	L/R	6	1/0	ATCTGCCCA	GGCtCAGGG	chr18	74383767	Intergenic	-

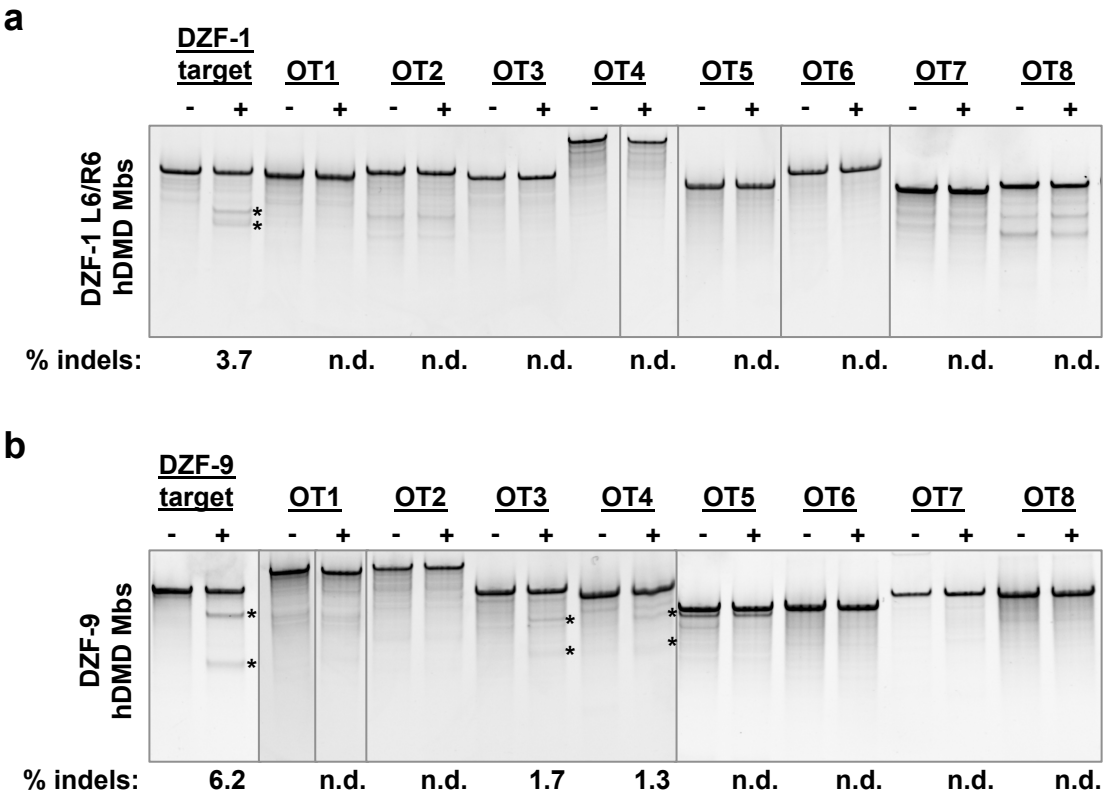


Figure 17: Evaluation of ZFN off-target effects in human cells. Human DMD myoblasts were electroporated with ten micrograms of DNA constructs encoding either DZF-1 L6/R6 or DZF-9. After 3 days, genomic DNA was analyzed by the Surveyor to measure off-target activity at eight different loci for DZF-1 L6/R6 (a) or DZF-9 (b). Asterisks indicate detectable Surveyor cleavage products.

4.5. Discussion

The rapid advancement of gene editing technologies has enabled at-will modification and correction of disease-related genes. This study introduces a novel method to correct the dystrophin gene by deleting exons from the genome, thereby permanently excluding the exons in the dystrophin transcript. Importantly, this genome editing method will likely be compatible with many existing gene- and cell-based

therapies in development for DMD. Autologous, genetically corrected myogenic cells present an exciting strategy to introduce functional dystrophin *in vivo,* currently a major challenge for DMD therapies. Here, we demonstrate that a clonally derived, genetically corrected population of DMD patient cells can generate human dystrophin expression *in vivo* that is properly localized to the sarcolemma membrane. Thus, gene correction by excising exons from the genome may be a viable method for creating an autologous population of corrected cells.

The utility of this approach will likely require enhancing the overall efficiency of deletions, particularly to expand this strategy to *in vivo* delivery of gene editing enzymes to correct the dystrophin gene *in situ*. Improving the activity and specificity of each nuclease pair may also further enhance the efficiency of this approach. All of the ZFNs engineered in this study displayed moderate cytotoxicity in human cells similar to two benchmark ZFNs targeting AAVS1 and CCR5. One ZFN pair, DZF-9, had observable activity in 2 of 8 tested off-target loci in human cells. Notably, all eight of these target sites had only 1 mismatch to the intended target site. It may be possible to reduce the observed off-target activity by utilizing recent advances in ZFN technology that have enabled replacement of zinc finger modules to increase specificity and decrease off-target toxicities [172]. The fidelity of this approach may be further enhanced by incorporating highly active orthogonal *Fok*I obligate heterodimer mutations that would increase the specificity of this approach by limiting or eliminating unintended off-target

pairings between monomers from each ZFN pair [91]. This may reduce the potential for generating unintended chromosomal rearrangements by creating simultaneous DSBs at unintended chromosomal loci. Other gene editing technologies, such as CRISPR/Cas9 [10, 11] or TALENs [8, 113], may present suitable alternatives to introduce similar genomic deletions with potentially reduced toxicity and/or off-target activity.

Based on these results, deletion of genetic sequences containing splice sites for exon 51 results in the complete exclusion of the exon from the transcript and restoration of dystrophin protein expression. This study demonstrates a proof of principle approach to correcting the DMD reading frame by introducing predictable and repeatable changes to the genome that eliminate exon 51 from the transcript. Importantly, genome editing enables permanent changes in the corrected cell and its progeny, thereby potentially eliminating the need for continuous administration of a therapeutic agent. In contrast to our previous study [42], the genetic changes here are intended to remove exon 51 and generate expression of a protein that is predictable based on the patient background deletion, similar to the changes caused by oligonucleotide-mediated exon skipping. The advantage of this method is that it reproducibly generates internally deleted proteins with known protein sequences and predictable functionality. This method presents a robust gene editing approach to restore the dystrophin gene that can be extended to address additional patient deletions common in DMD and serves as a blueprint for correcting the genetic basis of other monogenic hereditary disorders.

Chapter 5: Correction of Duchenne Muscular Dystrophy by Multiplex CRISPR/Cas9-Based Genome Editing

Co-authored with Ami M. Kabadi, Pratiksha I. Thakore, and Charles A. Gersbach

5.1. Synopsis

The recently described CRISPR/Cas9 gene editing platform presents a novel tool to correct the genetic basis of hereditary diseases. The increased versatility, efficiency, and multiplexing capabilities of the CRISPR/Cas9 system enable a variety of otherwise challenging gene correction strategies. We have used the CRISPR/Cas9 system to restore the expression of the dystrophin gene that is mutated in Duchenne muscular dystrophy. Single or multiplexed sgRNAs were designed to restore the dystrophin reading frame by targeting the mutational hotspot at exons 45-55 and introducing either intraexonic small insertions and deletions, or large deletions of one or more exons. Following treatment with Cas9 and one or more sgRNAs, dystrophin expression was restored in Duchenne patient muscle cells *in vitro.* Human dystrophin was detected *in vivo* following transplantation of genetically corrected patient cells into immunodeficient mice. Significantly, the unique multiplex gene editing capabilities of the CRISPR/Cas9 system enable efficiently generating large deletions of this mutational hotspot region that can

correct up to 62% of patient mutations by universal or patient-specific gene editing approaches.

5.2. Introduction

Genome editing technologies use synthetic nucleases to induce cellular DNA repair mechanisms and introduce site-specific, predefined genetic modifications in complex genomes [162]. These engineered enzymes are commonly based on zinc finger nucleases (ZFNs) [6], transcription activator-like effector nucleases (TALENs) [75], meganucleases [7], and most recently the RNA-guided CRISPR/Cas9 system [10, 11, 99-101, 103, 173]. The nucleases create site-specific double-strand breaks (DSBs) at predefined genomic sites that stimulate either non-homologous end joining (NHEJ) for targeted gene disruption or homologous recombination for highly efficient gene targeting. The simplicity and versatility of the CRISPR/Cas9 genome editing system has led to rapid adoption and expansion of this technology that has proven to be remarkably robust for manipulating gene sequences in human cells. This has enabled new possibilities such as efficient multiplex gene editing for simultaneously inactivating multiple genes [10, 11, 135]. In this study, we apply the CRISPR system to repair genes mutated in hereditary disease, including capitalizing on the unique multiplex capacity of this technology to create large genomic deletions that restore gene expression.

CRISPR/Cas9 systems have been adapted from multiple bacterial species, including *S. pyogenes*, *S. thermophilus*, and *N. meningitidis*, to efficiently generate targeted

gene modifications in human cells [10, 11, 99-101, 103, 136, 173]. These systems consist of a Cas9 nuclease that is co-expressed with a single guide RNA (sgRNA) molecule. Cas9 forms a complex with the 3′ end of the sgRNA, and the protein-RNA pair recognizes its genomic target by complementary base pairing between the 5′ end of the sgRNA sequence and a predefined 20 bp DNA sequence, known as the protospacer. By simply exchanging the 20 bp recognition sequence of the expressed sgRNA, the Cas9 nuclease can be directed to new genomic targets. The only restriction for protospacer targeting is that the sequence must be immediately followed by the protospacer-adjacent motif (PAM), a short sequence recognized by the Cas9 nuclease that is required for DNA cleavage. Several studies of the *S. pyogenes* CRISPR system have defined the PAM sequence for this Cas9 (SpCas9) as 5′-NRG-3′, where R is either A or G, and characterized the specificity of this system in human cells [137-142, 174]. A unique capability of the CRISPR/Cas9 system is the straightforward ability to simultaneously target multiple distinct genomic loci by co-expressing a single Cas9 protein with two or more sgRNAs [10, 11, 135].

One of the most promising applications of genome editing is the correction of genetic mutations associated with hereditary disease [6, 7, 75, 162]. Duchenne muscular dystrophy (DMD) is the most common hereditary disease and no effective treatments exist for this disorder. DMD is a severe X-linked disease that presents with progressive muscle wasting that typically leads to loss of ambulation in the second decade and death

within the third decade of life due to respiratory complications or heart failure. The molecular basis of DMD is a mutation in the dystrophin gene [2] that leads to the complete lack of function of this essential skeletal muscle protein. These mutations are most commonly frameshifts generated by large intragenic deletions of one or more exons. DMD is the prototypical example of a group of monogenetic hereditary diseases that can be corrected by removing internal, but unessential, regions of the mutated gene to restore the proper reading frame [4, 175]. For example, there is a class of common deletions in the exon 45-55 mutation hotspot region of the dystrophin gene that maintain the correct reading frame and lead to the expression of a truncated, but functional, dystrophin protein. Patients with this class of mutations are often asymptomatic or display mild symptoms associated with Becker muscular dystrophy, a substantially less severe disease than DMD. This has led to significant interest in developing an oligonucleotide-mediated exon skipping strategy that will restore the dystrophin reading frame during mRNA processing and convert DMD to a Becker-like phenotype [4]. Whereas early clinical trials in this area have focused on skipping only exon 51 [70, 71], which is applicable to 13% of DMD patients, other preclinical efforts have demonstrated multi-exon skipping of the complete exon 45-55 coding region with a combination treatment of up to 10 oligonucleotides [176, 177] that could potentially address greater than 60% of known DMD patient mutations [175]. However, there are significant technical and practical hurdles to designing and developing this type of

complex combination therapy, in addition to the general challenges of developing any oligonucleotide-based therapy that must be continuously administered for the lifetime of the patient. In contrast to these transient mRNA-targeted oligonucleotide-mediated exon skipping strategies, genome editing has the ability to make precise changes to gene sequences that will be permanent in treated cells and persist after cell division. Additionally, only two nucleases are necessary to delete a genomic region of any length, in contrast to exon skipping in which a distinct oligonucleotide must be designed for each exon to be removed from the mRNA transcript.

Genome editing using various designer nucleases has been proposed as a promising method to restore the native dystrophin gene in DMD patient cells [42, 90, 163]. However, an obstacle to implementing this approach has been successfully engineering the multiple nucleases targeted to the exons and introns necessary to address a large fraction of the DMD patient population. In this study, we take advantage of the versatility of the *S. pyogenes* CRISPR/Cas9 system to rapidly and efficiently generate targeted frameshifts and large deletions to address commonly occurring mutations in the dystrophin gene across exons 45-55. Skeletal myoblasts from DMD patients were treated with sgRNAs and SpCas9 to correct patient-specific mutations and edited cells were enriched by fluorescence-activated cell sorting. Gene editing by CRISPR/Cas9 resulted in restored dystrophin mRNA transcripts and protein expression. Significantly, we generated a large deletion of 336 kb across a mutational hotspot

containing exons 45-55 that is applicable to correction of greater than 60% of DMD patient mutations. This genomic deletion resulted in the loss of exons 45-55 in the corresponding dystrophin transcript and restored dystrophin expression in human DMD cells. Additionally, an enriched pool of gene-corrected cells demonstrated expression of human dystrophin *in vivo* following engraftment into immunodeficient mice. CRISPR/Cas9 gene editing did not have significant toxic effects in human myoblasts as observed by stable gene editing frequencies and minimal cytotoxicity of several sgRNAs. However, gene editing activity was confirmed at three out of 50 predicted off-target sites across five sgRNAs and CRISPR/Cas9-induced chromosomal translocations between on-target and off-target sites were detectable, indicating a need to increase the specificity of this technology. Collectively, this study demonstrates that the CRISPR/Cas9 technology is an efficient and versatile method for correcting a significant fraction of dystrophin mutations and can serve as a general platform for treating genetic disease.

5.3. Materials and Methods

5.3.1. Plasmid constructs

The expression cassettes for the *S. pyogenes* sgRNA and human codon optimized Cas9 (hCas9) nuclease were used as previously described [178]. In order to create a fluorescent reporter system to enrich CRISPR/Cas9-modified cells, a GeneBlock (IDT) was synthesized containing a portion of the 3′ end of the Cas9 coding sequence fused to

a T2A skipping peptide immediately upstream of a multiple cloning site and subsequently cloned into the hCas9 expression vector. An eGFP reporter gene was then cloned into the T2A vector to allow co-translation of Cas9 and eGFP proteins from the same expression vector.

5.3.2. Cell culture and transfection

HEK293T cells were obtained from the American Tissue Collection Center (ATCC) through the Duke Cell Culture Facility and were maintained in DMEM supplemented with 10% fetal bovine calf serum and 1% penicillin/streptomycin. Immortalized myoblasts [152] from a DMD patient harboring a deletion of exons 48-50 (Δ48-50) in the dystrophin gene were maintained in skeletal muscle media (PromoCell) supplemented with 20% fetal bovine calf serum (Sigma), 50 μg/ml fetuin, 10 ng/ml human epidermal growth factor (Sigma), 1 ng/ml human basic fibroblast growth factor (Sigma), 10 μg/ml human insulin (Sigma), 1% GlutaMAX (Invitrogen), and 1% penicillin/streptomycin (Invitrogen). All cell lines were maintained at 37°C and 5% CO_2. HEK293T cells were transfected with Lipofectamine 2000 (Invitrogen) with 400 ng of each expression vector according to the manufacturer's protocol in 24 well plates. Immortalized myoblasts were transfected with 5 micrograms of each expression vector by electroporation using the Gene Pulser XCell (BioRad) with PBS as an electroporation buffer using optimized conditions [42]. Transfection efficiencies were measured by delivering an eGFP expression plasmid (pmaxGFP, Clontech) and using flow cytometry.

These efficiencies were routinely ≥95% for HEK293T and ≥70% for the immortalized myoblasts.

5.3.3. Surveyor assay for endogenous gene modification

CRISPR/Cas9-induced lesions at the endogenous target site were quantified using the Surveyor nuclease assay [166], which detects mutations characteristic of nuclease-mediated NHEJ. After transfection, cells were incubated for 3 or 10 days at 37°C and genomic DNA was extracted using the DNeasy Blood and Tissue kit (QIAGEN). The target locus was amplified by 35 cycles of PCR with the AccuPrime High Fidelity PCR kit (Invitrogen) using primers specific to each locus (**Appendix C**). The resulting PCR products were randomly melted and reannealed in a thermal cycler with the program: 95°C for 240 s, followed by 85°C for 60 s, 75°C for 60s, 65°C for 60s, 55°C for 60 s, 45°C for 60 s, 35°C for 60 s, and 25°C for 60s with a -0.3°C/s rate between steps. Following reannealing, 8 μl of PCR product was mixed with 1 μl of Surveyor Nuclease S and 1 μl of Enhancer S (Transgenomic) and incubated at 42°C for 1 hour. After incubation, 6 μl of digestion product was loaded onto a 10% TBE polyacrylamide gel and run at 200V for 30 min. The gels were stained with ethidium bromide and quantified by densitometry using the ImageLab software suite (Bio-Rad) as previously described [166].

5.3.4. Fluorescence-activated cell sorting of myoblasts

DMD myoblasts were electroporated with 5 micrograms each of hCas9-T2A-GFP and sgRNA expression vectors and incubated at 37°C and 5% CO_2. Three days after electroporation, cells were trypsinized and collected for FACS sorting using a FACSvantage II sorting machine. GFP-positive cells were collected and grown for analysis.

5.3.5. PCR-based assay to detect genomic deletions

The exon 51 or exon 45-55 loci were amplified from genomic DNA by PCR (Invitrogen AccuPrime High Fidelity PCR kit) using primers flanking each locus. The flanking primers were CelI-CR1/2-F and CelI-CR5-R for exon 51 or CelI-CR6-F and CelI-CR36-R for exon 45-55 analysis (**Appendix C**). PCR products were separated on TAE-agarose gels and stained with ethidium bromide for analysis.

5.3.6. PCR-based detection of translocations

Loci with predicted possible translocations were amplified by a two-step nested PCR (Invitrogen AccuPrime High Fidelity PCR kit for each step) of genomic DNA from cells transfected with Cas9 alone (control) or Cas9 with sgRNA. In the first step, translocations that may occur at each on-target and off-target sgRNA target site were amplified by 35 cycles of PCR using combinations of Surveyor primers for each locus that were modified to include restriction sites to facilitate cloning and sequencing analysis (**Appendix C**). One microliter of each PCR reaction was subjected to a second

round of amplification by 35 rounds of PCR using nested primer sets custom designed for each individual predicted translocation (**Appendix C**). Each second nested PCR primer binds within the same approximate region within the primary amplicon; however, each pair was optimized using Primer3 online bioinformatics software to ensure specific detection of each translocation. PCR amplicons corresponding to the expected length of predicted translocations and only present in cells treated with sgRNA were purified (QIAGEN Gel Extraction kit) and analyzed by Sanger sequencing.

5.3.7. mRNA analysis

Immortalized myoblasts were differentiated into myofibers by replacing the growth medium with DMEM supplemented with 1% insulin-transferrin-selenium (Invitrogen #51500056) and 1% penicillin/streptomycin (Invitrogen #15140) for 5 days before the cells were trypsinized and collected. Total RNA was isolated from these cells using the RNeasy Plus Mini Kit (QIAGEN) according to the manufacturer's instructions. RNA was reverse transcribed to cDNA using the VILO cDNA synthesis kit (Life Technologies #11754) and 1.5 micrograms of RNA for 2 hours at 42°C according to the manufacturer's instructions. The target loci were amplified by 35 cycles of PCR with the AccuPrime High Fidelity PCR kit (Invitrogen) using primers annealing to exons 44 and 52 to detect exon 51 deletion by CR1/5 or CR2/5 or primers annealing to exons 44 and 60 to detect exon 45-55 deletion by CR6/36 (**Appendix C**). PCR products were run on TAE-

agarose gels and stained with ethidium bromide for analysis. The resolved PCR bands were cloned and analyzed by Sanger sequencing to verify the expected exon junctions.

5.3.8. Western blot analysis

To assess dystrophin protein expression, immortalized myoblasts were differentiated into myofibers as above for 6-7 days. Cells were trypsinized, collected and lysed in RIPA buffer (Sigma) supplemented with a protease inhibitor cocktail (Sigma) and the total protein amount was quantified using the bicinchoninic acid assay according to the manufacturer's instructions (Pierce). Samples were then mixed with NuPAGE loading buffer (Invitrogen) and 5% β-mercaptoethanol and heated to 85°C for 10 minutes. Twenty-five micrograms of protein were separated on 4-12% NuPAGE Bis-Tris gels (Invitrogen) with MES buffer (Invitrogen). Proteins were transferred to nitrocellulose membranes for 1-2 hours in 1X tris-glycine transfer buffer containing 10% methanol and 0.01% SDS. The blot was then blocked for 1 hour with 5% milk-TBST at room temperature. Blots were probed with the following primary antibodies: MANDYS8 to detect dystrophin (1:1000, Sigma D8168) or rabbit anti-GAPDH (1:5000, Cell Signaling 2118S). Blots were then incubated with mouse or rabbit horseradish peroxidase-conjugated secondary antibodies (Santa Cruz) and visualized using the ChemiDoc chemilumescent system (BioRad) and Western-C ECL substrate (BioRad).

5.3.9. Transplantation into immunodeficient mice

All animal experiments were conducted under protocols approved by the Duke Institutional Animal Care & Use Committee. Cells were trypsinized, collected and washed in 1X Hank's Balanced Salt Solution (HBSS, Sigma). Two million cells were pelleted and resuspended in five μL 1X HBSS (Sigma) supplemented with cardiotoxin (Sigma #C9759) immediately prior to injection. These cells were transplanted into the hind limb tibialis anterior (TA) muscle of NOD.SCID.gamma (NSG) mice (Duke CCIF Breeding Core) by intramuscular injection. Four weeks after injection, mice were euthanized and the TA muscles were harvested.

5.3.10. Immunofluorescence staining

Harvested TA muscles were incubated in 30% glycerol overnight at 4°C before mounting and freezing in Optimal Cutting Temperature compound. Serial 10 micron sections were obtained by cryosectioning of the embedded muscle tissue at -20°C. Cryosections were then washed in PBS to remove the OCT compound and subsequently blocked for 30-60 minutes at room temperature in PBS containing 10% heat-inactivated fetal bovine serum for spectrin detection or 5% heat-inactivated fetal bovine serum for dystrophin detection. Cryosections were incubated overnight at 4°C with the following primary antibodies that are specific to human epitopes only: anti-spectrin (1:20, Leica NCL-SPEC1) or anti-dystrophin (1:2, Leica NCL-DYS3). After primary staining, spectrin or dystrophin expression was detected using a tyramide-based immunofluorescence

signal amplification detection kit (Life Technologies, TSA Kit #22, catalog #T-20932,). Briefly, cryosections were incubated with 1:200 goat anti-mouse biotin-XX secondary (Life Technologies #B2763) in blocking buffer for one hour at room temperature. The signal was then amplified using streptavidin-HRP conjugates (1:100, from TSA Kit) in blocking buffer for one hour at room temperature. Finally, cryosections were incubated with tyramide-AlexaFluor488 conjugates (1:100, TSA kit) in manufacturer-provided amplification buffer for 10 minutes at room temperature. Stained cryosections were then mounted in ProLong AntiFade (Life Technologies #P36934) and visualized with conventional fluorescence microscopy.

5.3.11. Cytotoxicity assay

To quantitatively assess potential sgRNA or SpCas9 nuclease-associated cytotoxicity, HEK293T cells were transfected with 10 ng of a GFP reporter and 100 ng SpCas9 expression vector and 100 ng sgRNA expression vector using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). The percentage of GFP positive cells was assessed at 2 and 5 days by flow cytometry. The survival rate was calculated as the decrease in GFP positive cells from days 2 to 5 and normalized to cells transfected with an empty nuclease expression vector as described [155].

5.4. Results

5.4.1. Targeting CRISPR/Cas9 to Hotspot Mutations in the Human Dystrophin Gene

To utilize the CRISPR/Cas9 gene editing platform for correcting a wide range of dystrophin mutations, we created dozens of sgRNAs targeted to the hotspot mutation region between exons 45-55 (**Figure 18**). We selected the previously described *S. pyogenes* system that utilizes a human-codon optimized SpCas9 nuclease [10, 11] and a chimeric single-guide RNA (sgRNA) expression vector to guide efficient site-specific gene editing. Similar to our previous study targeting exon 51 with TALENs [42], we selected protospacers to target the 5′ and 3′ ends of exons 45 through 55 which meet the 5′-NRG-3′ PAM requirement of SpCas9. Small insertions or deletions created by NHEJ-based DNA repair within these exons can generate targeted frameshift mutations that address various dystrophin mutations surrounding each exon (**Figure 18a-b**). For example, CR3 was designed to correct dystrophin mutations or deletions surrounding exon 51 by introducing small insertions or deletions in the 5′ end of exon 51 to restore the downstream dystrophin reading frame (**Figure 18b**). Additionally, we designed sgRNAs to employ the multiplex capability of the CRISPR/Cas9 system and specifically delete individual exons or a series of exons to restore the dystrophin reading frame, similar to the methods of oligonucleotide-based exon skipping. For this purpose, sgRNAs were targeted to the intronic regions surrounding exon 51 (**Figure 18c**) or exons 45-55 (**Figure 18d**). These sgRNAs were intentionally targeted to sites nearest to the downstream or

upstream exon intended to be included in the resulting transcript to minimize the likelihood that the background patient deletion would include the intronic sgRNA target sites.

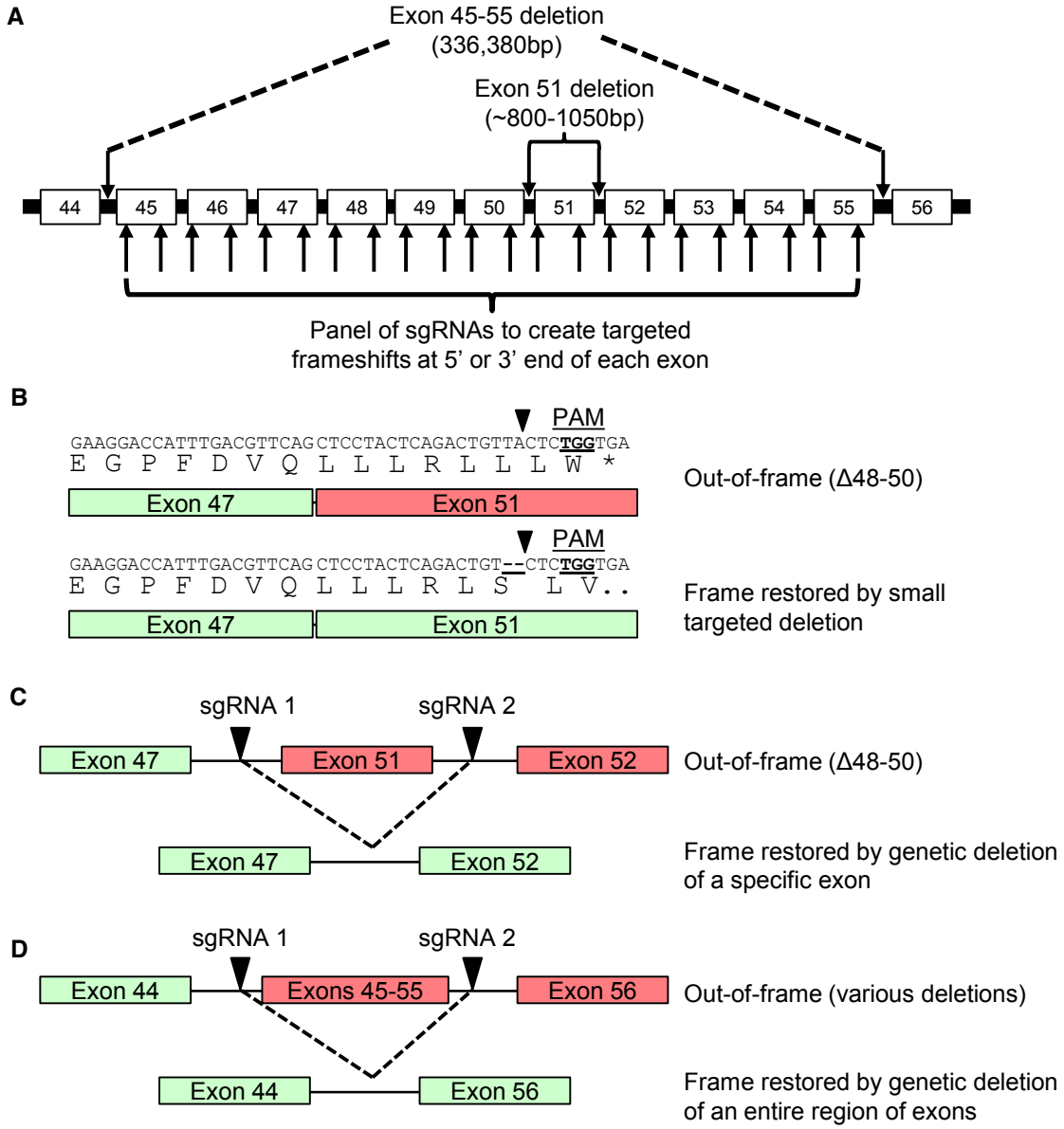


Figure 18: CRISPR/Cas9 targeting of the dystrophin gene. (a) sgRNA sequences were designed to bind sequences in the exon 45-55 mutational hotspot

region of the dystrophin gene, such that gene editing could restore dystrophin expression from a wide variety of patient-specific mutations. Arrows within introns indicate sgRNA targets designed to delete entire exons from the genome. Arrows within exons indicate sgRNA targets designed to create targeted frameshifts in the dystrophin gene. (b) Example of frame correction following introduction of small insertions or deletions by NHEJ DNA repair in exon 51 using the CR3 sgRNA. (c) Schematic of multiplex sgRNA targets designed to delete exon 51 and restore the dystrophin reading frame in a patient mutation with the deletion of exons 48-50. (d) Schematic of multiplex sgRNA targets designed to delete the entire exon 45-55 region to address a variety of DMD patient mutations.

5.4.2. Screening of sgRNAs Targeted to the Dystrophin Gene in Human Cells

We initially assessed gene editing frequency in the human HEK293T cell line to rapidly determine different sgRNA targeting efficiencies. As quantified by the Surveyor assay 3 days post-transfection, we found that 29/32 (~90%) of sgRNAs tested were able to mediate highly efficient gene modification at the intended locus (**Table 3, Appendix C**). The gene editing frequencies were stable for almost all of the sgRNAs (<25% signal change from day 3 to day 10, **Table 3, Appendix C**), indicating that gene editing mediated by each individual sgRNA was well-tolerated. A notable exception is CR33, which had no detectable activity at day 10, although activity may be below the sensitivity of the Surveyor assay (est. ~1%).

Table 3: Measured activity of sgRNAs in human cells. HEK293Ts were transfected with constructs encoding human codon-optimized SpCas9 and the indicated sgRNA. Each sgRNA was designed to modify the dystrophin gene as indicated. The frequency of gene modification at day 3 or day 10 post-transfection was determined by the Surveyor assay. The ratio of measured Surveyor signal at day 3 and day 10 was calculated to quantify the stability of gene editing frequencies for each sgRNA in human cells.

Target	sgRNA #	% modified alleles at day 3	% modified alleles at day 10	% change day 10/day 3
Multiplex deletion of exon 51				
Int 50	CR1	6.6	9.3	41.8
Int 50	CR2	10.3	14.0	36.2
Ex 51	CR4	11.9	14.4	21.3
Int 51	CR5	12.4	13.3	7.8
Multiplex deletion of exons 45-55				
Int 44	CR6	16.1	16.9	4.3
Int 44	CR33	1.3	<1	n.d.
Int 44	CR34	13.2	11.0	-16.6
Int 55	CR7	6.8	7.1	5.3
Int 55	CR35	22.5	20.9	-7.1
Int 55	CR36	26.4	24.7	-6.4
Targeted frameshifts				
Ex 45	CR10	14.9	16.3	9.3
Ex 45	CR11	<1	<1	n.d.
Ex 46	CR12	<1	<1	n.d.
Ex 46	CR13	16.9	18.4	9.2
Ex 47	CR14	17.2	17.6	2.9
Ex 47	CR15	15.4	15.3	-0.9
Ex 48	CR16	11.5	10.9	-5.0
Ex 48	CR17	<1	<1	n.d.
Ex 49	CR18	1.8	2.2	20.1
Ex 49	CR19	33.7	38.4	13.9
Ex 50	CR20	14.9	13.7	-7.6
Ex 50	CR21	24.1	20.8	-13.5
Ex 51	CR3	13.0	16.7	28.0
Ex 51	CR31	18.9	16.9	-10.2
Ex 52	CR22	25.9	20.3	-21.6
Ex 52	CR23	25.2	24.0	-4.8
Ex 53	CR24	24.8	23.6	-4.6
Ex 53	CR25	2.6	2.9	9.5
Ex 54	CR26	24.5	22.0	-10.1
Ex 54	CR27	13.4	12.6	-5.9
Ex 55	CR28	21.6	19.8	-8.4
Ex 55	CR29	19.2	19.6	2.2

5.4.3. Enrichment of Gene-Edited Cells Using a Fluorescence-Based Reporter System

We next sought to use selected sgRNAs to correct specific mutations in DMD patient myoblast cell lines. After transfection into DMD myoblasts, we observed unexpectedly low or undetectable gene modification activity as measured by the Surveyor assay (**Figure 19c**, bulk population). Therefore, we used flow cytometry to select for transfected cells co-expressing GFP through a 2A ribosomal skipping peptide linked to the SpCas9 protein (**Figure 19a**), similar to previously described methods to enrich gene-modified cell populations using fluorescent reporters [179, 180]. Importantly, the addition of this fluorescent reporter to the SpCas9 expression vector did not seem to significantly impact gene editing activity in HEK293T cells (**Figure 19b**). A low percentage of transfected myoblasts (~0.5-2%) expressed the fluorescent reporter at 3 days after electroporation, despite high transfection efficiencies of control GFP expression plasmids (typically >70%, **Figure 19d,** pmaxGFP). Given the high levels of CRISPR/Cas9 activity in the easily transfected HEK293T line, inefficient transgene expression after electroporation of SpCas9-T2A-GFP and sgRNA constructs into the DMD cells may explain the low observed gene editing efficiencies in unsorted cells. After sorting the GFP-positive DMD myoblasts, we observed a substantial increase in detectable activity at most sgRNA target loci (**Figure 19c**). Therefore, all subsequent

experiments used cells sorted for SpCas9 expression by expression of this fluorescent reporter.

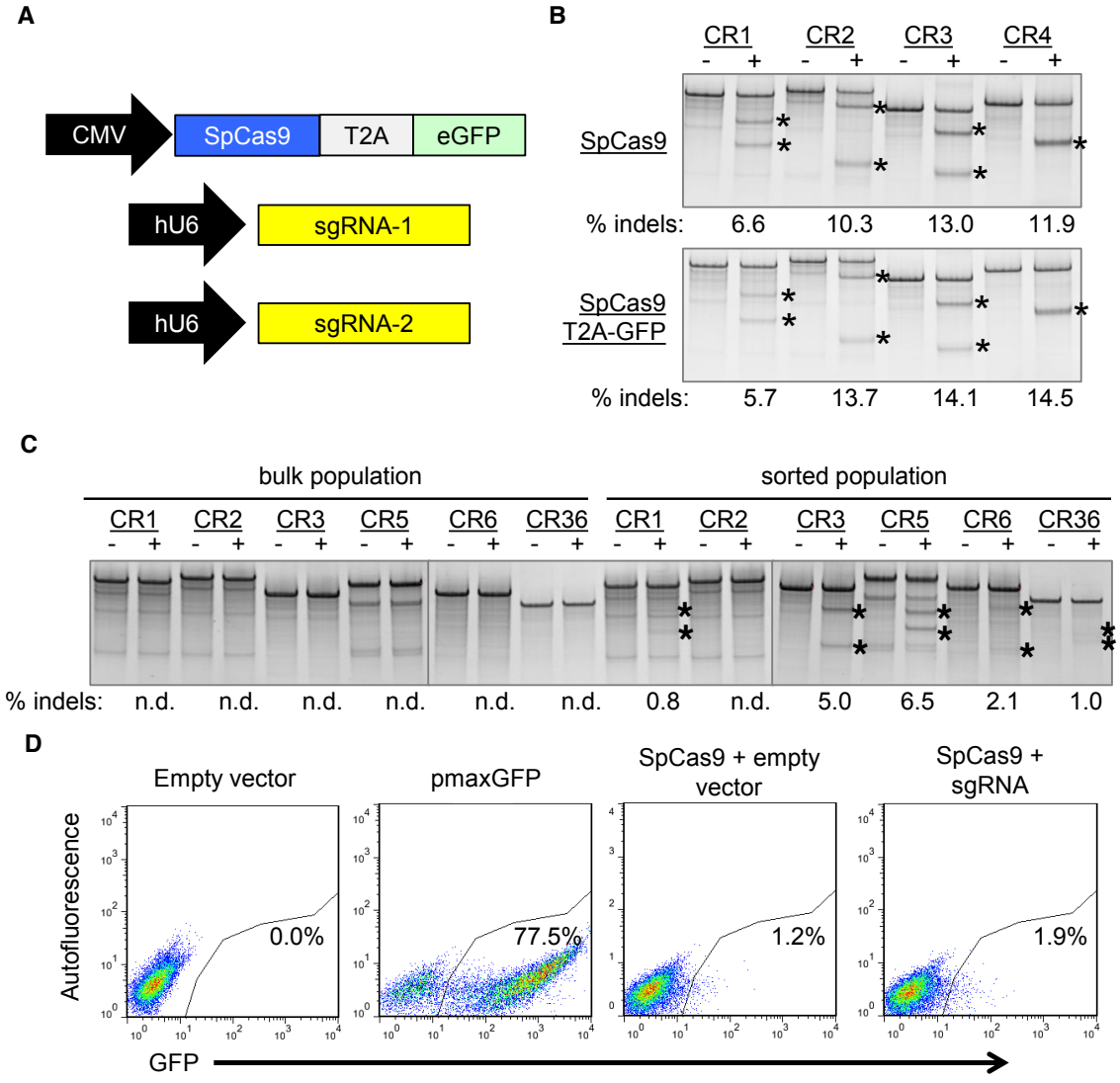


Figure 19: Fluorescence-activated flow sorting to enrich genetically modified DMD myoblasts. (a) A plasmid expressing a human-codon optimized SpCas9 protein linked to a GFP marker using a T2A ribosomal skipping peptide sequence was co-electroporated into human DMD myoblasts with one or two plasmids carrying sgRNA expression cassettes. (b) The indicated sgRNA expression cassettes were independently co-transfected into HEK293Ts with a separate plasmid expressing SpCas9 with (bottom) or without (top) a GFP marker linked to SpCas9 by a T2A

ribosomal skipping peptide sequence. Gene modification frequencies were assessed at 3 days post-transfection by the Surveyor assay. (c) DMD myoblasts with deletions of exons 48-50 in the dystrophin gene were treated with sgRNAs that correct the dystrophin reading frame in these patient cells. Gene modification was assessed at 20 days post-electroporation in unsorted (bulk) or GFP+ sorted cells. (d) GFP expression in DMD myoblasts 3 days after electroporation with indicated expression plasmids. Transfection efficiencies and sorted cell populations are indicated by the gated region.

5.4.4. Restoration of Dystrophin Expression by Targeted Frameshifts

We have shown previously that small insertions and deletions created by NHEJ DNA repair can be used to create targeted frameshifts to correct aberrant reading frames [42]. Similar to this approach, we designed a sgRNA, CR3, to restore the dystrophin reading frame by introducing small insertions and deletions within exon 51 (**Figures 18b, 20a**). The types of insertions and deletions generated by CRISPR/Cas9 at this locus were assessed by Sanger sequencing of alleles from the genomic DNA of HEK293T cells co-transfected with expression plasmids for SpCas9 and the CR3 sgRNA (**Figure 20b**). Notably, the insertions and deletions resulted in conversion to all three reading frames (**Figures 20b,c**), consistent with our previous results using TALENs [42]. To demonstrate genetic correction in a relevant patient cell line, expression plasmids for SpCas9 and the CR3 sgRNA were electroporated into a DMD myoblast line with a deletion of exons 48-50 that is correctable by creating frameshifts in exon 51. The treated cells were sorted, verified to have gene modification activity by the Surveyor assay (CR3, **Figure 20c** sorted population), and differentiated into myotubes to test for restored dystrophin expression. Expression of dystrophin protein was observed concomitant with the

detectable nuclease activity (**Figure 20d**). Taken together with the data from **Table 3**, the *S. pyogenes* CRISPR/Cas9 system presents a powerful method to quickly generate targeted frameshifts to address a variety of patient mutations and restore expression of the human dystrophin gene.

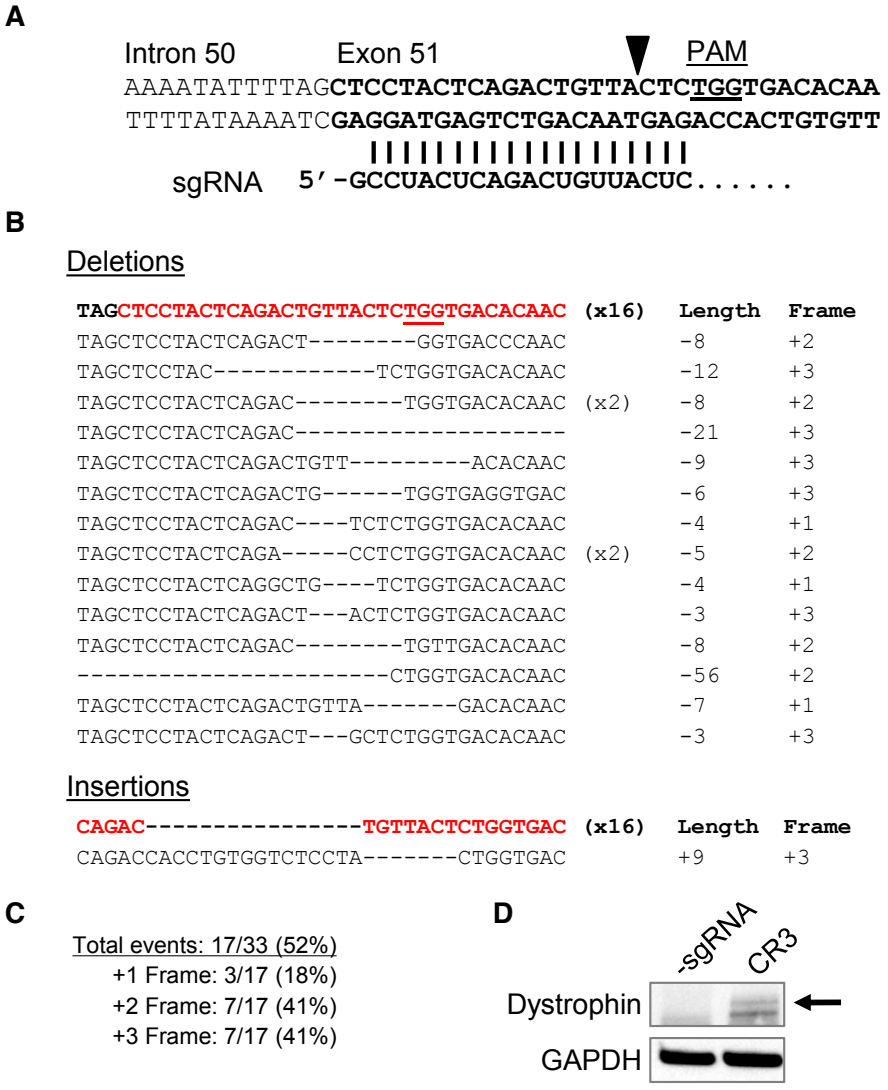


Figure 20: Targeted frameshifts to restore the dystrophin reading frame using CRISPR/Cas9. (a) The 5′ region of exon 51 was targeted using a sgRNA, CR3, that binds immediately upstream of the first out-of-frame stop codon. PAM: protospacer-

adjacent motif. (b) The exon 51 locus was PCR amplified from HEK293T cells treated with SpCas9 and CR3 expression cassettes. Sequences of individual clones were determined by Sanger sequencing. The top sequence (bolded, exon in red) is the native, unmodified sequence. The number of clones for each sequence is indicated in parentheses. (c) Summary of total gene editing efficiency and reading frame conversions resulting from gene modification shown in (b). (d) Western blot for dystrophin expression in human DMD myoblasts treated with SpCas9 and the CR3 sgRNA expression cassette (Figure 19c) to create targeted frameshifts to restore the dystrophin reading frame. Dystrophin expression was probed using an antibody against the rod-domain of the dystrophin protein after 6 days of differentation.

5.4.5. Multiplex CRISPR/Cas9 Gene Editing Mediates Genomic Deletion of Exon 51 and Rescues Dystrophin Protein Expression

The multiplexing capability of the CRISPR/Cas9 system presents a novel method to efficiently generate genomic deletions of specific exons for targeted gene correction. DMD patient myoblasts with background deletions correctable by exon 51 skipping were treated with two combinations of sgRNAs flanking exon 51 (CR1/CR5 or CR2/CR5) and sorted to enrich for gene-edited cells as in **Figure 19**. As detected by end-point PCR of the genomic DNA from these treated cells, the expected genomic deletions were only present when both sgRNAs were electroporated into the cells with SpCas9 (**Figure 21a**). Sanger sequencing confirmed the expected junction of the distal chromosomal segments (**Figure 21b**) for both deletions. After differentiating the sorted myoblasts, a deletion of exon 51 from the mRNA transcript was detected only in the cells treated with both sgRNAs (**Figure 21c**). Finally, restored dystrophin protein expression was detected in the treated cells concomitant with observed genome- and mRNA-level deletions of exon 51 (**Figure 21d**).

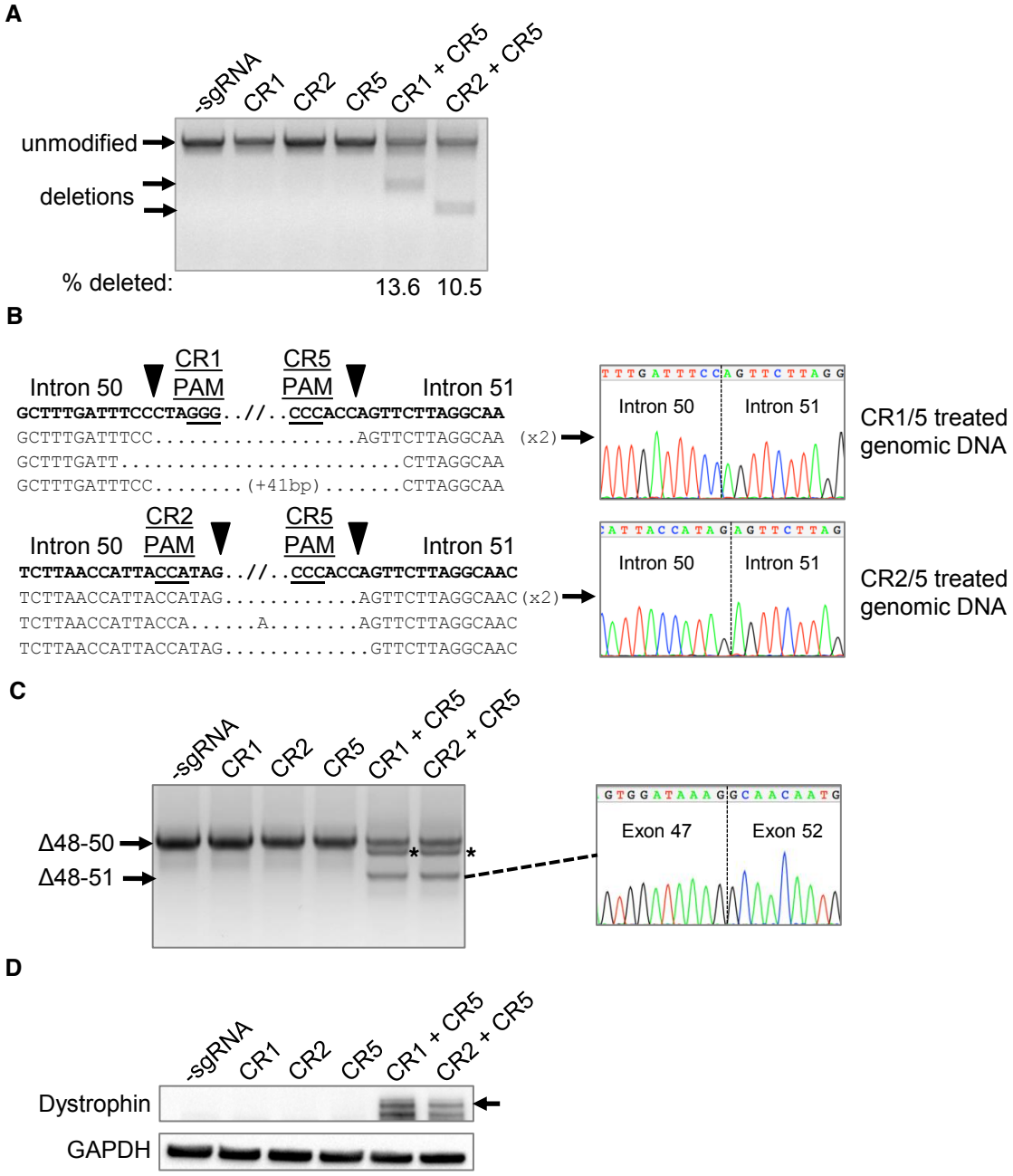


Figure 21: Deletion of exon 51 from the human genome using multiplex CRISPR/Cas9 gene editing. (a) End-point genomic PCR across the exon 51 locus in human DMD myoblasts with a deletion of exons 48-50. The top arrow indicates the expected position of full-length PCR amplicons and the two lower arrows indicate the expected position of PCR amplicons with deletions caused by the indicated sgRNA combinations. (b) PCR products from (a) were cloned and individual clones were

sequenced to determine insertions and deletions present at the targeted locus. The top row shows the wild-type unmodified sequence and the triangles indicate SpCas9 cleavage sites. At the right are representative chromatograms showing the sequences of the expected deletion junctions. (c) End-point RT-PCR analysis of dystrophin mRNA transcripts in CRISPR/Cas9-modified human Δ48-50 DMD myoblasts treated with the indicated sgRNAs. A representative chromatogram of the expected deletion PCR product is shown at the right. Asterisk: band resulting from hybridization of the deletion product strand to the unmodified strand. (d) Rescue of dystrophin protein expression by CRISPR/Cas9 genome editing was assessed by western blot for the dystrophin protein with GAPDH as a loading control. The arrow indicates the expected restored dystrophin protein band.

5.4.6. Dystrophin Rescue by a Multi-Exon Large Genomic Deletion

Although addressing patient-specific mutations is a powerful use of the CRISPR/Cas9 system, it would be advantageous to develop a single method that can address a myriad of common patient deletions. For example, a promising strategy is to exclude the entire exon 45-55 region as a method to correct up to 62% of known patient deletions [4, 176, 177]. Therefore we tested if multiplex CRISPR/Cas9-based gene editing may be able to generate efficient deletion of the exon 45-55 locus in human cells. After transfection into HEK293T cells, we detected the expected deletion of ~336,000 bp by PCR of the genomic DNA (**Figure 22a**). Similarly, we were able to detect this deletion by PCR of the genomic DNA from SpCas9/sgRNA-treated DMD patient cells harboring a background deletion of exons 48-50 of unknown length (**Figure 22a**). Sanger sequencing of this deletion band from the genomic DNA of treated DMD cells revealed the expected junctions of intron 44 and intron 55 immediately adjacent to the sgRNA target sites (**Figure 22b**). After differentiation of treated DMD cells, the expected deletion of exons 45-55 was detected in the dystrophin mRNA transcript and verified to be a fusion of

exons 44 and 56 by Sanger sequencing (**Figure 22c**). Restored protein expression was observed by western blot in the sorted cell populations containing the CRISPR/Cas9-induced deletion of exons 45-55 from the genome and resulting mRNA transcripts (**Figure 22d**). These data demonstrate that multiplex CRISPR/Cas9 editing presents a single universal method to restore the dystrophin reading frame in more than 60% of DMD patient mutations.

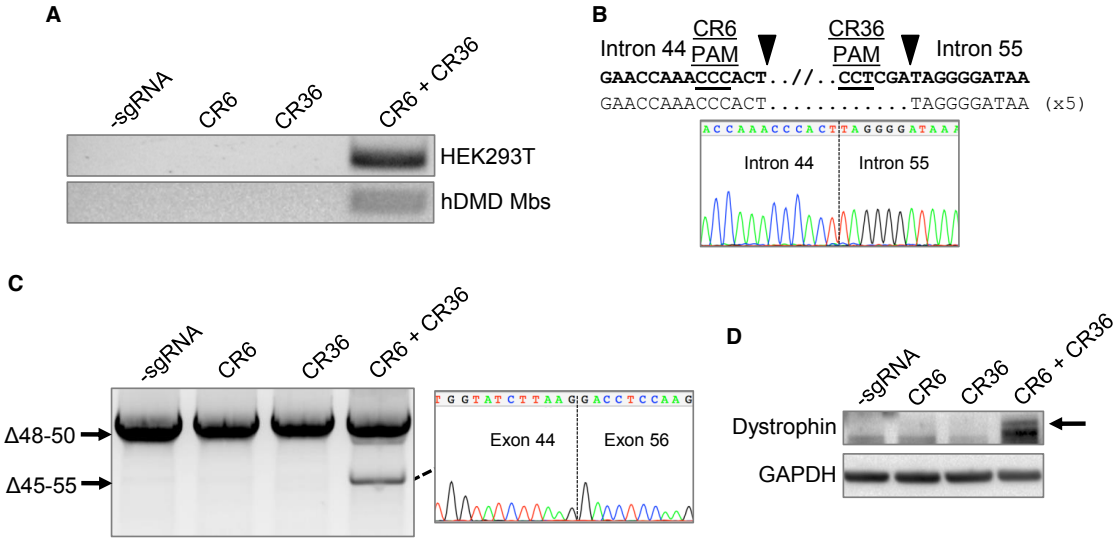


Figure 22: Deletion of the entire exon 45-55 region in human DMD myoblasts by multiplex CRISPR/Cas9 gene editing. (a) End-point genomic PCR of genomic DNA to detect deletion of the region between intron 44 and intron 55 after treating HEK293Ts or DMD myoblasts with the indicated sgRNAs. (b) Individual clones of PCR products of the expected size for the deletions from DMD myoblasts in (a) were analyzed by Sanger sequencing to determine the sequences of genomic deletions present at the targeted locus. Below is a representative chromatograms showing the sequence of the expected deletion junctions. (c) End-point RT-PCR analysis of dystrophin mRNA transcripts in CRISPR/Cas9-modified human Δ48-50 DMD myoblasts treated with the indicated sgRNAs. A representative chromatogram of the expected deletion PCR product is shown at the right. (d) Analysis of restored dystrophin protein expression by western blot following electroporation of DMD myoblasts with sgRNAs targeted to intron 44 and/or intron 55.

5.4.7. Transplantation of Corrected Myoblasts into Immunodeficient Mice

A promising method for DMD therapy is to correct a population of autologous patient muscle progenitor cells that can be engrafted into the patient's skeletal muscle tissue to rescue dystrophin expression [19, 36, 181]. To demonstrate the ability of the corrected cells to express human dystrophin *in vivo*, we transplanted a population of DMD myoblasts that were treated with sgRNAs CR1 and CR5, which flank exon 51, and sorted for expression of GFP as before (**Figure 23**). After 4 weeks, muscle fibers positive for human spectrin, which is expressed by both corrected and uncorrected cells, were detected in cryosections of injected muscle tissue (**Figures 24-25**). A number of these fibers were also positive for human dystrophin with expression localized to the sarcolemma, demonstrating functional gene correction in these cells (**Figures 24-25**). No fibers positive for human dystrophin were observed in sections from mice injected with the untreated DMD myoblasts (**Figures 24-25**), indicating that the CRISPR/Cas9-modified cells were the source of human dystrophin expression.

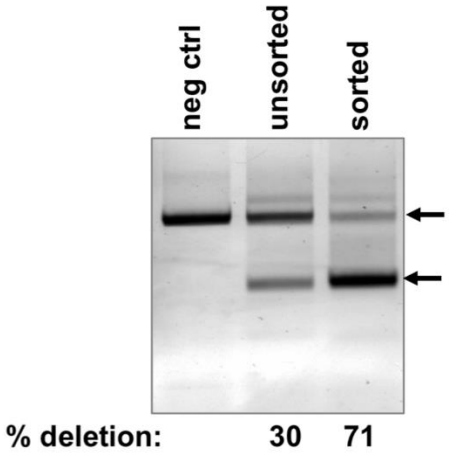


Figure 23: Verification of flow cytometry-based enrichment of gene-modified DMD myoblasts used for *in vivo* cell transplantation experiment. DMD myoblasts were treated with Cas9 with or without sgRNA expression vectors for CR1 and CR5 and sorted for GFP+ cells by flow cytometry. Deletions at the exon 51 locus were detected by end-point PCR using primers flanking the locus. Neg ctrl: DMD myoblasts treated with Cas9 only and sorted for GFP+ cells.

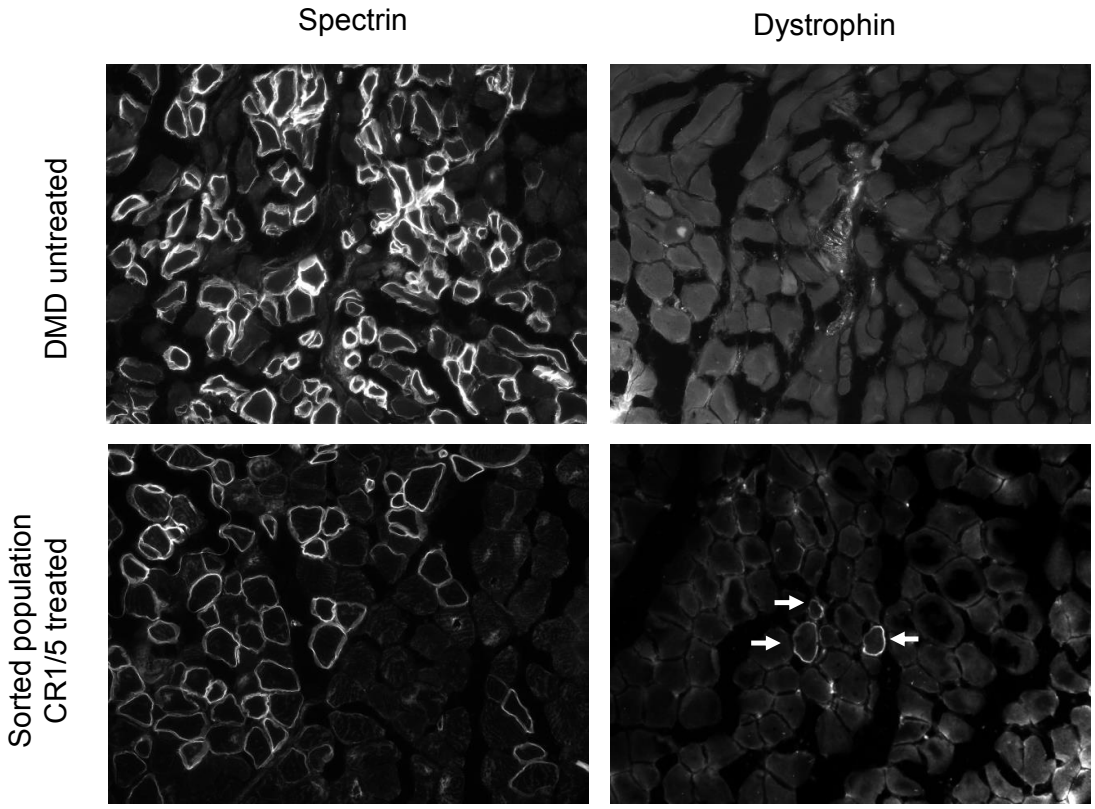


Figure 24: Expression of restored human dystrophin *in vivo* following transplantation of CRISPR/Cas9-treated human DMD myoblasts into immunodeficient mice. Human Δ48-50 DMD myoblasts were treated with SpCas9, CR1, and CR5 to delete exon 51 and sorted for GFP expression as shown in Figure 19. These sorted cells and untreated control cells were injected into the hind limbs of immunodeficient mice and assessed for human-specific protein expression in muscle fibers after 4 weeks post-transplantation. Cryosections were stained with anti-human spectrin, which is expressed by both uncorrected and corrected myoblasts that have fused into mouse myofibers, or anti-human dystrophin antibodies as indicated. White arrows indicate muscle fibers positive for human dystrophin.

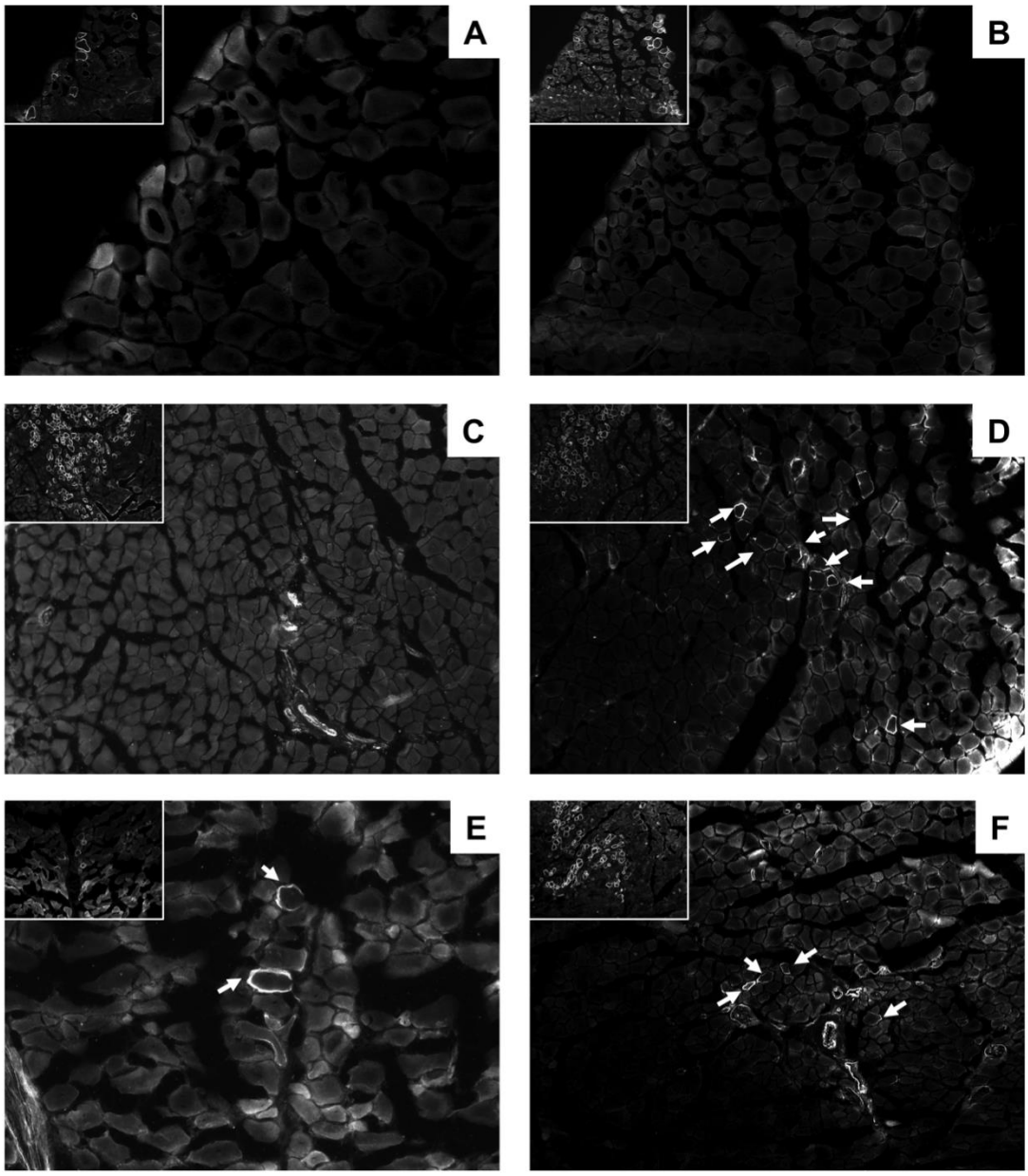


Figure 25: Additional immunofluorescence images probing human dystrophin expression. Serial sections from regions stained with anti-human spectrin are shown inset in top left. (a-c) Sections from muscles injected with untreated human DMD myoblasts. (d-f) Sections from muscles injected with CR1/5 treated human DMD myoblasts enriched by flow cytometry. White arrows indicate dystrophin positive fibers.

5.4.8. Off-target and Cytotoxicity Analysis

We assessed the relative cytotoxicity of the CRISPR/Cas9 system in human cells for select sgRNAs by adapting a flow cytometry-based GFP retention assay as previously described [42]. Minimal cytotoxicity was observed for SpCas9 co-expressed with or without sgRNAs after transfection into human cells (**Figure 26a**). Several recent studies have documented activity of the CRIPSR/Cas9 system at off-target loci in human cells [137-142, 174]. Publicly available tools are available to assess and prioritize potential CRISPR/Cas9 activity at off-target loci based on predicted positional bias of a given mismatch in the sgRNA protospacer sequence and the total number of mismatches to the intended target site [137]. We used this public webserver to predict the most likely off-target sites for the sgRNAs used to correct the dystrophin gene in this study **(Table 4**). The top ten potential off-target sites were assessed by the Surveyor assay in HEK293T cells treated with SpCas9 and individual sgRNA expression cassettes for CR1, CR3, CR5, CR6, or CR36. CR1, CR3 and CR36 each had one of these ten predicted off-target loci demonstrate significant levels of gene modification, consistent with other studies investigating the specificity of the *S. pyogenes* CRISPR/Cas9 system [137-142] (**Table 4 and Appendix C**). Interestingly, the CR3 off-target sequence had substantial homology and similar modification frequencies to the intended on-target (9.3% at OT-1 vs. 13.3% at intended site (**Table 4 and Appendix C**). Notably, CR3-OT1 was the only one of these three off-target sites to show significant levels of activity in the

sorted hDMD cells by the Surveyor assay (**Figure 26b**). Thus, our selected sgRNAs have relatively favorable specificities, however, we cannot rule out activity at predicted off-target loci that may exist below the sensitivity of the Surveyor assay or off-target activity at other loci that were not assessed here.

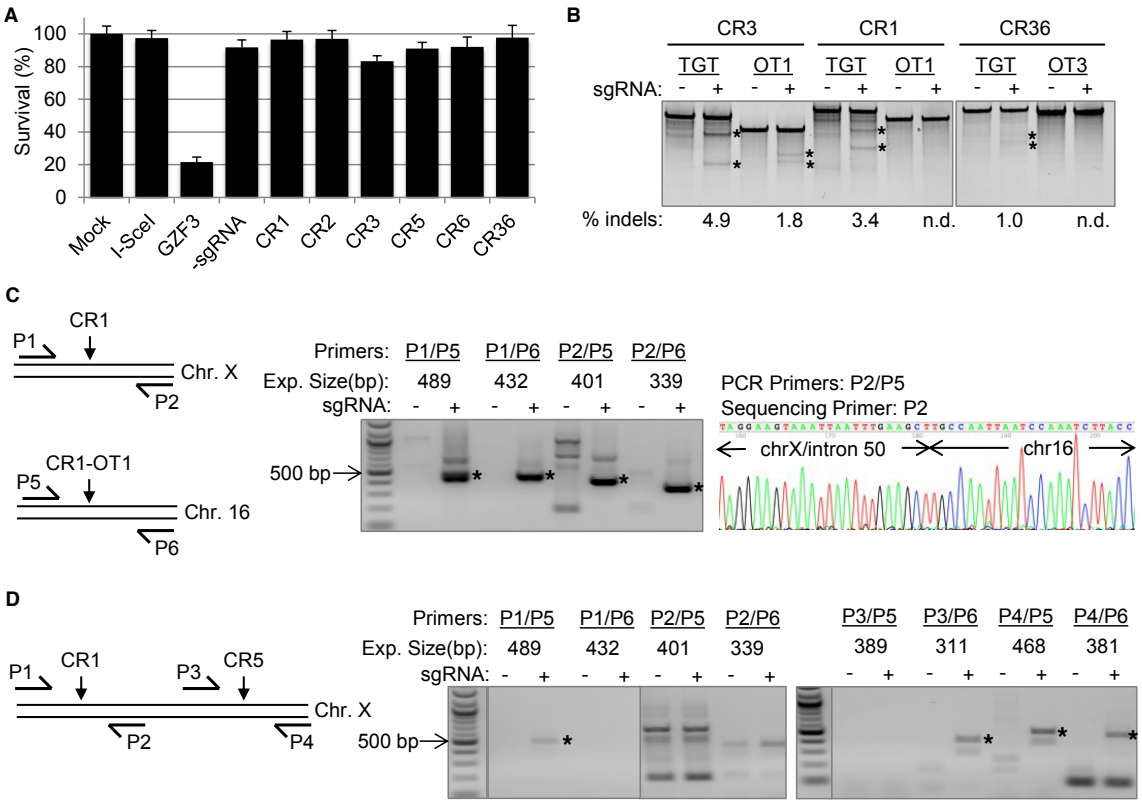


Figure 26: Evaluation of CRISPR/Cas9 toxicity and off-target effects for CR1/CR5-mediated deletion of exon 51 in human cells. (a) Results of a cytotoxicity assay in HEK293T cells treated with human-optimized SpCas9 and the indicated sgRNA constructs. Cytotoxicity is based on survival of GFP-positive cells that are co-transfected with the indicated nuclease. I-SceI is a well-characterized non-toxic meganuclease and GZF3 is a known toxic zinc finger nuclease. (b) Surveyor analysis at off-target sites in sorted hDMD cells treated with expression cassettes encoding Cas9 the indicated sgRNAs. These three off-target sites tested in hDMD cells were identified from a panel of 50 predicted sites tested in HEK293T cells (Table 4). TGT: on-target locus for indicated sgRNA. OT:off-target locus. (c, d) End-point nested PCR to detect chromosomal translocations in (c) HEK293T cells treated with Cas9 and CR1 or (d) sorted hDMD cells treated with Cas9, CR1, and CR5. The schematic depicts the relative location of nested primer pairs customized for each translocation event. The expected size of each band was estimated based on the primer size and the location of the predicted sgRNA cut site at each locus. Asterisks indicate bands detected at the expected size. The identities of the bands in (c) were verified by Sanger sequencing

from each end (Figure 29). A representative chromatogram for the P2/P5 translocation in HEK293T cells is shown.

Nuclease activity at off-target sites may cause unintended chromosomal rearrangements by distal re-ligation between cleaved target and off-target loci on distinct chromosomes [174, 182]. This presents a significant concern for deletion-based gene correction strategies due to the increased potential for off-target activity by using two or more nucleases, such as in multiplex CRISPR/Cas9 gene editing. We probed for potential translocations using a highly sensitive nested genomic PCR assay to detect translocations at the validated off-target loci (**Table 4**) during both single and multiplex CRISPR/Cas9 editing strategies. Using this assay, we readily detected translocations between on-target and off-target sites in the model HEK293T cell line that also shows high levels of off-target activity (**Figures 26c, 27-29**). Sanger sequencing of the PCR amplicons confirmed the identity of the predicted translocation event for each primer pair (**Figures 28-29**). A subset of the translocations detected in the HEK293T cells were also detectable by nested PCR in the sorted hDMD myoblasts, although the signal was considerably weaker and we were unable to confirm the sequence identity due to low yield of product (**Figures 26d and 27a,c**). Notably, we did not detect translocations using this assay in HEK293T or sorted hDMD cells treated with CR6 or CR6/CR36, respectively (**Figure 27**), that had low levels of off-target activity at CR6-OT3 only in HEK293T cells (**Table 4**). These results underscore the importance of selecting highly specific sgRNAs, particularly for multiplex editing applications, and show that this

approach can benefit from ongoing efforts to improve the specificity of the CRISPR/Cas9 system [139, 142, 143, 173, 174]. However, taken together with the cytotoxicity and stable gene editing activities observed in **Table 3**, these data suggest that the selected sgRNAs are able to correct the dystrophin gene without significant toxicity and with only a single strongly predicted off-target site with detectable levels of activity.

Table 4: Summary of top 10 off target sites predicted *in silico* and activity at each site as detected by the Surveyor assay in HEK293T cells transfected with Cas9 and the indicated sgRNA expression cassettes. n.d.: not detected.

Guide	Target	Sequence	PAM	Score	Chr	Gene	Intron /Exon	# MMs	% indels
CR3	Guide	GCCTACTCAGACTGTTACTC	-	-	-	-	-	-	-
	Target	tCCTACTCAGACTGTTACTC	TGG	-	X	DMD	Exon	1	13.0
	OT1	tCCTACTCAcACTGTTACTC	AGG	7.4	1	STRIP1	Intron	2	9.3
	OT2	aCCTgCTCAcACTGTTACTC	CAG	2.5	2	ARHGAP25	Intron	3	n.d.
	OT3	GCaTtCTCAaACTGTTACTC	AGG	2.4	13	None	None	3	n.d.
	OT4	GgaTtCTCAcACTGTTACTC	GGG	1.3	14	PGPEP1	Exon	4	n.d.
	OT5	aCaTACTtAtACTGTTACTC	TAG	1.3	19	MDGA2	Intron	4	n.d.
	OT6	tatTcCTaAGACTGTTACTC	AAG	0.9	8	LPPR1	Intron	5	n.d.
	OT7	aaggACTaAGACTGTTACTC	GGG	0.9	9	RNF122	Intron	5	n.d.
	OT8	GagctCTCAtACTGTTACTC	TAG	0.8	3	DNMBP	Exon	5	n.d.
	OT9	GCaaAaTgAGACTGTTACTC	CAG	0.8	5	SLC12A2	Intron	4	n.d.
	OT10	cCtcAtTCAGACTGTTACTC	AAG	0.8	4	KCNIP4	Intron	4	n.d.
CR1	Guide	GATTGGCTTTGATTTCCCTA	-	-	-	-	-	-	-
	Target	cATTGGCTTTGATTTCCCTA	GGG	-	X	DMD	Intron	1	8.3
	OT1	aATTGGCATTGATTTCCCTA	GAG	7.1	16	None	None	2	0.8
	OT2	cATTGGCTTTaATTTCCCTA	TAG	4.8	4	None	None	2	n.d.
	OT3	GATaGGCTgTGATTTCCCTA	GAG	3.9	9	None	None	2	n.d.
	OT4	GAaTaGCcTTGATTTCCCTA	AAG	2.4	1	None	None	3	n.d.
	OT5	aATTtGCTTTGATTTCCCTg	AGG	1.5	1	TIMM17A	Intron	3	n.d.
	OT6	GATgtGCTTTGATTTCCCTt	GGG	1.4	17	MYO1D	Intron	3	n.d.
	OT7	aATTGGtTTTaATTTCCCTA	AAG	1.1	8	PIK1A	Intron	3	n.d.
	OT8	aATTGGgTTTGATTTCCCTt	TGG	1.1	11	MS4A1	Intron	3	n.d.
	OT9	GATgGGtTTTtATTTCCCTA	GAG	1.0	11	None	None	3	n.d.
	OT10	GAaTGGtTTTGATTTCCCTg	GAG	1.0	11	None	None	3	n.d.
CR5	Guide	GCAGTTGCCTAAGAACTGGT	-	-	-	-	-	-	-
	Target	aCAGTTGCCTAAGAACTGGT	GGG	-	X	DMD	Intron	1	14.0
	OT1	cCAGTTGtCTAAGAACTGGg	GAG	1.5	5	NRG1	Intron	3	n.d.
	OT2	GCAGTTGCCTgtGAACTGGT	AGG	1.4	X	None	None	2	n.d.
	OT3	GCAGaTGCagAAGAACTGGT	GAG	1.4	19	SMIM7	Intron	3	n.d.
	OT4	GCAGTTcCagAAGAACTGGT	GAG	0.9	11	GLB1L2	Intron	3	n.d.
	OT5	caAcTTGCCTAtGAACTGGT	AGG	0.7	8	ASAP1	Intron	4	n.d.
	OT6	aCAccTGCCTAAGAACTGGa	GGG	0.7	11	None	None	4	n.d.
	OT7	tCAGgTGgCTAAGAACTGGg	TGG	0.7	14	NIN	Intron	4	n.d.
	OT8	GaAGTTGgCcAAGAACTGGa	GAG	0.6	7	None	None	4	n.d.
	OT9	GCtGcTGCCcAAGAACTGGc	AGG	0.6	11	AMOTL1	Intron	4	n.d.
	OT10	tCAGcTGgCTAAGAACgGGT	AAG	0.6	7	ACTR3C	Intron	4	n.d.
CR6	Guide	GGGGCTCCACCCTCACGAGT	-	-	-	-	-	-	-
	Target	aGGGCTCCACCCTCACGAGT	GGG	-	X	DMD	Intron	1	19.9
	OT1	GcaGCTCagCCCTCACGAGT	CAG	0.8	3	None	None	4	n.d.
	OT2	GGGGCTtCAgCaTCACGAGT	GAG	0.8	8	None	None	3	n.d.
	OT3	GGGGCTCtcCCCTCACtAGT	GAG	0.6	8	None	None	3	n.d.
	OT4	GGGGaTCCACCtTCACcAGT	CAG	0.6	2	None	None	3	n.d.
	OT5	aGGGCTggACCCTCACaAGT	AAG	0.4	16	AXIN1	Intron	4	n.d.
	OT6	tGGtCTCCtCCCcCACGAGT	GGG	0.4	2	None	None	4	n.d.
	OT7	aGGGCTCCcaCCcCACGAGT	GAG	0.3	5	None	None	4	n.d.
	OT8	GaGGCTCCAtaCTCACcAGT	GAG	0.3	11	None	None	4	n.d.
	OT9	GGaGCTgCcCCtTCACGAGT	GGG	0.3	3	None	None	4	n.d.
	OT10	atGaCTCCACCCTCAaGAGT	AAG	0.3	8	AGPAT5	None	4	n.d.
CR36	Guide	GCCTTCTTTATCCCCTATCG	-	-	-	-	-	-	-
	Target	GCCTTCTTTATCCCCTATCG	AGG	-	X	DMD	Intron	0	20.6
	OT1	GtCTgCTgTgTCCCCTATCG	GGG	1.3	21	None	None	4	n.d.
	OT2	cCCTTCTcTATCCCCTgTCG	TGG	1.3	8	None	None	3	n.d.

	OT3	GCCTTCTTTATCCCCTcTCt	TGG	0.9	10	None	None	2	0.5
	OT4	GCgcTCTTTtTCCCCTATCt	TAG	0.6	16	None	None	4	n.d.
	OT5	GCCcTCTgTcTCCCCTgTCG	CAG	0.5	1	NFASC	None	4	n.d.
	OT6	tCCATCTtTgTCCCCTATtG	AGG	0.5	10	None	None	4	n.d.
	OT7	aCCtTCTCTcTCCCCTATaG	AGG	0.5	5	LOC100996485	Intron	4	n.d.
	OT8	GttTTCTTTtTCCCCTATgG	GAG	0.5	3	None	None	4	n.d.
	OT9	tgCTTCTTaATCCCCTATCa	AAG	0.4	7	None	None	4	n.d.
	OT10	aCCTTCTTacTCCCCTATCc	GGG	0.4	10	ADARB2	None	4	n.d.

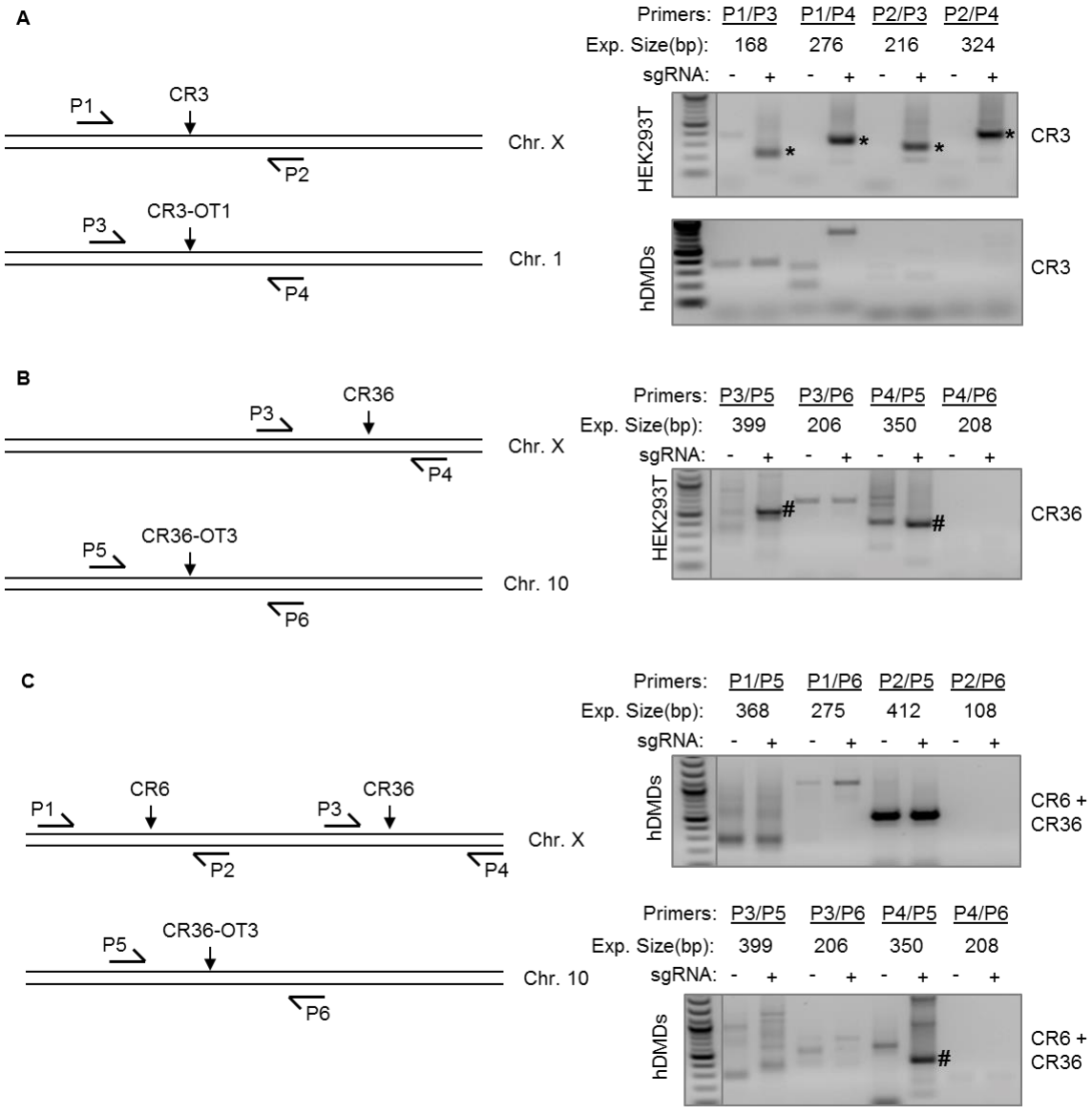


Figure 27: End-point nested PCR to detect chromosomal translocations caused by CRISPR/Cas9 off-target activity for CR3 and CR6/CR36 in human cells. Nested end-point PCR analysis was used to detect translocations in (a) HEK293T or sorted hDMD cells treated with Cas9 and CR3 as indicated, (b) HEK293T cells treated with Cas9 and CR36 alone, or (c) sorted hDMD cells treated with Cas9, CR6, and CR36 expression cassettes. The second nested PCR reaction for translocation was amplified using custom primers for each predicted translocation locus to maximize specificity

(See Appendix C). The schematic depicts the relative location of nested primer pairs used to probe for the presence of translocations. Each possible translocation event was first amplified from genomic DNA isolated from cells treated with or without the indicated sgRNA(s). A second nested PCR reaction was performed using primers within the predicted PCR amplicons that would result from translocations. Expected size was estimated based on the indicated primer binding site and the predicted sgRNA cut site at each locus. *indicates bands detected at the expected size and verified by Sanger sequencing from each end. #indicates amplicons in which Sanger sequencing showed sequences other than the predicted translocation, likely a result of mispriming during the nested PCR.

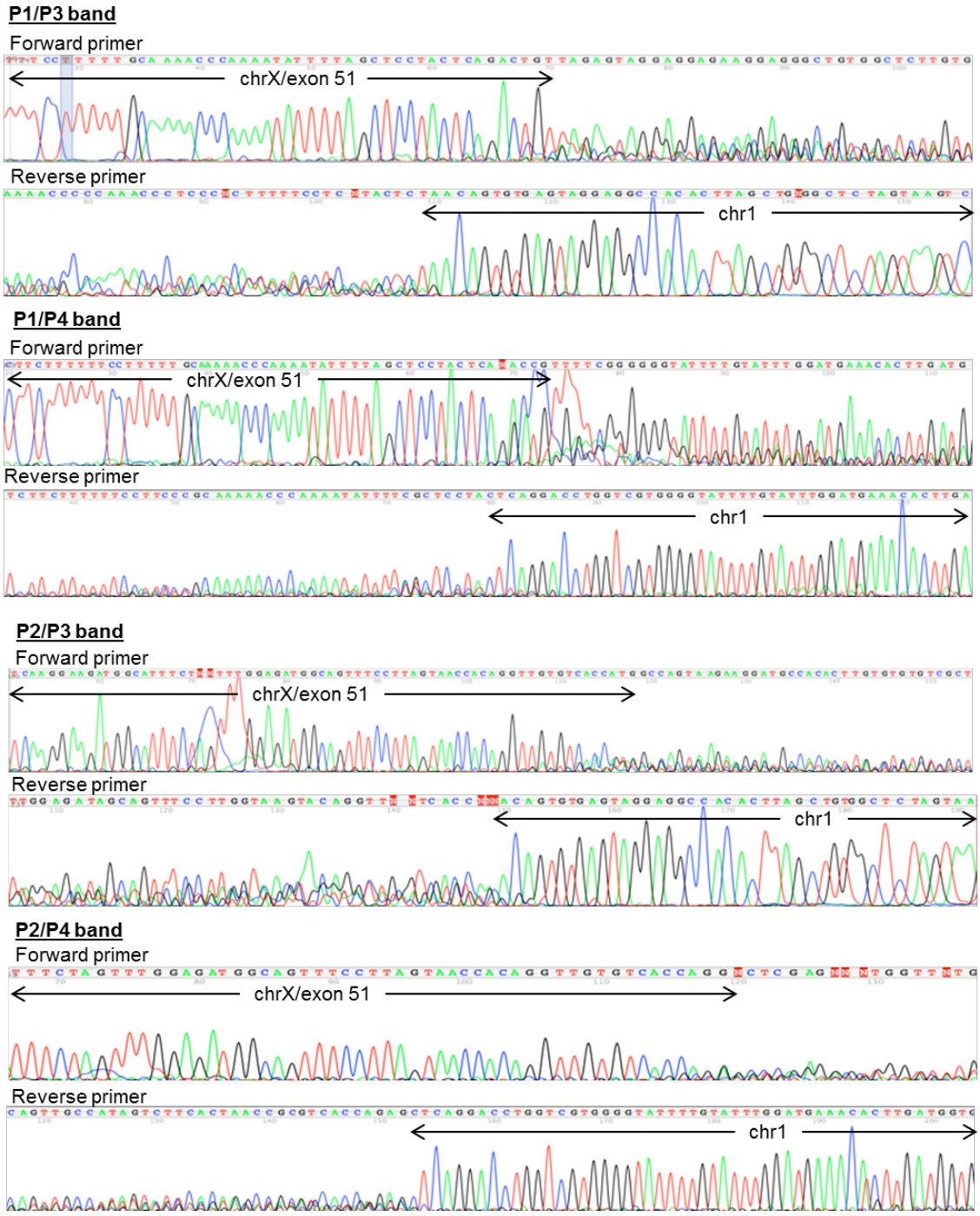


Figure 28: Sanger sequencing chromatograms for bands detected in Figure 27 resulting from translocations between CR3 and CR3-OT1, on chromosomes X and 1, respectively, in HEK293T cells treated with Cas9 and CR3 gene cassettes. Arrows

show regions of homology to the indicated chromosome nearby the expected break points caused by the appropriate sgRNAs. Note that sequencing reads become out of phase near the break point due to the error-prone nature of DNA repair by non-homologous end-joining.

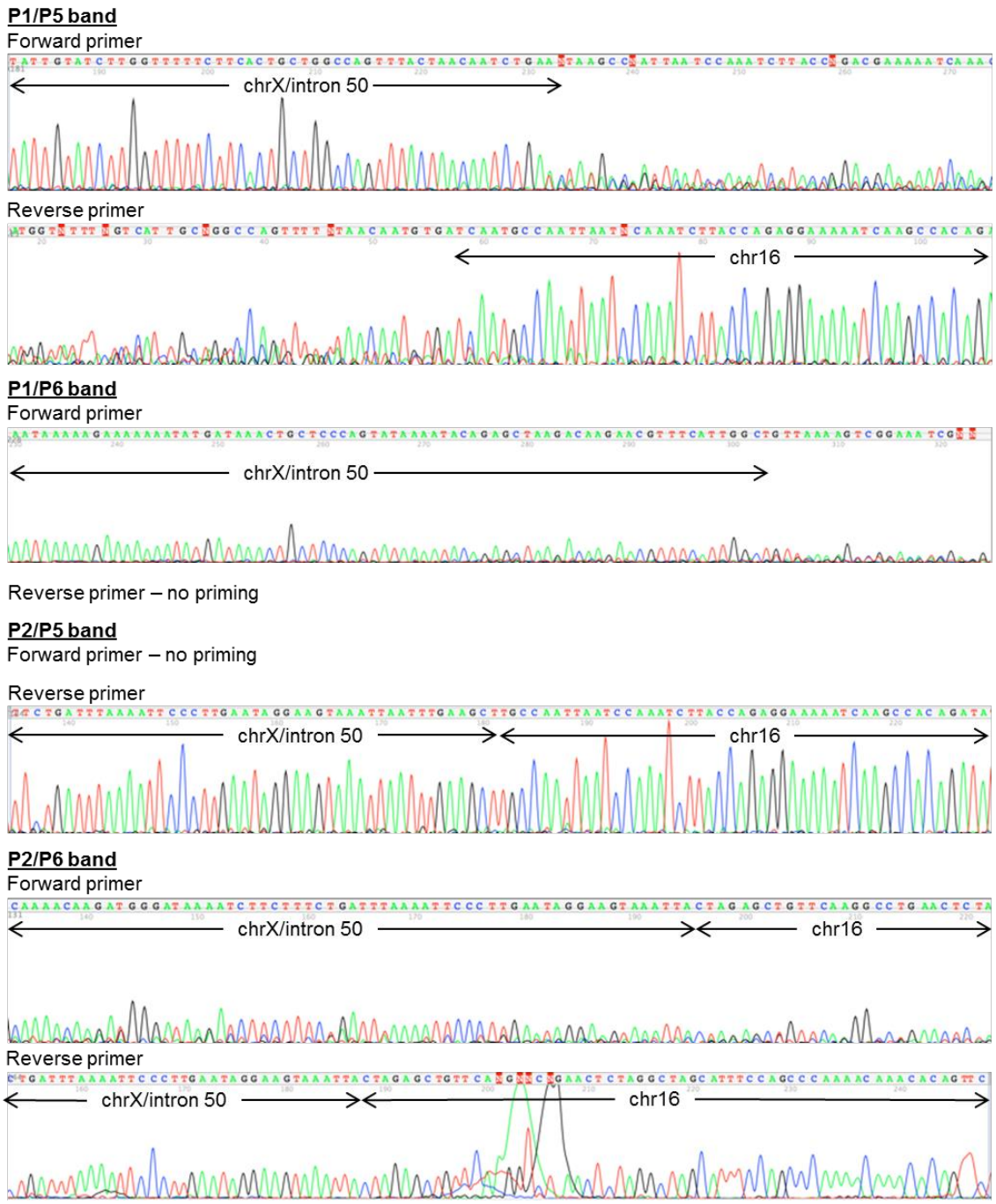


Figure 29: Sanger sequencing chromatograms for bands detected in Figure 26c resulting from translocations between CR1 and CR1-OT1, on chromosomes X and 16, respectively, in HEK293T cells treated with Cas9 and CR1 gene cassettes. Arrows

show regions of homology to the indicated chromosome nearby the expected break points caused by the appropriate sgRNAs. Note that sequencing reads become out of phase near the break point due to the error-prone nature of DNA repair by non-homologous end-joining.

5.5. Discussion

Genome editing is a powerful tool for correcting genetic disease and the recent development of the CRISPR/Cas9 system is dramatically accelerating progress in this area. Here we demonstrate the correction of DMD, the most common genetic disease that also currently has no approved therapeutic options. Many gene- and cell-based therapies for DMD are in preclinical development and clinical trials, and genome editing methods are compatible with many of these approaches. For example, genome editing may be combined with patient-specific cell-based therapies for DMD [19, 36, 181]. The CRISPR/Cas9 system has been previously demonstrated to function in human pluripotent stem cells [100, 103] and other human cell lines [10, 11, 99, 101], as well as human skeletal myoblasts in this study. Furthermore, genome editing of autologous primary T cells with ZFNs is currently in clinical trials (NCT01252641,NCT00842634 and NCT01044654) [13]. Importantly, gene editing with CRISPR/Cas9 did not abolish the myogenic capacity of these cells, as demonstrated by efficient dystrophin *expression in vitro* and *in vivo* after transplantation into immunodeficient mice. Thus, this strategy should be compatible with cell-based therapies for DMD that are under development, although achieving efficient cell engraftment is still a challenge [19, 36, 181]. Additionally, direct transfection of the sgRNA and Cas9 mRNA, in contrast to the

plasmid-based delivery method used here, may increase specificity and safety by reducing the duration of Cas9 expression and eliminating the possibility of random plasmid integration. Alternatively, delivery of the CRISPR/Cas9 system directly to skeletal and/or cardiac muscle by viral, plasmid, or RNA delivery vectors is a promising strategy for *in vivo* genome editing and translation of this approach [5, 19, 183]. The large size of *S. pyogenes* Cas9 gene (~4.2 kilobases) presents a challenge to its use in size-restricted adeno-associated viral vectors. However, Cas9 genes from other species [136], such as *N. meningitidis* and *S. thermophilus,* are short enough to efficiently package both Cas9 and sgRNA expression cassettes into single AAV vectors for *in vivo* gene editing applications, as has been done with the smaller ZFNs [81, 87].

The *S. pyogenes* CRISPR/Cas9 system enabled efficient modification of nearly 90% of tested targets, consistent with other reports of robust activity of this system at diverse loci. The robustness and versatility of this technology is a significant advancement towards at-will creation of patient-specific gene editing. However, further enhancing gene editing frequencies and minimizing off-target activity remain as major challenges to customized clinical applications. Notably, recent studies have shown that low levels of dystrophin, including as little as 4% of wild-type expression, are sufficient to improve survival, motor function, and cardiac function in a mouse model [184-186]. Therefore the levels of CRISPR/Cas9 activity reported in this study may be sufficient for therapeutic benefit, although we suspect that optimized delivery and expression vehicles, as well as

ongoing improvements to the CRISPR/Cas9 system [173], will lead to significant enhancements of gene correction levels.

The use of multiplexing with CRISPR/Cas9 to delete exons also presents a unique set of opportunities and challenges. While Chapter 4 explores deletion of essential exon splicing sequences, this is the first study to report the deletion of entire exons from the genome to restore dystrophin expression. Previous efforts restored the reading frame of the dystrophin gene with small indels generated by NHEJ-based DNA repair following the action of a single nuclease [42, 90]. A primary benefit of this new approach is that the protein product of the edited gene is predictable and already characterized in Becker muscular dystrophy patients with the naturally occurring deletion, in contrast to the random indels created by intraexonic action of a single nuclease that will lead to the creation of novel epitopes from each DNA repair event. Furthermore, the product resulting from the exon deletions will lead to restored dystrophin for every successful gene editing event, whereas modifying the gene with random indels within exons will only restore the reading frame in the one-third of editing events that leads to the correct reading frame. Additional studies on enhancing and/or biasing gene repair towards distal chromosomal re-ligation will be helpful, particularly for long deletions such as the excision of the 336 kb containing exons 45-55 demonstrated in this study. Moreover, the introduction of multiple DSBs increases the requirement for stringent specificity of the gene editing reagents in order to minimize the opportunity for the unintended

chromosomal rearrangements that were readily detected in the treated HEK293T cells (**Figures 26c and 27-29**).

We observed that all of the sgRNAs tested were not associated with significant cytotoxic effects in human cells. We identified three potential off-target sites out of 50 total tested sites for the five sgRNAs used here to restore dystrophin expression. Furthermore, chromosomal translocations between the intended on-target sites and these off-target sites were detectable by highly sensitive nested PCR assays in HEK293T cells expressing high levels of Cas9 and sgRNAs. This is consistent with previous work showing translocations with two nucleases intentionally targeted to distinct chromosomes [182]. These results are also consistent with other studies that have characterized off-target activity and translocations generated by SpCas9 [137-142, 174] and compare favorably with other gene editing systems, including ZFNs and TALENs, which are sometimes cytotoxic [187] and are known to also act at off-target sites [12]. Notably, the off-target activity and translocations identified in HEK293T cells, which is an immortalized and aneuploid cell line that expresses very high levels of Cas9 and sgRNA, did not occur at as high a level and in some cases were undetectable in the hDMD myoblasts. This corroborates previous studies showing that lower levels of Cas9 and sgRNA can reduce off-target effects [137, 140]. Importantly, this level of specificity may be tolerable given the severity of DMD, the lack of an apparent cytotoxic effect in human cells, and the absence of adverse events in ongoing clinical trials for genome

editing with ZFNs (NCT01252641, NCT00842634 and NCT01044654) [13]. Furthermore, improvements to the specificity of the CRISPR/Cas9 system, including the use of dual nickases, truncated sgRNAs, and careful sgRNA selection methods [139, 142, 143, 174], are also under development. CRISPR/Cas9 specificity may be further enhanced by rational design or directed evolution of the *S. pyogenes* sgRNA and/or Cas9 nuclease, or characterization of novel Cas9 nucleases from other species with more stringent sgRNA or PAM requirements [136].

This study demonstrates that CRISPR/Cas9 genome editing is a robust, easily programmable method to rapidly generate targeted frameshifts or genomic deletions that can address a variety of dystrophin mutations. Importantly, this method can reproducibly modify the dystrophin gene by deleting exons 45-55, thereby addressing more than half of DMD patient deletions with a single genome editing strategy. Further advancements in the delivery, specificity, and efficiency of these reagents will enhance the utility of this method for correcting the dystrophin gene and creating other custom genetic modifications. Thus, CRISPR/Cas9 genome editing offers an exciting new avenue for gene therapies to treat Duchenne muscular dystrophy and other hereditary disorders.

Chapter 6: Summary and Future Studies

6.1. Overview

This thesis capitalizes on the rapid advancement of genome editing tools to create novel molecular therapies for Duchenne muscular dystrophy. Genetic correction of the native dystrophin gene is a powerful method to permanently restore dystrophin expression that is compatible with many leading cell and gene therapies for this disorder. Presently, there is still a need to find effective therapies for DMD and gene correction at the genomic level is a promising alternative to transient methods such as AAV gene transfer of minidystrophin genes or small oligonucleotides to rescue the reading frame of the dystrophin gene.

Chapter 3 explores the use of TALEN technologies to generate intraexonic small insertions and deletions that restore the dystrophin reading frame. First, several combinations of TALEN pairs were generated and tested for efficient gene modification in human cells. An optimal TALEN pair was then transfected into several human cell lines derived from DMD patients. These cell lines included myoblasts and fibroblasts, two clinically relevant cell sources for cell-based therapies for DMD. TALEN gene modification was able to restore the expression of dystrophin in clonally derived patient myoblasts, and in treated bulk populations of myoblasts as well as fibroblasts that were reprogrammed into the myogenic lineage by forced overexpression of the master

myogenic transcription factor, MyoD. Importantly, exome sequencing of clonally derived DMD myoblasts revealed no off-target effects caused by these TALENs.

Building on this work, we also wanted to create homogeneous changes to the corrected dystrophin protein. Chapter 4 describes a strategy to generate genetic deletions of sequences essential to exon splicing as a method to delete exon 51 and restore the dystrophin reading frame. To do this, ZFNs were generated and screened for chromosomal activity. Two ZFN pairs were transfected into DMD patient derived myoblasts and a single clonal population was isolated containing the expected genetic deletion. This genetic deletion resulted in the loss of exon 51 from the mRNA transcript and restoration of dystrophin expression. Furthermore, these corrected cells were able to engraft *in vivo* following transplantation into immunodeficient mice and express human dystrophin that is correctly localized to the sarcolemma membrane. Thus, this study demonstrates that genetic corrections can result in restored dystrophin functionality and that ZFN-mediated gene modification can be achieved without a significant impact of the myogenic capabilities of targeted cells.

In Chapter 5, we expanded on this approach using the recently described CRISPR/Cas9 system to rapidly generate intraexonic small insertions and deletions, single exon deletions, and multiple exon deletions across a mutational hotspot in the dystrophin gene. Importantly, the *S. pyogenes* CRISPR/Cas9 gene editing platform is able to generate gene modifications with a high rate of success, thereby allowing expansion

of the proof-of-principle studies in Chapters 3 and 4. Thus, we were able to generate unique nucleases targeting the 5′ and 3′ ends of nearly every exon between the mutational hotspot of exons 45-55, enabling a patient-specific therapeutic strategy that is otherwise difficult and time-consuming to achieve with other gene editing technologies. Furthermore, we utilized the unique multiplex editing capabilities of this system to rapidly and efficiently delete exon 51 or exons 45-55. Using these approaches, we were able to restore dystrophin expression *in vitro* in human DMD patient myoblasts and *in vivo* after transplantation of a bulk corrected population of hDMD myoblasts. Taken together, these approaches can address over 60% of patient mutations. Notably, multiplex deletion of exons 45-55 presents a universal approach to correct these >60% of mutations and generate a predictable protein product that is known to have high functionality in some Becker muscular dystrophy patients [176, 188], a less severe form of DMD. We also show that there is a need to improve the specificity of these approaches, as we were able to readily detect off-target cleavage at unintended genomic loci that resulted in chromosomal rearrangements in human cells.

Nuclease-mediated gene correction is an effective method to restore the dystrophin gene. This thesis demonstrates that the reading frame of the dystrophin gene can be corrected in a variety of methods, including small insertions and deletions and large deletions of one or more exons. We show that ZFNs, TALENs, and CRISPR/Cas9 are efficient systems for generating genetic modifications in complex genomes.

However, the large size and repetitive sequence composition of TALENs presents a significant challenge for delivering these gene-editing tools *in vivo*. Therefore, ZFNs and CRISPR/Cas9 are excellent candidates for clinical translation as genetic tools to cure DMD. A recent clinical trial established that ZFN-modified cells met safety criteria gene editing platform after *ex vivo* gene modification and transplantation of autologous ZFN-modified cells [13]. Chapter 5 demonstrates that CRISPR/Cas9 can efficiently delete one or more entire exons, thereby restoring a clinically relevant, predictable protein. The multiplexing capability and straightforward sgRNA design of CRISPR/Cas9 gene editing is a highly efficient method to introduce genetic deletions. In particular, genetic deletion of exon 51 is attractive due to the high efficiency of deletion with CRISPR/Cas9 and the positive results in ongoing oligonucleotide-based transient exon 51 skipping clinical trials. Further work remains to increase the specificity of these gene editing platforms and to develop strategies for cell-based therapies or gene transfer of gene editing tools to restore dystrophin *in vivo.*

6.2. Improving the efficiency of gene correction

A major factor that will influence the therapeutic efficacy of genomic intervention in DMD is the percentage of muscle fibers expressing restored dystrophin protein. A notable set of studies have investigated the effect of dystrophin expression levels on phenotypic outcome in a mosaic mouse model of DMD that expresses variable levels of dystrophin to explore. From these studies, it has been suggested that

dystrophin levels greater than 4-12% of normal can ameliorate skeletal muscle histopathology and cardiomyopathic symptoms in mosaic mice that randomly express normal or mutant dystrophin [184, 186], while dystrophin levels of greater than 20% are necessary to protect muscle from exercise-induced damage [189]. A study investigating phenotypic severity in Becker muscular dystrophy patients, which have variable levels of expressed dystrophin, demonstrates that 10% or less of normal dystrophin levels universally results in a high disease severity, including early onset of loss of ambulation, cardiac involvement, and death [21]. Thus, genetic correction strategies that correct >4% of genomes will provide therapeutic benefit and protection, with a secondary goal of achieving optimal gene correction rates of >20% of treated nuclei to ensure complete therapeutic protection.

In Chapter 3, TALENs mediated *in situ* modification in 6.8-39.3% of DMD myoblasts, of which 1/3 of total events are expected to result in frame correction, thereby creating an effective correction rate of 2.2-13.1% of cells. In Chapter 4, we were able to generate an isogenic genetically corrected population that can engraft into host muscle tissue to restore human dystrophin expression. A limitation of this study is that clonally corrected cell populations must be efficiently delivered to dystrophic muscle *in vivo*. This is a significant challenge, requiring numerous injections in each muscle and would be particularly difficult for engrafting corrected cells in the heart. However, as discussed in Chapter 2, there have been significant advances in iPSC and dystrophic muscle

homing progenitors that may be compatible with this approach. Finally, Chapter 5 demonstrates a variety of gene modifications, including efficient deletion of exon 51 in 10.5-13.6% of DMD myoblasts (each deletion event is expected to result in genetic correction) and modification of intraexonic targets in 1.8-33.7% of treated human cells (0.6-11.2% total corrective events). Together, the methods presented in Chapters 3-5 are able to generate *in situ* genetic correction rates of 0.6-13.6% *in vitro* in well controlled conditions. Future studies will need to explore the efficiency of gene modifications *in vivo* following delivery of targeted nucleases to skeletal and cardiac muscle.

The genome editing strategies presented in this work mediate highly efficient correction, easily surpassing the 4% of treated genomes needed for therapeutic effect. However, improving specificity of these approaches will be important to minimize unintended genotoxicity that may be caused by engineered nucleases. Currently, ZFNs, TALENs, and CRISPR/Cas9 are capable of generating single-site double-strand breaks in >50% of treated cells, but this is heavily locus- or reagent-dependent. Moving forward, the discovery or engineering of improved gene editing systems and approaches can be used to increase single-site gene editing as well as genetic deletions to >20% correction levels that will provide complete protection, particularly for large deletions such as the 330kb deletion of exons 45-55 in Chapter 5. For example, there have been a number of recent studies improving the delivery, activity, and/or specificity of ZFNs [172], TALENs [103], and CRISPR/Cas9 [142, 190]. Furthermore, next generation gene editing tools, such

as recombinases [109-111, 191], may improve genome editing efficiencies as well. Finally, efficient delivery methods to deliver gene editing enzymes, such as adeno-associated virus, or novel cell-based methods such as dystrophic muscle-homing progenitors or iPSC-derived progenitors [29] will need to be optimized to reach protective levels of restored dystrophin expression. AAV is of significant interest for gene transfer to muscle and there has been significant progress in engineering novel AAV serotypes with improved tropism to skeletal and cardiac muscle [53, 192-196]. Importantly, these studies demonstrate that customized AAV vectors can be developed to evade pre-existing vector immunity and optimize vector biodistribution to muscle. Despite this, genome editing is an exciting method to generate autologous cell therapies for DMD, and could be realized as a custom therapeutic that allows extensive *ex vivo* characterization and safety profiling before administration to the patient.

6.3. Functional analysis of restored dystrophin proteins

DMD patient mutations that disrupt the dystrophin reading frame are commonly deletions of one or more non-essential exons. Chapters 3-5 explore restoration of the dystrophin reading frame around these mutational hotspot of the dystrophin gene that accounts for >60% of patient mutations. Additional work is needed to characterize the resulting proteins generated by these methods. As shown in Chapter 5, genome editing offers the exciting possibility of deleting large regions of the dystrophin gene at will, thereby creating the possibility to characterize of the function of a range of predefined,

internally deleted dystrophin proteins through generation of isogenic myoblast cell lines. These modified isogenic cell populations could be combined with novel *in vitro* analyses [197] to assess the effect of internal dystrophin deletions on skeletal muscle functions such as passive and contractile force as well as calcium regulation. Moreover, genome editing can be used to rapidly create animal models [162] that could be applied to investigate phenotypic outcomes, such as reduction in centrally nucleated fibers, increased muscle contractility, and reduction of creatine kinase levels, in whole organisms with genetically modified dystrophin genes, including mouse models engineered to express the entire 2.4 megabase human dystrophin gene [198].

While this information will be valuable, it is known that there are variable phenotypic outcomes in Becker muscular dystrophy (BMD) patient siblings with the same underlying mutations in the dystrophin gene [21, 22]. This observation suggests that there is uncertain functionality in internally deleted dystrophin proteins that will be patient-dependent. However, analysis of BMD patients with mutations similar to those created by exon skipping, and by extension genome editing, suggests that restoration of internally deleted dystrophin proteins alone will have significant therapeutic benefit [22]. Thus, *in vitro* analysis in human DMD patient cells and *in vivo* analysis in model organisms will provide significant insight into protein function, but may not fully elucidate absolute phenotypic outcomes. Nevertheless, these efforts will help identify

optimal genomic editing outcomes with strongly predicted functionality of the corrected dystrophin proteins.

6.4. In vivo *genetic correction of native dystrophin gene in* mdx *and humanized DMD mouse models*

One of the most promising gene therapy strategies for DMD is to introduce functional dystrophin expression in dystrophic tissue using the small nonpathogenic adeno-associated virus. AAV serotype 8 (AAV8) vectors have demonstrated high efficiency genome editing in *in vivo* applications to deliver ZFNs to liver tissue [81, 87]. AAV serotypes, such as AAV1, 6, 8, or 9, have high muscle-specific tropism that can be exploited to direct *in situ* genetic correction by delivering engineered nucleases directly to muscle tissue *in vivo*. Thus, there is an opportunity to deliver gene editing nucleases directly to muscle tissue *in vivo* to restore dystrophin expression by *in situ* correction. The canonical *mdx* animal model of DMD carries a premature stop codon in exon 23 of the dystrophin gene. Importantly, this exon can be skipped from the transcript without disrupting the reading frame [199], and is therefore an excellent proof-of-principle model to test genetic deletion strategies to delete exons and restore dystrophin expression similar to the approaches in Chapters 4 and 5. To expand on this approach, a humanized DMD animal model has also been developed that carries the entire 2.4 megabase human dystrophin gene [198] and will be useful in studies investigating gene targeting of the human dystrophin gene. However, at present, these animal models express only functional human dystrophin genes and do not recapitulate the mutations

or phenotype of DMD. Thus, humanized DMD animal models are currently useful only to investigate gene modification of the human dystrophin gene *in vivo* rather than serving as a therapeutic model of the human disease.

Further work should focus on identifying optimal conditions for creating genetic modifications in skeletal and cardiac tissue. Two major parameters to investigate are the impact of AAV serotype and dose on gene corrections in muscle and non-muscle types. An important consideration for these studies will be to monitor the ratio of on-target to off-target activity to balance safety and efficacy. The use of muscle-specific promoters will be useful in limiting non-muscle gene correction and off-target effects, however these promoters will need to retain activity in all dystrophin-expressing cell types, including skeletal and cardiac tissue. Preliminary work suggests evidence that NHEJ-like gene repair is evident in non-dividing skeletal muscle (see **Figure 30c**, Chapter 6.5). Further studies will be necessary to evaluate the overall robustness of NHEJ repair pathways and also to investigate if homology-directed repair (HDR) DNA repair pathways are functional in non-dividing muscle fibers. Finally, it is presently not known if AAV can efficiently transduce muscle progenitors to create a self-renewing pool that can repopulate genetically corrected muscle fibers following damage and repair. Gene editing offers a novel method to probe the ability of AAV to efficiently transduce muscle progenitors *in vivo* because it leaves behind a permanent genetic imprint after expression

of the engineered nuclease. While not comprehensive, these studies will be critical to translate engineered nucleases as effective genomic therapies to treat DMD.

6.5. Targeted addition of dystrophin to predefined safe harbor genomic loci

Our work demonstrates efficient restoration of the mutant dystrophin reading frames within a mutational hotspot that represents >60% of patient mutations. However, other patient mutations can result in the loss of essential exons or large regions of the dystrophin gene that impair any functionality of a frame restored dystrophin gene. In these cases, restoration of the reading frame would be predicted to have little or no effect on restoring the functionality of dystrophin in DMD patients. Functional replacement of the dystrophin gene by expression of exogenous dystrophin protein may be an attractive complementary method to address these rare mutations. To date, these efforts include expression of functional miniaturized or reassembled full-length dystrophin genes by transient adeno-associated virus gene transfer [48, 60, 200] and by permanent knock-in to the CCR5 safe harbor locus [31], or by genomic integration of full-length dystrophin gene cassettes using the ΦC31 integrase [107, 201]. Other useful avenues to explore could focus on high efficiency integration of the entire dystrophin gene into predefined genomic loci using engineered site-specific recombinases [109, 111, 112, 191, 202] and transposases [203, 204]. These strategies remain difficult to implement due to the low efficiency of integration of large gene constructs by nuclease-mediated homologous recombination and difficulty in engineering highly specific recombinases and

transposases to ensure high fidelity integration. Thus, the continued development of new, highly specific genome editing technologies, in combination with novel gene transfer vectors specific to muscle tissue, will be a valuable component towards creating a single, universal, highly functional correction of all types of dystrophin genetic defects that cause DMD.

Minidystrophin proteins have been designed to retain maximum functionality while reducing the overall size of the coding region to enable efficient gene transfer using adeno-associated virus and thus are of interest to *in vivo* gene correction strategies. Our preliminary work demonstrates the feasibility of two genome editing methods to introduce minidystrophin expression by promoterless knock-in of a gene cassette to: (1) the safe harbor Rosa26 locus in mice or (2) the 5'UTR of the skeletal muscle isoform Dp427m of human dystrophin.

The Rosa26 locus is an attractive site to knock-in a minidystrophin gene cassette because it expresses a long non-coding RNA with no known essential functions and permits ubiquitous expression of a knocked-in gene without needing an exogenous promoter to drive expression. Previously, we engineered a highly active zinc-finger nuclease pair targeting the Rosa26 locus [153]. AAV-based gene transfer of a bicistronic gene cassette expressing this ZFN pair (AAV-ROSA) was able to efficiently modify proliferating and differentiated skeletal myoblasts (**Figures 30a, b**). Local injection of AAV-ROSA into the tibialis anterior of mice resulted in efficient gene modification

(**Figure 30c**), demonstrating that NHEJ based gene repair is possible in skeletal muscle. As mentioned in 6.4, further studies are needed to characterize tissue-specific gene modification using systemic delivery methods. As discussed above, some critical parameters to be assessed will be the dose and AAV capsid used to deliver nuclease and donor template and expression levels of minidystrophin from the Rosa26 locus. Ultimately, future studies will aim to determine if AAV-delivery of the Rosa26 ZFN and a minidystrophin gene cassette can restore functional dystrophin expression *in vivo* and improve the DMD phenotype in the canonical *mdx* animal model.

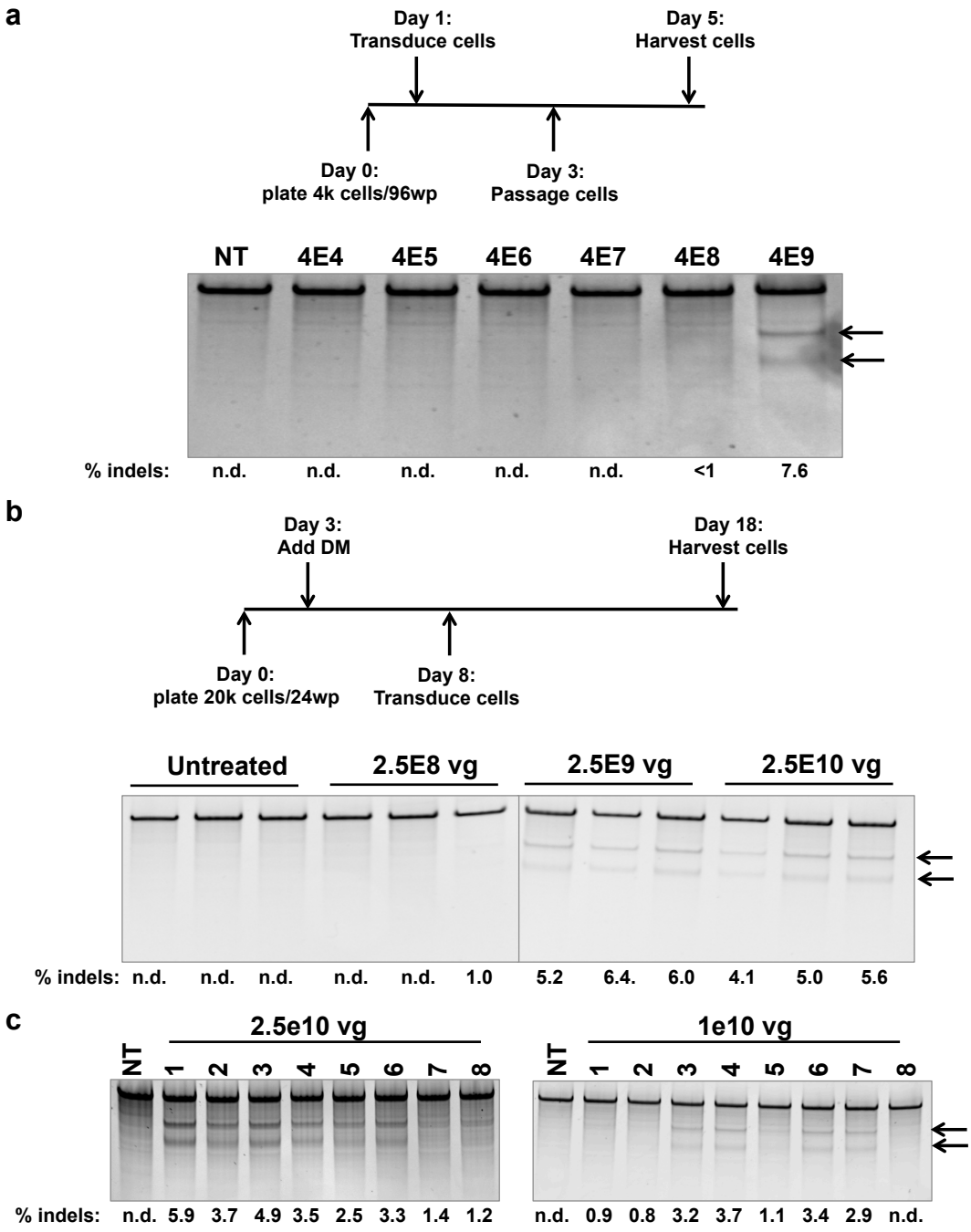


Figure 30: Surveyor analysis of Rosa26 ZFN activities in skeletal *muscle in vitro* and *in vivo* following delivery of AAV-SASTG-ROSA. Arrows indicate expected bands resulting from Surveyor cleavage. n.d.: not detected. (a) Proliferating C2C12s were transduced with the indicated amount of virus and harvested at 4 days

post-infection. Arrows indicate expected bands sizes resulting from Surveyor cleavage. (b) C2C12s were incubated in differentiation medium for 5 days and then transduced with the indicated amount of AAV-SASTG-ROSA virus in 24 well plates. Samples were collected at 10 days post-transduction. (c) The indicated amount of AAV-SASTG-ROSA was injected directly into the tibialis anterior of C57BL/6J mice and muscles were harvested 4 weeks post-infection. The harvested TA muscles were partitioned into 8 separate pieces for genomic DNA analysis, each shown in a separate lane.

Using TALENs, we were able to isolate monoclonal cell lines expressing a minidystrophin gene cassette from the native dystrophin promoter (**Figure 31**). However, in order to expand this as a viable therapeutic platform, several factors must be improved. First, the overall efficiency of integration is presently high but requires antibiotic selection. Investigation of AAV-based delivery *in vitro* and *in vivo* will be important to demonstrate efficient *in situ* knock-in of minidystrophin without selection. Second, we observed highly variable expression levels of minidystrophin in different clones (**Figure 31c**). Further studies may investigate including promoter enhancement elements and/or mRNA stabilization elements in the knocked-in cassette to address these issues.

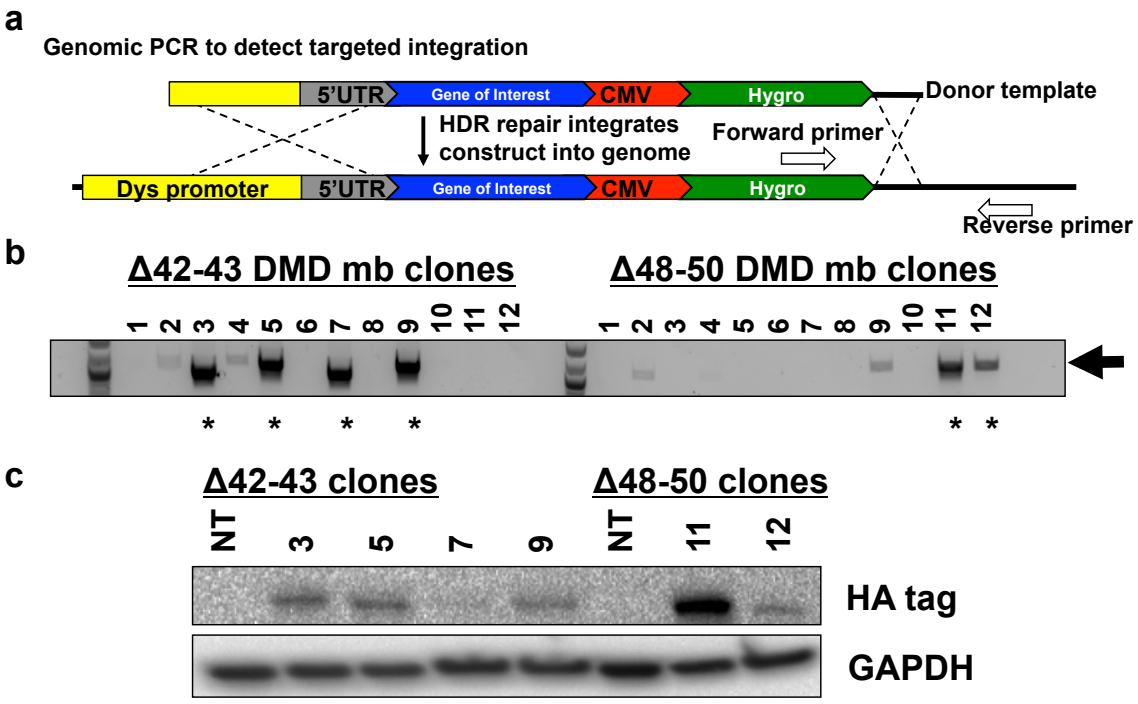


Figure 31: TALEN mediated integration of minidystrophin at the 5'UTR of the Dp427m skeletal muscle isoform of dystrophin in skeletal myoblast cell lines derived from human DMD patients carrying different deletions in the dystrophin gene. DMD patient cells were electroporated with constructs encoding a TALEN pair active at the 5'UTR locus and a donor template carrying the minidystrophin gene. (a) Schematic showing how minidystrophin is integrated into the 5'UTR. (b) Hygromycin-resistant clonal cell lines were isolated and screened by PCR for successful site-specific integrations at the 5'UTR using the primers shown in (a). Asterisks indicate clones selected for further analysis in (c). (c) Clonally isolated DMD myoblasts with detected integration events were differentiated for 6 days and assessed for expression of an HA tag fused to the C terminus of minidystrophin.

Ultimately, these studies will need to be expanded to humanized animal models to determine the efficacy of these approaches in more human-like genomic contexts. The AAVS1 locus is a non-coding safe harbor locus in the human genome similar to Rosa26, for which animal models [205] as well as well-characterized ZFN pairs [169] exist.

Notably, the strategy outlined above does not utilize homology arms to integrate minidystrophin to the Rosa26 locus, therefore only the ZFN utilized would need to be changed to adapt this study. As discussed in Chapter 6.3, a humanized DMD animal model [198] will also be useful in studies investigating gene targeting of minidystrophin to the 5'UTR in an animal model. Thus, there are exciting and available opportunities to expand on the studies presented in this thesis to introduce functional dystrophin expression to address nearly any DMD patient deletion.

6.6. Immunity to restored dystrophin protein products

Rescuing native dystrophin expression will generate novel protein epitopes at the site of gene correction, as well as expression of the absent C-terminus of the dystrophin protein. Therefore it will be important to consider potential immune responses to restored dystrophin proteins [160, 206] following permanent genetic correction of the reading frame. For example, one clinical study found pre-existing dystrophin immunity in approximately 29% of enrolled patients [206]. Interestingly, this study also showed that commonly used steroid treatments for DMD may attenuate potential anti-dystrophin responses. Additionally, exon skipping clinical studies observed a minimal immune response to the restored native gene product in DMD patients with restored dystrophin protein expression [70, 71]. The methods described in Chapters 3-5 result in modification of the native dystrophin gene that may further reduce potential immune responses as compared to the addition of foreign components

contained in exogenous expression cassettes, such as in minidystrophin gene transfer. Finally, it may be possible to introduce tolerance to restored dystrophin protein products by immunomodulation [207-209], and may be enhanced by transient T-cell suppression therapy [210]. Thus, future studies will need to account for dystrophin immunity in assessing the efficiency of gene correction and phenotypic improvement in DMD patients.

Appendix A

Table 5: Exome capture statistics. DOWT is the parent DMD myoblast cell line used as the reference sample for analysis. DO32, DO106, DO127, and DO141 are the four clonally derived DMD myoblast lines carrying predetermined on-target NHEJ events at the exon 51 dystrophin locus.

Sample Name	D0106	DO127	DO141	DO32	DOWT	Agilent-Human All Exon V4
capture efficiency						
Reads onTarget	79.33	79.28	79.27	75.95	79.35	75
Reads On-Target+/-100bp	86.84	88.96	89.15	86.11	89.15	85
Coverage						
1x	99.87	99.88	99.88	99.87	99.88	99
10x	97.4	97.71	97.48	97.53	97.46	90
20x	91.26	92.28	91.41	91.62	91.41	80
30x	82.66	84.45	82.8	83.27	82.94	
50x	63.51	66.35	63.54	64.35	63.9	
100x	28.13	31.12	27.85	28.45	28.62	

Table 6: Target sequences and RVDs for TALENs in this study. All target sequences are preceded by a prerequisite 5′-T.

	Target sequence (5'-3')	**RVDs**
TN1	attttagctcctact	NI NG NG NG NG NI NN HD NG HD HD NG NI HD NG
TN2	tttagctcctactcaga	NG NG NG NI NN HD NG HD HD NG NI HD NG HD NI NN NI
TN3	agctcctactcagact	NI NN HD NG HD HD NG NI HD NG HD NI NN NI HD NG
TN4	cctactcagactgtt	HD HD NG NI HD NG HD NI NN NI HD NG NN NG NG
TN5	aaccacaggttgtgtca	NI NI HD HD NI HD NI NN NN NG NG NN NG NN NG HD NI
TN6	agtaaccacaggttgt	NI NN NG NI NI HD HD NI HD NI NN NN NG NG NN NG
TN8	ccttagtaaccacaggt	HD HD NG NG NI NN NG NI NI HD HD NI HD NI NN NN NG

Appendix B

Table 7: Summary of target sites for ZFNs targeted in Chapter 5.

ZFN target	Target site	Spacer length (bp)
DZF-1	5'-CAA ACT AGA AAT GCC ATC TTCCTT GAT GTT GGA GGT ACC TGC 3'-GTT TGA TCT TTA CGG TAG AAGGAA CTA CAA CCT CCA TGG ACG	6
DZF-2	5'-ATG ATC ATC AAG CAG AAG GTATGA GAA AAA ATG ATA AAA GTT 3'-TAC TAG TAG TTC GTC TTC CATACT CTT TTT TAC TAT TTT CAA	6
DZF-3	5'-GAC TGT TAC TCT GGT GAC ACAACCT GTG GTT ACT AAG GAA ACT 3'-CTG ACA ATG AGA CCA CTG TGTTGGA CAC CAA TGA TTC CTT TGA	7
DZF-4	5'-CTT TAC CAC TTC CAC AAT GTATATG ATT GTT ACT GAG AAG GCT 3'-GAA ATG GTG AAG GTG TTA CATATAC TAA CAA TGA CTC TTC CGA	7
DZF-5	5'-CAC ATT CAC ATT CAC AATATA GTT ATG GAT ATG GAT GTA 3'-GAG TAA GTG TAA GTG TTATAT CAA TAC CTA TAC CTA CAT	6
DZF-6	5'-AAC TTC ACC AAT TCCATA GGA ATA AAA GTA ATT TGA 3'-TTG AAG TGG TTA AGGTAT CCT TAT TTT CAT TAA ACT	6
DZF-7	5'-AAC CCC ATC AAAAAGT GGG GGA AGG 3'-TTG GGG TAG TTTTTCA CCC CCT TCC	7
DZF-8	5'-ATC ATC TCC TCTGGTG GAT GAG GCT 3'-TAG TAG AGG AGACCAC CTA CTC CGA	7
DZF-9	5'-ATC TGC CCA TGACT GGC GCA GGG 3'-TAG ACG GGT ACTGA CCG CGT CCC	5
DZF-10	5'-GCC ATC TTC CTTGAT GTT GGA GGT 3'-CGG TAG AAG GAACTA CAA CCT CCA	6
DZF-11	5'-TGC TTC AGC CTCCTGA GTA GCT GGG 3'-ACG AAG TCG GAGGACT CAT CGA CCC	7
DZF-12	5'-GCC TCA GCC TCCCAAA GTG GTG GGA 3'-CGG AGT CGG AGGGTTT CAC CAC CCT	7
DZF-13	5'-CTC AGC CTC CCAAA GTG GTG GGA 3'-GAG TCG GAG GGTTT CAC CAC CCT	5

Table 8: Sequences of zinc-finger recognition helices to supplement Barbas modular assembly kit.

Triplet target	Helix sequence	Source
TGC	QRNALAG	ZiFiT F3 ID #631
TCT	QQRSLVG	ZiFiT F3 ID #790

Table 9: Primers used in the study in Chapter 4.

Primer name	Primer sequence	Notes
CelI-DZF-1/3/10-F	GAGTTTGGCTCAAATTGTTACTCTT	Forward Surveyor primer for DZF-1, DZF-3 and DZF-10
CelI-DZF-1/3/10-R	GGGAAATGGTCTAGGAGAGTAAAGT	Reverse Surveyor primer for DZF-1, DZF-3 and DZF-10
CelI-DZF-4/5/6-F	CCTCAGTGTAATCCATTTGGTAAAA	Forward Surveyor primer forDZF-4, DZF-5 and DZF-6
CelI-DZF-4-R	CTGCTACTTACTGGGAATTTGACAT	Reverse Surveyor primer for DZF-4
CelI-DZF-5/6-R	CAAAGTTGTGCTGAAGGTATTTAGG	Reverse Surveyor primer for DZF-5 and DZF-6
CelI-DZF-7-F	AACCATTGGAATTTACAGGATGAT	Forward Surveyor primer for DZF-7
CelI-DZF-7-R	GGCTGAGTTAAATGGTATTTCTGG	Reverse Surveyor primer for DZF-7
CelI-DZF-8-F	ACTTGCACCTCATTCTAATTGTGA	Forward Surveyor primer for DZF-8
CelI-DZF-8-R	CCTCCTACCTGAATGTTAGAGACAA	Reverse Surveyor primer for DZF-8
CelI-DZF-9-F	GATGCAAGAGATAGAGCAGTGAGA	Forward Surveyor primer for DZF-9
CelI-DZF-9-R	GTTTGGAAAAAGACAGAAAGGAAG	Reverse Surveyor primer for DZF-9
CelI-DZF-11-F	CCAATGACTTAAGGTTTCTTCACA	Forward Surveyor primer for DZF-11
CelI-DZF-11-R	CTGAATCATTGATGAAAAAGACCA	Reverse Surveyor primer for DZF-11
CelI-DZF-12/13-F	CCAACATGAGACTTTCTTTTTGTTT	Forward Surveyor primer for DZF-12 and DZF-13
CelI-DZF-12/13-R	AGCTGGAATATGCTTTTACTTTCC	Reverse Surveyor primer for DZF-12 and DZF-13
Dys-E44-F	TGGCGGCGTTTTCATTAT	Forward RT-PCR primer binding in exon 44
Dys-E52-R	TTCGATCCGTAATGATTGTTCTAGCC	Reverse RT-PCR primer binding in exon 52
SSA-fwd	CTAGCAAAATAGGCTGTCCC	Forward primer to construct SSA luciferase
SSA-luc-DZF-1-rev	GAGGAGGAATTCAGCAGGTACCTCCAACATCA AGGAAGATGGCATTTCTAGTTTGGTCACATAGG ACCTCTCACACACAG	Reverse primer to construct SSA luciferase
SSA-luc-DZF-2-rev	GAGGAGGAATTCCAACTTTTATCATTTTTTCTCA TACCTTCTGCTTGATGATCATCTCACATAGGACC TCTCACACACAG	Reverse primer to construct SSA luciferase
SSA-luc-DZF-3-rev	GAGGAGGAATTCCAGTTTCCTTAGTAACCACAG GTTGTGTCACCAGAGTAACAGTCTTCACATAGG ACCTCTCACACACAG	Reverse primer to construct SSA luciferase
SSA-luc-DZF-4-rev	GAGGAGGAATTCAAGCCTTCTCAGTAACAATC ATATACATTGTGGAAGTGGTAAAGATCACATAG GACCTCTCACACACAG	Reverse primer to construct SSA luciferase
SSA-luc-DZF-5-rev	GAGGAGGAATTCATACATCCATATCCATAACTA TATTGTGAATGTGAATGTGTTCACATAGGACCT CTCACACACAG	Reverse primer to construct SSA luciferase
SSA-luc-DZF-6-rev	GAGGAGGAATTCATCAAATTACTTTTATTCCTA TGGAATTGGTGAAGTTTTCACATAGGACCTCTC ACACACAG	Reverse primer to construct SSA luciferase
DZF1-OT-1-F	CATGCTAGCTCCTACAAAGCACTG	Forward Surveyor primer for DZF-1 off-target site 1
DZF1-OT-1-R	GGGAAATGGGTACTGAAGAAGACG	Reverse Surveyor primer for DZF-1 off-target site 1
DZF1-OT-2-F	CTGTGCTGCCTATTGCTTTCTGTC	Forward Surveyor primer for DZF-1

		off-target site 2
DZF1-OT-2-R	CTGGTTGTGTGCCTAGTGATGG	Reverse Surveyor primer for DZF-1 off-target site 2
DZF1-OT-3-F	CCCATTACTGCATTTGCGGTCTTG	Forward Surveyor primer for DZF-1 off-target site 3
DZF1-OT-3-R	TCAACCTTGCCTGCACGGAG	Reverse Surveyor primer for DZF-1 off-target site 3
DZF1-OT-4-F	CCTCTTCTCTTGGGATCTGTGAGT	Forward Surveyor primer for DZF-1 off-target site 4
DZF1-OT-4-R	GAGAACCCAATGTAATGTGTTCACTGAGC	Reverse Surveyor primer for DZF-1 off-target site 4
DZF1-OT-5-F	AAAGACACCTTTTCTGCCCTCACG	Forward Surveyor primer for DZF-1 off-target site 5
DZF1-OT-5-R	GTGCCCAGCCCAATTCTTTCTTGC	Reverse Surveyor primer for DZF-1 off-target site 5
DZF1-OT-6-F	GGTCCTGGTCCAAAGCAATTCTG	Forward Surveyor primer for DZF-1 off-target site 6
DZF1-OT-6-R	CGCCCGGCCAGATTTGTCTA	Reverse Surveyor primer for DZF-1 off-target site 6
DZF1-OT-7-F	CCACACACACAGGACACTGATC	Forward Surveyor primer for DZF-1 off-target site 7
DZF1-OT-7-R	CCAGAAGGCAGCCACTAGAAAC	Reverse Surveyor primer for DZF-1 off-target site 7
DZF1-OT-8-F	CTAGAATTACAGGCGTGAGCCACT	Forward Surveyor primer for DZF-1 off-target site 8
DZF1-OT-8-R	GCAGCTGAGTTGCAGGCATAAG	Reverse Surveyor primer for DZF-1 off-target site 8
DZF9-OT-1-F	ACGTTCCTGGGAAACACAGGG	Forward Surveyor primer for DZF-9 off-target site 1
DZF9-OT-1-R	CCACCAAAAGGCAGCTCCATAAAC	Reverse Surveyor primer for DZF-9 off-target site 1
DZF9-OT-2-F	GCACAGGGTACACACCCATTAAC	Forward Surveyor primer for DZF-9 off-target site 2
DZF9-OT-2-R	AGTCTCTCCATCCCCGAGGT	Reverse Surveyor primer for DZF-9 off-target site 2
DZF9-OT-3-F	CTGGTTTCTGCACCACATATTGCC	Forward Surveyor primer for DZF-9 off-target site 3
DZF9-OT-3-R	CACATGGCCGGCAGGAGAAA	Reverse Surveyor primer for DZF-9 off-target site 3
DZF9-OT-4-F	AAAAGGGAGCAGGTCAGCACAC	Forward Surveyor primer for DZF-9 off-target site 4
DZF9-OT-4-R	GGGGGAATTTGGGGAAACTTTCCT	Reverse Surveyor primer for DZF-9 off-target site 4
DZF9-OT-5-F	TGAGTCAGATGGCCCAGGGA	Forward Surveyor primer for DZF-9 off-target site 5
DZF9-OT-5-R	CTTGGAGCCTCCACAGGTGA	Reverse Surveyor primer for DZF-9 off-target site 5
DZF9-OT-6-F	TGGACTGAGGGAACCCCTCT	Forward Surveyor primer for DZF-9 off-target site 6
DZF9-OT-6-R	CAGATTCCCAGGGAAGCTCG	Reverse Surveyor primer for DZF-9 off-target site 6
DZF9-OT-7-F	GCGCGCTCGGGTGAAAAAATTAAG	Forward Surveyor primer for DZF-9 off-target site 7
DZF9-OT-7-R	TCCCCTTCTTCCACCTCCAG	Reverse Surveyor primer for DZF-9

		off-target site 7
DZF9-OT-8-F	GGCTCCCCTCCTTGTTAATGTTG	Forward Surveyor primer for DZF-9 off-target site 8
DZF9-OT-8-R	AGTGAGGACGATGACCAGCG	Reverse Surveyor primer for DZF-9 off-target site 8

Appendix C

Table 10: List of sgRNA targets in Chapter 5. PAM: protospacer-adjacent motif.

Name	Target	Strand	19bp protospacer	PAM	Target Finder
CR1	Intron 50	+	GATTGGCTTTGATTTCCCTA	GGG	Manual Inspection
CR2	Intron 50	-	GTGTAGAGTAAGTCAGCCTA	TGG	Manual Inspection
CR3	Exon 51-5'	+	GCCTACTCAGACTGTTACTC	TGG	Manual Inspection
CR4	Exon 51-3'	+	GTTGGACAGAACTTACCGAC	TGG	Manual Inspection
CR5	Intron 51	-	GCAGTTGCCTAAGAACTGGT	GGG	Manual Inspection
CR6	Intron 44	-	GGGGCTCCACCCTCACGAGT	GGG	UCSC Browser (Cong et al. Science 2013)
CR7	Intron 55	+	GTTTGCTTCGCTATAAAACG	AGG	UCSC Browser (Cong et al. Science 2013)
CR10	Exon 45	+	GCCAGGATGGCATTGGGCAG	CGG	Manual Inspection
CR11	Exon 45	+	GCTGAATCTGCGGTGGCAGG	AGG	Manual Inspection
CR12	Exon 46	-	GTTCTTTTGTTCTTCTAGCC	TGG	Manual Inspection
CR13	Exon 46	+	GGAAAAGCTTGAGCAAGTCA	AGG	Manual Inspection
CR14	Exon 47	+	GGAAGAGTTGCCCCTGCGCC	AGG	Excel File from Mali et al. Science 2013
CR15	Exon 47	+	GACAAATCTCCAGTGGATAA	AGG	Manual Inspection
CR16	Exon 48	-	GTGTTTCTCAGGTAAAGCTC	TGG	Manual Inspection
CR17	Exon 48	+	GGAAGGACCATTTGACGTTA	AGG	Manual Inspection
CR18	Exon 49	-	GAACTGCTATTTCAGTTTCC	TGG	Manual Inspection
CR19	Exon 49	+	GCCAGCCACTCAGCCAGTGA	AGG	Manual Inspection
CR20	Exon 50	+	GGTATGCTTTTCTGTTAAAG	AGG	Manual Inspection
CR21	Exon 50	+	GCTCCTGGACTGACCACTAT	TGG	Manual Inspection
CR22	Exon 52	+	GGAACAGAGGCGTCCCCAGT	TGG	Manual Inspection
CR23	Exon 52	+	GGAGGCTAGAACAATCATTA	CGG	Manual Inspection
CR24	Exon 53	+	GACAAGAACACCTTCAGAAC	CGG	Manual Inspection
CR25	Exon 53	-	GGGTTTCTGTGATTTTCTTT	TGG	Manual Inspection
CR26	Exon 54	+	GGGCCAAAGACCTCCGCCAG	TGG	Manual Inspection
CR27	Exon 54	+	GTTGGAGAAGCATTCATAAA	AGG	Manual Inspection
CR28	Exon 55	-	GTCGCTCACTCACCCTGCAA	AGG	Manual Inspection
CR29	Exon 55	+	GAAAAGAGCTGATGAAACAA	TGG	Manual Inspection
CR31	Exon 51	+	GGAGATGATCATCAAGCAGA	AGG	Manual Inspection
CR33	Intron 44	-	GCACAAAAGTCAAATCGGAA	TGG	UCSC Browser (Cong et al. Science 2013)
CR34	Intron 44	-	GATTTCAATATAAGATTCGG	AGG	UCSC Browser (Cong et al. Science 2013)
CR35	Intron 55	-	GCTTAAGCAATCCCGAACTC	TGG	UCSC Browser (Cong et al. Science 2013)
CR36	Intron 55	-	GCCTTCTTTATCCCCTATCG	AGG	UCSC Browser (Cong et al. Science 2013)

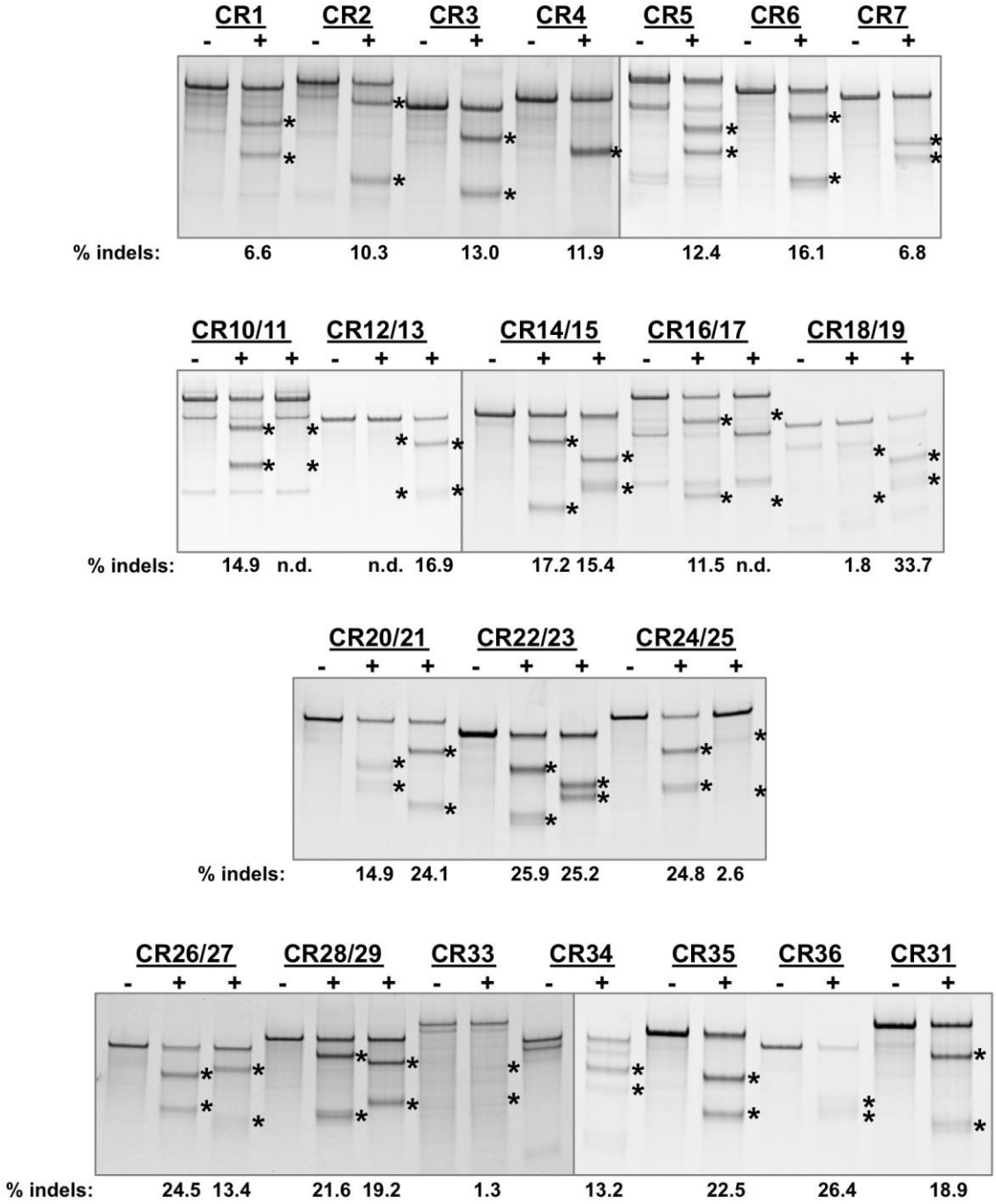


Figure 32: Images of TBE-PAGE gels used to quantify Surveyor assay results to measure day 3 gene modification in Table 3. Asterisks mark expected sizes of bands indicative of nuclease activity.

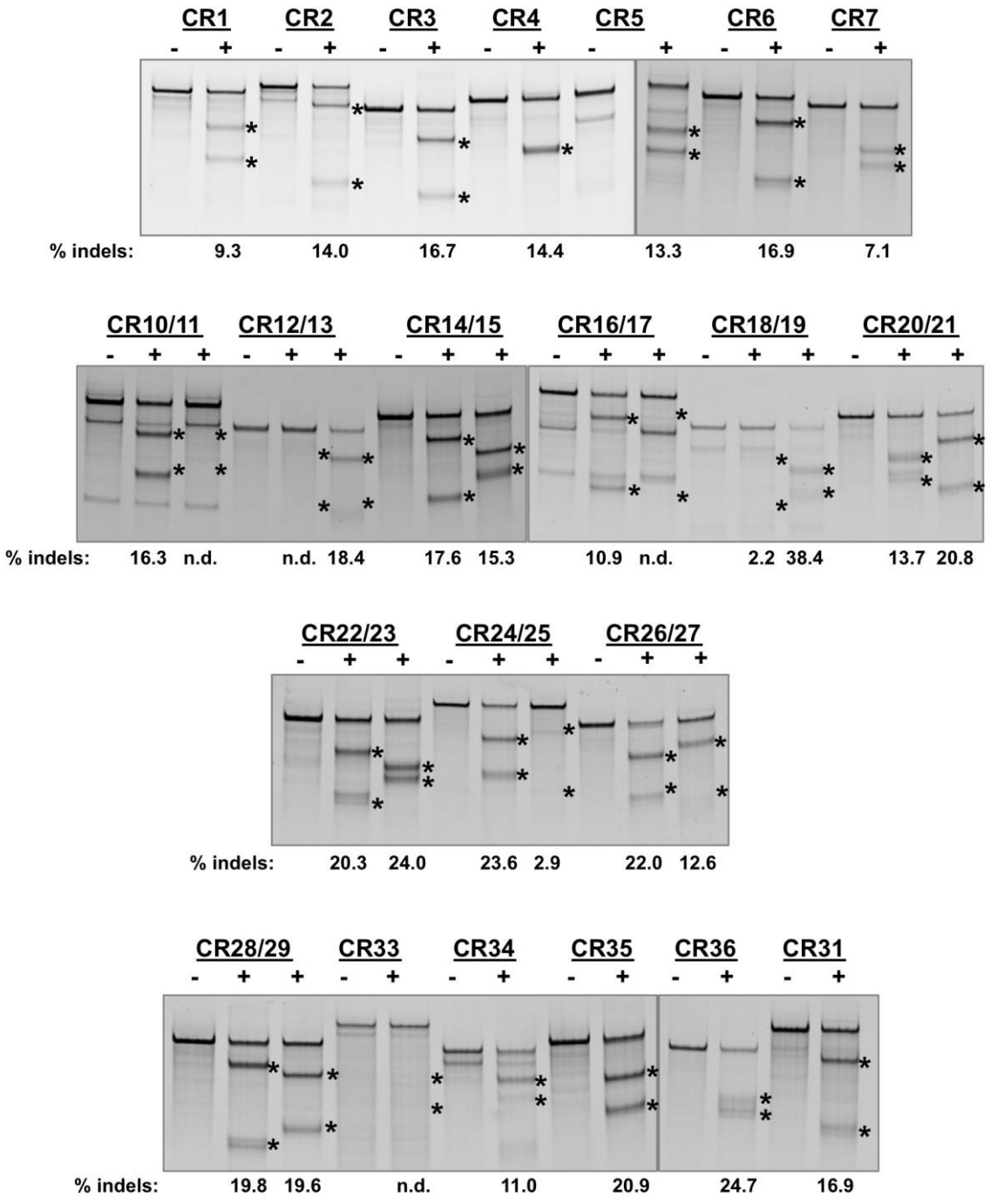


Figure 33: Images of TBE-PAGE gels used to quantify Surveyor assay results to measure day 10 gene modification in Table 3. Asterisks mark expected sizes of bands indicative of nuclease activity.

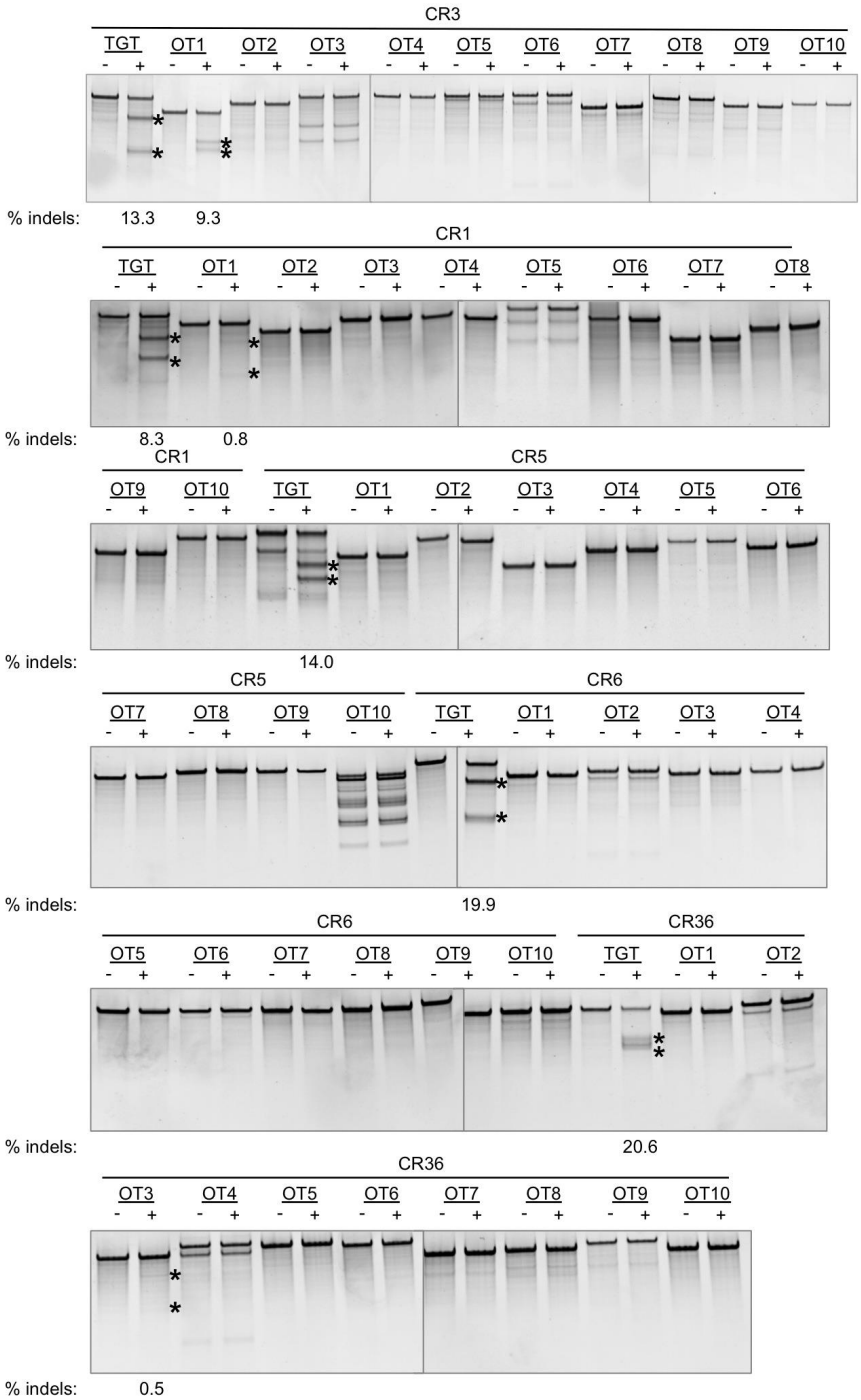


Figure 34: Images of TBE-PAGE gels used to quantify Surveyor assay results to measure on-target and off-target gene modification in Table 4. Asterisks mark expected sizes of bands indicative of nuclease activity.

References

1. Monaco, A.P., Bertelson, C.J., Colletti-Feener, C. & Kunkel, L.M. Localization and cloning of Xp21 deletion breakpoints involved in muscular dystrophy. *Hum Genet* **75**, 221-227 (1987).

2. Hoffman, E.P., Brown, R.H., Jr. & Kunkel, L.M. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**, 919-928 (1987).

3. Spurney, C.F. Cardiomyopathy of Duchenne muscular dystrophy: current understanding and future directions. *Muscle Nerve* **44**, 8-19 (2011).

4. Lu, Q.L. et al. The status of exon skipping as a therapeutic approach to duchenne muscular dystrophy. *Mol Ther* **19**, 9-15 (2011).

5. Pichavant, C. et al. Current status of pharmaceutical and genetic therapeutic approaches to treat DMD. *Mol Ther* **19**, 830-840 (2011).

6. Perez-Pinera, P., Ousterout, D.G. & Gersbach, C.A. Advances in targeted genome editing. *Curr Opin Chem Biol* **16**, 268-277 (2012).

7. Silva, G. et al. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther* **11**, 11-27 (2011).

8. Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**, e82 (2011).

9. Reyon, D. et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol* **30**, 460-465 (2012).

10. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013).

11. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013).

12. Perez, E.E. et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**, 808-816 (2008).

13. Tebas, P. et al. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* **370**, 901-910 (2014).

14. Aartsma-Rus, A. & van Ommen, G.J. Antisense-mediated exon skipping: a versatile tool with therapeutic and research applications. *RNA* **13**, 1609-1624 (2007).

15. Taniguchi-Ikeda, M. et al. Pathogenic exon-trapping by SVA retrotransposon and rescue in Fukuyama muscular dystrophy. *Nature* **478**, 127-131 (2011).

16. Blake, D.J., Weir, A., Newey, S.E. & Davies, K.E. Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiol Rev* **82**, 291-329 (2002).

17. Ervasti, J.M. Dystrophin, its interactions with other proteins, and implications for muscular dystrophy. *Biochim Biophys Acta* **1772**, 108-117 (2007).

18. Rando, T.A. The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies. *Muscle Nerve* **24**, 1575-1594 (2001).

19. Konieczny, P., Swiderski, K. & Chamberlain, J.S. Gene and cell-mediated therapies for muscular dystrophy. *Muscle Nerve* **47**, 649-663 (2013).

20. Barnea, E., Zuk, D., Simantov, R., Nudel, U. & Yaffe, D. Specificity of expression of the muscle and brain dystrophin gene promoters in muscle and brain cells. *Neuron* **5**, 881-888 (1990).

21. van den Bergen, J.C. et al. Dystrophin levels and clinical severity in Becker muscular dystrophy patients. *J Neurol Neurosurg Psychiatry* (2013).

22. van den Bergen, J.C. et al. Clinical characterisation of Becker muscular dystrophy patients predicts favourable outcome in exon-skipping therapy. *J Neurol Neurosurg Psychiatry* **85**, 92-98 (2014).

23. Kinali, M. et al. Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *Lancet Neurol* **8**, 918-928 (2009).

24. van Deutekom, J.C. et al. Local dystrophin restoration with antisense oligonucleotide PRO051. *N Engl J Med* **357**, 2677-2686 (2007).

25. Palmieri, B. & Tremblay, J.P. Myoblast transplantation: a possible surgical treatment for a severe pediatric disease. *Surg Today* **40**, 902-908 (2010).

26. Roy, R. et al. Antibody formation after myoblast transplantation in Duchenne-dystrophic patients, donor HLA compatible. *Transplant Proc* **25**, 995-997 (1993).

27. Miller, R.G. et al. Myoblast implantation in Duchenne muscular dystrophy: the San Francisco study. *Muscle Nerve* **20**, 469-478 (1997).

28. Skuk, D. et al. Dystrophin expression in muscles of duchenne muscular dystrophy patients after high-density injections of normal myogenic cells. *J Neuropathol Exp Neurol* **65**, 371-386 (2006).

29. Tedesco, F.S. et al. Transplantation of genetically corrected human iPSC-derived progenitors in mice with limb-girdle muscular dystrophy. *Sci Transl Med* **4**, 140ra189 (2012).

30. Tedesco, F.S. et al. Stem cell-mediated transfer of a human artificial chromosome ameliorates muscular dystrophy. *Sci Transl Med* **3**, 96ra78 (2011).

31. Benabdallah, B.F. et al. Targeted Gene Addition of Microdystrophin in Mice Skeletal Muscle via Human Myoblast Transplantation. *Mol Ther Nucleic Acids* **2**, e68 (2013).

32. Darabi, R. et al. Human ES- and iPS-Derived Myogenic Progenitors Restore DYSTROPHIN and Improve Contractility upon Transplantation in Dystrophic Mice. *Cell Stem Cell* **10**, 610-619 (2012).

33. Dezawa, M. et al. Bone marrow stromal cells generate muscle cells and repair muscle degeneration. *Science* **309**, 314-317 (2005).

34. Cerletti, M. et al. Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* **134**, 37-47 (2008).

35. Negroni, E. et al. In vivo myogenic potential of human CD133+ muscle-derived stem cells: a quantitative study. *Mol Ther* **17**, 1771-1778 (2009).

36. Peault, B. et al. Stem and progenitor cells in skeletal muscle development, maintenance, and therapy. *Mol Ther* **15**, 867-877 (2007).

37. Zhu, C.H. et al. Cellular senescence in human myoblasts is overcome by human telomerase reverse transcriptase and cyclin-dependent kinase 4: consequences in aging muscle and therapeutic strategies for muscular dystrophies. *Aging Cell* **6**, 515-523 (2007).

38. Berghella, L. et al. Reversible immortalization of human myogenic cells by site-specific excision of a retrovirally transferred oncogene. *Hum Gene Ther* **10**, 1607-1617 (1999).

39. Lombardo, A. et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* **25**, 1298-1306 (2007).

40. Gaj, T., Guo, J., Kato, Y., Sirk, S.J. & Barbas, C.F., 3rd Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. *Nat Methods* **9**, 805-807 (2012).

41. Liu, J., Gaj, T., Patterson, J.T., Sirk, S.J. & Barbas, C.F., 3rd Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. *PLoS One* **9**, e85755 (2014).

42. Ousterout, D.G. et al. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther* **21**, 1718-1726 (2013).

43. Filareto, A. et al. An ex vivo gene therapy approach to treat muscular dystrophy using inducible pluripotent stem cells. *Nat Commun* **4**, 1549 (2013).

44. Kafri, T., Blomer, U., Peterson, D.A., Gage, F.H. & Verma, I.M. Sustained expression of genes delivered directly into liver and muscle by lentiviral vectors. *Nat Genet* **17**, 314-317 (1997).

45. Naldini, L. et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263-267 (1996).

46. Asokan, A. et al. Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nat Biotechnol* **28**, 79-82 (2010).

47. Messina, E.L. et al. Adeno-Associated Viral Vectors Based on Serotype 3b Use Components of the Fibroblast Growth Factor Receptor Signaling Complex for Efficient Transduction. *Hum Gene Ther* **23**, 1031-1042 (2012).

48. Bowles, D.E. et al. Phase 1 Gene Therapy for Duchenne Muscular Dystrophy Using a Translational Optimized AAV Vector. *Mol Ther* **20**, 443-455 (2012).

49. Qiao, C., Koo, T., Li, J., Xiao, X. & Dickson, J.G. Gene therapy in skeletal muscle mediated by adeno-associated virus vectors. *Methods Mol Biol* **807**, 119-140 (2011).

50. Qiao, C. et al. Hydrodynamic limb vein injection of adeno-associated virus serotype 8 vector carrying canine myostatin propeptide gene into normal dogs enhances muscle growth. *Hum Gene Ther* **20**, 1-10 (2009).

51. Wang, Z. et al. Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. *Nat Biotechnol* **23**, 321-328 (2005).

52. Watchko, J. et al. Adeno-associated virus vector-mediated minidystrophin gene therapy improves dystrophic muscle contractile function in mdx mice. *Hum Gene Ther* **13**, 1451-1460 (2002).

53. Yang, L. et al. A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection. *Proc Natl Acad Sci U S A* **106**, 3946-3951 (2009).

54. Liu, M. et al. Adeno-associated virus-mediated microdystrophin expression protects young mdx muscle from contraction-induced injury. *Mol Ther* **11**, 245-256 (2005).

55. Odom, G.L., Gregorevic, P., Allen, J.M., Finn, E. & Chamberlain, J.S. Microutrophin delivery through rAAV6 increases lifespan and improves muscle function in dystrophic dystrophin/utrophin-deficient mice. *Mol Ther* **16**, 1539-1545 (2008).

56. Wang, Z. et al. Sustained AAV-mediated dystrophin expression in a canine model of Duchenne muscular dystrophy with a brief course of immunosuppression. *Mol Ther* **15**, 1160-1166 (2007).

57. Ragot, T. et al. Efficient adenovirus-mediated transfer of a human minidystrophin gene to skeletal muscle of mdx mice. *Nature* **361**, 647-650 (1993).

58. T Seto, J., N Ramos, J., Muir, L., S Chamberlain, J. & L Odom, G. Gene Replacement Therapies for Duchenne Muscular Dystrophy Using Adeno-Associated Viral Vectors. *Current Gene Therapy* **12**, 139-151 (2012).

59. Gruber, K. Europe gives gene therapy the green light. *Lancet* **380**, e10 (2012).

60. Wang, B., Li, J. & Xiao, X. Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model. *Proc Natl Acad Sci U S A* **97**, 13714-13719 (2000).

61. Harper, S.Q. et al. Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy. *Nat Med* **8**, 253-261 (2002).

62. Wolff, J.A. et al. Direct gene transfer into mouse muscle in vivo. *Science* **247**, 1465-1468 (1990).

63. Aihara, H. & Miyazaki, J. Gene transfer into muscle by electroporation in vivo. *Nat Biotechnol* **16**, 867-870 (1998).

64. Liu, F., Nishikawa, M., Clemens, P.R. & Huang, L. Transfer of full-length Dmd to the diaphragm muscle of Dmd(mdx/mdx) mice through systemic administration of plasmid DNA. *Mol Ther* **4**, 45-51 (2001).

65. Taniyama, Y. et al. Development of safe and efficient novel nonviral gene transfer using ultrasound: enhancement of transfection efficiency of naked plasmid DNA in skeletal muscle. *Gene Ther* **9**, 372-380 (2002).

66. Lemieux, P. et al. A combination of poloxamers increases gene expression of plasmid DNA in skeletal muscle. *Gene Ther* **7**, 986-991 (2000).

67. Lorden, E.R., Levinson, H.M. & Leong, K.W. Integration of drug, protein, and gene delivery systems with regenerative medicine. *Drug Delivery and Translational Research*, 1-19 (2013).

68. McNeer, N.A. et al. Systemic delivery of triplex-forming PNA and donor DNA by nanoparticles mediates site-specific genome editing of human hematopoietic cells in vivo. *Gene Ther* **20**, 658-669 (2013).

69. Ruszczak, C., Mirza, A. & Menhart, N. Differential stabilities of alternative exon-skipped rod motifs of dystrophin. *Biochim Biophys Acta* **1794**, 921-928 (2009).

70. Cirak, S. et al. Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. *Lancet* **378**, 595-605 (2011).

71. Goemans, N.M. et al. Systemic administration of PRO051 in Duchenne's muscular dystrophy. *N Engl J Med* **364**, 1513-1522 (2011).

72. Bidou, L., Allamand, V., Rousset, J.P. & Namy, O. Sense from nonsense: therapies for premature stop codon diseases. *Trends Mol Med* **18**, 679-688 (2012).

73. Welch, E.M. et al. PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **447**, 87-91 (2007).

74. Malik, V. et al. Gentamicin-induced readthrough of stop codons in Duchenne muscular dystrophy. *Ann Neurol* **67**, 771-780 (2010).

75. Joung, J.K. & Sander, J.D. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol* **14**, 49-55 (2013).

76. Moehle, E.A. et al. 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 (2007).

77. Urnov, F.D. et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646-651 (2005).

78. Sebastiano, V. et al. In Situ Genetic Correction of the Sickle Cell Anemia Mutation in Human Induced Pluripotent Stem Cells Using Engineered Zinc Finger Nucleases. *Stem Cells* **29**, 1717-1726 (2011).

79. Zou, J., Mali, P., Huang, X., Dowey, S.N. & Cheng, L. Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. *Blood* **118**, 4599-4608 (2011).

80. Yusa, K. et al. Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* **478**, 391-394 (2011).

81. Li, H. et al. In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217-221 (2011).

82. Ellis, B.L., Hirsch, M.L., Porter, S.N., Samulski, R.J. & Porteus, M.H. Zinc-finger nuclease-mediated gene correction using single AAV vector transduction and enhancement by Food and Drug Administration-approved drugs. *Gene Ther* **20**, 35-42 (2013).

83. Hirsch, M.L., Green, L., Porteus, M.H. & Samulski, R.J. Self-complementary AAV mediates gene targeting and enhances endonuclease delivery for double-strand break repair. *Gene Ther* **17**, 1175-1180 (2010).

84. Handel, E.M. et al. Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. *Hum Gene Ther* **23**, 321-329 (2012).

85. Rahman, S.H. et al. The nontoxic cell cycle modulator indirubin augments transduction of adeno-associated viral vectors and zinc-finger nuclease-mediated gene targeting. *Hum Gene Ther* **24**, 67-77 (2013).

86. Asuri, P. et al. Directed evolution of adeno-associated virus for enhanced gene delivery and gene targeting in human pluripotent stem cells. *Mol Ther* **20**, 329-338 (2012).

87. Anguela, X.M. et al. Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood* **122**, 3283-3287 (2013).

88. Benedetti, S., Hoshiya, H. & Tedesco, F.S. Repair or replace? Exploiting novel gene and cell therapy strategies for muscular dystrophies. *FEBS J* **280**, 4263-4280 (2013).

89. Chapdelaine, P., Pichavant, C., Rousseau, J., Paques, F. & Tremblay, J.P. Meganucleases can restore the reading frame of a mutated dystrophin. *Gene Ther* **17**, 846-858 (2010).

90. Rousseau, J. et al. Endonucleases: tools to correct the dystrophin gene. *J Gene Med* **13**, 522-537 (2011).

91. Sollu, C. et al. Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion. *Nucleic Acids Res* **38**, 8269-8276 (2010).

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

93. Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. & Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646 (2010).

94. Mussolino, C. & Cathomen, T. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol* **23**, 644-650 (2012).

95. Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**, 731-734 (2011).

96. Mussolino, C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res* **39**, 9283-9293 (2011).

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

98. Boch, J. et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509 (2009).

99. Cho, S.W., Kim, S., Kim, J.M. & Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232 (2013).

100. Hou, Z. et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proc Natl Acad Sci U S A* **110**, 15644-15649 (2013).

101. Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* **2**, e00471 (2013).

102. Horii, T., Tamura, D., Morita, S., Kimura, M. & Hatada, I. Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *Int J Mol Sci* **14**, 19774-19781 (2013).

103. Yang, L. et al. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res* **41**, 9049-9061 (2013).

104. Redondo, P. et al. Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* **456**, 107-111 (2008).

105. Bertoni, C. & Rando, T.A. Dystrophin gene repair in mdx muscle precursor cells in vitro and in vivo mediated by RNA-DNA chimeric oligonucleotides. *Hum Gene Ther* **13**, 707-718 (2002).

106. Kayali, R., Bury, F., Ballard, M. & Bertoni, C. Site-directed gene repair of the dystrophin gene mediated by PNA-ssODNs. *Hum Mol Genet* **19**, 3266-3281 (2010).

107. Quenneville, S.P. et al. Nucleofection of muscle-derived stem cells and myoblasts with phiC31 integrase: stable expression of a full-length-dystrophin fusion gene by human myoblasts. *Mol Ther* **10**, 679-687 (2004).

108. Bertoni, C. et al. Enhancement of plasmid-mediated gene therapy for muscular dystrophy by directed plasmid integration. *Proc Natl Acad Sci U S A* **103**, 419-424 (2006).

109. Gaj, T., Mercer, A.C., Gersbach, C.A., Gordley, R.M. & Barbas, C.F., 3rd Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc Natl Acad Sci U S A* **108**, 498-503 (2011).

110. Gersbach, C.A., Gaj, T., Gordley, R.M. & Barbas, C.F., 3rd Directed evolution of recombinase specificity by split gene reassembly. *Nucleic Acids Res* **38**, 4198-4206 (2010).

111. Gersbach, C.A., Gaj, T., Gordley, R.M., Mercer, A.C. & Barbas, C.F. Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. *Nucleic Acids Res* **39**, 7868-7878 (2011).

112. Mercer, A.C., Gaj, T., Fuller, R.P. & Barbas, C.F., 3rd Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res* **40**, 11163-11172 (2012).

113. Miller, J.C. et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143-148 (2011).

114. Miller, J.C. et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* **25**, 778-785 (2007).

115. Doyon, Y. et al. Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods* **8**, 74-79 (2011).

116. Ramirez, C.L. et al. Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res* **40**, 5560-5568 (2012).

117. Wang, J.B. et al. Targeted gene addition to a predetermined site in the human genome using a ZFN-based nicking enzyme. *Genome Res* **22**, 1316-1326 (2012).

118. Guo, J., Gaj, T. & Barbas, C.F., 3rd Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol* **400**, 96-107 (2010).

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

120. Alwin, S. et al. Custom zinc-finger nucleases for use in human cells. *Mol Ther* **12**, 610-617 (2005).

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

122. Maeder, M.L. et al. Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* **31**, 294-301 (2008).

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

124. Mandell, J.G. & Barbas, C.F., 3rd Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res* **34**, W516-523 (2006).

125. Gonzalez, B. et al. Modular system for the construction of zinc-finger libraries and proteins. *Nat Protoc* **5**, 791-810 (2010).

126. Kim, S., Lee, M.J., Kim, H., Kang, M. & Kim, J.S. Preassembled zinc-finger arrays for rapid construction of ZFNs. *Nat Methods* **8**, 7 (2011).

127. Ramirez, C.L. et al. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods* **5**, 374-375 (2008).

128. Sander, J.D. et al. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* **8**, 67-69 (2011).

129. Doyle, E.L. et al. TAL Effector Specificity for base 0 of the DNA Target Is Altered in a Complex, Effector- and Assay-Dependent Manner by Substitutions for the Tryptophan in Cryptic Repeat -1. *PLoS One* **8**, e82120 (2013).

130. Lamb, B.M., Mercer, A.C. & Barbas, C.F., 3rd Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucleic Acids Res* **41**, 9779-9785 (2013).

131. Bedell, V.M. et al. In vivo genome editing using a high-efficiency TALEN system. *Nature* **491**, 114-118 (2012).

132. Hockemeyer, D. et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nature biotechnology* **29**, 731-734 (2011).

133. Briggs, A.W. et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res* **40**, e117 (2012).

134. Schmid-Burgk, J.L., Schmidt, T., Kaiser, V., Honing, K. & Hornung, V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol* **31**, 76-81 (2013).

135. Wang, H. et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910-918 (2013).

136. Esvelt, K.M. et al. Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat Methods* **10**, 1116-1121 (2013).

137. Hsu, P.D. et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827-832 (2013).

138. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826 (2013).

139. Mali, P. et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833-838 (2013).

140. Pattanayak, V. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843 (2013).

141. Cradick, T.J., Fine, E.J., Antico, C.J. & Bao, G. CRISPR/Cas9 systems targeting beta-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res* **41**, 9584-9592 (2013).

142. Ran, F.A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389 (2013).

143. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M. & Joung, J.K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* (2014).

144. Holt, N. et al. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol* **28**, 839-847 (2010).

145. Choi, S.M. et al. Efficient drug screening and gene correction for treating liver disease using patient-specific stem cells. *Hepatology* **57**, 2458-2468 (2013).

146. Osborn, M.J. et al. TALEN-based gene correction for epidermolysis bullosa. *Mol Ther* **21**, 1151-1159 (2013).

147. Dupuy, A. et al. Targeted gene therapy of xeroderma pigmentosum cells using meganuclease and TALEN. *PLoS One* **8**, e78678 (2013).

148. Bacman, S.R., Williams, S.L., Pinto, M., Peralta, S. & Moraes, C.T. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med* **19**, 1111-1113 (2013).

149. Christian, M. et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761 (2010).

150. Ding, Q. et al. A TALEN Genome-Editing System for Generating Human Stem Cell-Based Disease Models. *Cell Stem Cell* **12**, 238-251 (2013).

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

152. Mamchaoui, K. et al. Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet Muscle* **1**, 34 (2011).

153. Perez-Pinera, P., Ousterout, D.G., Brown, M.T. & Gersbach, C.A. Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Res* **40**, 3741-3752 (2012).

154. Doyle, E.L. et al. TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res* **40**, W117-122 (2012).

155. Cornu, T.I. & Cathomen, T. Quantification of zinc finger nuclease-associated toxicity. *Methods Mol Biol* **649**, 237-245 (2010).

156. Kim, Y., Kweon, J. & Kim, J.S. TALENs and ZFNs are associated with different mutation signatures. *Nat Methods* **10**, 185 (2013).

157. Lattanzi, L. et al. High efficiency myogenic conversion of human fibroblasts by adenoviral vector-mediated MyoD gene transfer. An alternative strategy for ex vivo gene therapy of primary myopathies. *J Clin Invest* **101**, 2119-2128 (1998).

158. Kimura, E. et al. Cell-lineage regulated myogenesis for dystrophin replacement: a novel therapeutic approach for treatment of muscular dystrophy. *Hum Mol Genet* **17**, 2507-2517 (2008).

159. Bhakta, M.S. et al. Highly active zinc-finger nucleases by extended modular assembly. *Genome Res* **23**, 530-538 (2013).

160. Mendell, J.R. et al. Dystrophin immunity in Duchenne's muscular dystrophy. *N Engl J Med* **363**, 1429-1437 (2010).

161. Flanigan, K.M. et al. Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort. *Hum Mutat* **30**, 1657-1666 (2009).

162. Gaj, T., Gersbach, C.A. & Barbas, C.F., 3rd ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397-405 (2013).

163. Popplewell, L. et al. Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. *Hum Gene Ther* **24**, 692-701 (2013).

164. Chavez, C.L. & Calos, M.P. Therapeutic applications of the PhiC31 integrase system. *Curr Gene Ther* **11**, 375-381 (2011).

165. Handel, E.M., Alwin, S. & Cathomen, T. Expanding or restricting the target site repertoire of zinc-finger nucleases: the inter-domain linker as a major determinant of target site selectivity. *Mol Ther* **17**, 104-111 (2009).

166. Guschin, D.Y. et al. A rapid and general assay for monitoring endogenous gene modification. *Methods Mol Biol* **649**, 247-256 (2010).

167. Fine, E.J., Cradick, T.J., Zhao, C.L., Lin, Y. & Bao, G. An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage. *Nucleic Acids Res* (2013).

168. Sander, J.D. et al. ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Res* **38**, W462-468 (2010).

169. Hockemeyer, D. et al. Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* **27**, 851-857 (2009).

170. Sander, J.D. et al. In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. *Nucleic Acids Res* **41**, e181 (2013).

171. Gabriel, R. et al. An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* **29**, 816-823 (2011).

172. Persikov, A.V., Rowland, E.F., Oakes, B.L., Singh, M. & Noyes, M.B. Deep sequencing of large library selections allows computational discovery of diverse sets of zinc fingers that bind common targets. *Nucleic Acids Res* **42**, 1497-1508 (2014).

173. Mali, P., Esvelt, K.M. & Church, G.M. Cas9 as a versatile tool for engineering biology. *Nat Methods* **10**, 957-963 (2013).

174. Cho, S.W. et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* **24**, 132-141 (2014).

175. Aartsma-Rus, A. et al. Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum Mutat* **30**, 293-299 (2009).

176. Aoki, Y. et al. Bodywide skipping of exons 45-55 in dystrophic mdx52 mice by systemic antisense delivery. *Proc Natl Acad Sci U S A* **109**, 13763-13768 (2012).

177. Aartsma-Rus, A. et al. Exploring the frontiers of therapeutic exon skipping for Duchenne muscular dystrophy by double targeting within one or multiple exons. *Mol Ther* **14**, 401-407 (2006).

178. Perez-Pinera, P. et al. RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat Methods* **10**, 973-976 (2013).

179. Ding, Q. et al. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393-394 (2013).

180. Kim, H. et al. Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat Methods* **8**, 941-943 (2011).

181. Tedesco, F.S., Dellavalle, A., Diaz-Manera, J., Messina, G. & Cossu, G. Repairing skeletal muscle: regenerative potential of skeletal muscle stem cells. *J Clin Invest* **120**, 11-19 (2010).

182. Brunet, E. et al. Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci U S A* **106**, 10620-10625 (2009).

183. Gregorevic, P. et al. Systemic delivery of genes to striated muscles using adeno-associated viral vectors. *Nat Med* **10**, 828-834 (2004).

184. van Putten, M. et al. Low dystrophin levels increase survival and improve muscle pathology and function in dystrophin/utrophin double-knockout mice. *FASEB J* **27**, 2484-2495 (2013).

185. Li, D., Yue, Y. & Duan, D. Marginal level dystrophin expression improves clinical outcome in a strain of dystrophin/utrophin double knockout mice. *PLoS One* **5**, e15286 (2010).

186. van Putten, M. et al. Low dystrophin levels in heart can delay heart failure in mdx mice. *J Mol Cell Cardiol* **69C**, 17-23 (2014).

187. Cornu, T.I. et al. DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* **16**, 352-358 (2008).

188. Ferreiro, V. et al. Asymptomatic Becker muscular dystrophy in a family with a multiexon deletion. *Muscle Nerve* **39**, 239-243 (2009).

189. van Putten, M. et al. The effects of low levels of dystrophin on mouse muscle function and pathology. *PLoS One* **7**, e31937 (2012).

190. Nishimasu, H. et al. Crystal structure of cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935-949 (2014).

191. Gordley, R.M., Gersbach, C.A. & Barbas, C.F., 3rd Synthesis of programmable integrases. *Proc Natl Acad Sci U S A* **106**, 5053-5058 (2009).

192. Li, C. et al. Single amino acid modification of adeno-associated virus capsid changes transduction and humoral immune profiles. *J Virol* **86**, 7752-7759 (2012).

193. Piacentino, V., 3rd et al. X-linked inhibitor of apoptosis protein-mediated attenuation of apoptosis, using a novel cardiac-enhanced adeno-associated viral vector. *Hum Gene Ther* **23**, 635-646 (2012).

194. Qiao, C. et al. Single tyrosine mutation in AAV8 and AAV9 capsids is insufficient to enhance gene delivery to skeletal muscle and heart. *Hum Gene Ther Methods* **23**, 29-37 (2012).

195. Qiao, C. et al. Adeno-associated virus serotype 6 capsid tyrosine-to-phenylalanine mutations improve gene transfer to skeletal muscle. *Hum Gene Ther* **21**, 1343-1348 (2010).

196. Yang, L., Li, J. & Xiao, X. Directed evolution of adeno-associated virus (AAV) as vector for muscle gene therapy. *Methods Mol Biol* **709**, 127-139 (2011).

197. Hinds, S., Bian, W., Dennis, R.G. & Bursac, N. The role of extracellular matrix composition in structure and function of bioengineered skeletal muscle. *Biomaterials* **32**, 3575-3583 (2011).

198. t Hoen, P.A. et al. Generation and characterization of transgenic mice with the full-length human DMD gene. *J Biol Chem* **283**, 5899-5907 (2008).

199. Mann, C.J. et al. Antisense-induced exon skipping and synthesis of dystrophin in the mdx mouse. *Proc Natl Acad Sci U S A* **98**, 42-47 (2001).

200. Koo, T., Popplewell, L., Athanasopoulos, T. & Dickson, G. Triple trans-splicing adeno-associated virus vectors capable of transferring the coding sequence for full-length dystrophin protein into dystrophic mice. *Hum Gene Ther* **25**, 98-108 (2014).

201. Quenneville, S.P., Chapdelaine, P., Rousseau, J. & Tremblay, J.P. Dystrophin expression in host muscle following transplantation of muscle precursor cells modified with the phiC31 integrase. *Gene Ther* **14**, 514-522 (2007).

202. Gaj, T., Mercer, A.C., Sirk, S.J., Smith, H.L. & Barbas, C.F., 3rd A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. *Nucleic Acids Res* **41**, 3937-3946 (2013).

203. Kettlun, C., Galvan, D.L., George, A.L., Jr., Kaja, A. & Wilson, M.H. Manipulating piggyBac transposon chromosomal integration site selection in human cells. *Mol Ther* **19**, 1636-1644 (2011).

204. Owens, J.B. et al. Transcription activator like effector (TALE)-directed piggyBac transposition in human cells. *Nucleic Acids Res* **41**, 9197-9207 (2013).

205. Bakowska, J.C. et al. Targeted transgene integration into transgenic mouse fibroblasts carrying the full-length human AAVS1 locus mediated by HSV/AAV rep(+) hybrid amplicon vector. *Gene Ther* **10**, 1691-1702 (2003).

206. Flanigan, K.M. et al. Anti-dystrophin T cell responses in Duchenne muscular dystrophy: prevalence and a glucocorticoid treatment effect. *Hum Gene Ther* **24**, 797-806 (2013).

207. LoDuca, P.A., Hoffman, B.E. & Herzog, R.W. Hepatic gene transfer as a means of tolerance induction to transgene products. *Curr Gene Ther* **9**, 104-114 (2009).

208. Cao, O. et al. Induction and role of regulatory CD4+CD25+ T cells in tolerance to the transgene product following hepatic in vivo gene transfer. *Blood* **110**, 1132-1140 (2007).

209. Sun, B. et al. Immunomodulatory gene therapy prevents antibody formation and lethal hypersensitivity reactions in murine pompe disease. *Mol Ther* **18**, 353-360 (2010).

210. Mingozzi, F. et al. Modulation of tolerance to the transgene product in a nonhuman primate model of AAV-mediated gene transfer to liver. *Blood* **110**, 2334-2341 (2007).

Biography

Born: Binghamton, NY on 01/31/1987

Duke University, Durham, NC, May 2012 - Master of Science, Biomedical Engineering

Cornell University, Ithaca, NY, January 2009 - Bachelor of Science, Biological Engineering

Articles and Book Chapters:

1. **David G. Ousterout**, Charles A. Gersbach. "Development of TALENs and their use in biotechnology." Methods in Molecular Biology, Book Chapter. Manuscript in preparation.
2. **David G. Ousterout***, Ami M. Kabadi*, Isaac B. Hilton, Charles A. Gersbach. "Multiplex CRISPR/Cas9 Gene Editing using a Single Lentiviral Vector." Manuscript in preparation.
3. **David G. Ousterout**, Ami M. Kabadi, Pratiksha I. Thakore, Pablo Perez-Pinera, Matthew T. Brown, Charles A. Gersbach. "Genetic Correction of Duchenne Muscular Dystrophy by Genomic Excision of Exon 51 using Zinc-Finger Nucleases." Manuscript in preparation.
4. **David G. Ousterout**, Ami M. Kabadi, Pratiksha I. Thakore, Charles A. Gersbach. "Correction of Duchenne Muscular Dystrophy by Multiplex CRISPR/Cas9-Based Genome Editing." Submitted.
5. **David G. Ousterout**, Charles A. Gersbach. "Gene Editing for Neuromuscular Diseases." In Genome Editing: The Next Step in Gene Therapy. (Cathomen, Hirsch, and Porteus, eds.) Springer. In Press.
6. Pablo Perez-Pinera, Dewran D. Kocak, Christopher M. Vockley, Andrew F., Ami M. Kabadi, Lauren R. Polstein, Pratiksha I. Thakore, Katie A. Glass, **David G. Ousterout,** Kam W. Leong, Farshid Guilak, Gregory E. Crawford, Timothy E. Reddy, Charles A. Gersbach. "RNA-guided gene activation by CRISPR-Cas9-based transcription factors." *Nature Methods,* **10**(10): 973-976, October 2013.
7. **David G. Ousterout**, Pablo Perez-Pinera, Pratiksha I. Thakore, Ami M. Kabadi, Matthew T. Brown, Xiaoxia Qin, Olivier Fedrigo, Vincent Mouly, Jacques P. Tremblay, Charles A. Gersbach. "Reading Frame Correction by Targeted Genome Editing Restores Dystrophin Expression in Cells from Duchenne Muscular Dystrophy Patients." *Molecular Therapy,* **21**(9): 1718-26, September 2013.
8. Pablo Perez-Pinera, **David G. Ousterout**, Jonathan M. Brunger, Alicia M. Farin, Katherine A. Glass, Farshid Guilak, Gregory E. Crawford, Alexander J. Hartemink, and Charles A. Gersbach. "Synergistic and Tunable Gene Activation in Human Cells by Combinations of Synthetic Transcription Factors." *Nature Methods*, **10**(3): 239-42, March 2013.
9. Bhakta, M.S., Henry, I.M., **Ousterout, D.G.**, Das, K.T., Lockwood, S.H., Meckler, J.F., Wallen, M.C., Zykovich, A., Yu, Y., Leo, L., Xu, L., Gersbach, C.A. and Segal, D. J. "Highly Active Zinc Finger Nucleases by Extended Modular Assembly." *Genome Research,* **23**(3): 530-8, December 2012.
10. Perez-Pinera, P., **Ousterout, D.G.,** Gersbach, C.A. "Advances in Targeted Genome Editing." *Current Opinion in Chemical Biology,* 16(3-4): 268-77, August 2012.
11. Perez-Pinera, P., **Ousterout, D.G.,** Brown, M.T., Gersbach, C.A. "Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases." *Nucleic Acids Research*, 16(3-4): 268-77, April 2012.
12. Regan, A.D., **Ousterout, D.G.,** Whittaker, G.R. "Feline lectin activity is critical for the cellular entry of feline infectious peritonitis virus." *Journal of Virology*, **84**(15): 7917-21, May 2010.

*Co-first authorship

Scholarships:

1. American Heart Association Mid-Atlantic Affiliate Pre-doctoral Fellowship. July 2012 - May 2014.
2. Agilent Microarray Grant Program Award ($5000), May 2013.
3. American Society of Gene and Cell Therapy Travel Grant, May 2013.
4. American Society of Gene and Cell Therapy Travel Grant, May 2012.
5. Howard Clark Graduate Fellowship in Biomedical Engineering, 2010 – 2011.