PREVENTION OF DUCHENNE MUSCULAR DYSTROPHY BY
CRISPR/CAS THERAPEUTIC GENOME EDITING

APPROVED BY SUPERVISORY COMMITTEE

Eric N. Olson, Ph.D.

Elizabeth Chen, Ph.D.

Hesham A. Sadek, M.D., Ph.D.

Joshua T. Mendell, M.D., Ph.D.

DEDICATION

In deep memory of my beloved brother and grandfather,

Yi-Ci ZHAO (1981-2003) and Cai-Xiu YU (1928-2015) for showing me the essence of love.

In dedication to my mother and father,

Min YU and Xian ZHANG for giving me the freedom and confidence to chase my dream.

ACKNOWLEDGEMENTS

Four years ago, I was driving on my way from Minneapolis to Dallas to start my Ph.D. study in University of Texas Southwestern Medical Center. During this fifteen-hour long trip, I saw new dawn shining through the light mist covered on the land of Flint Hills in eastern Kansas, as well as heavy thunderstorm sweeping across the vast land of Oklahoma. It reflected my Ph.D. study as a fantastic adventure instead of a monotonous journey. This expedition is accompanied by so many wonderful people around me, who gave me tremendous support, confidence and encouragement.

My deep gratitude first goes to my mentor, Dr. Eric Olson, who inspired me in many aspects. First of all, his sensitivity in science transformed my way of defining great scientific questions. He influenced me how to solve important questions that may completely revolutionize the scientific field. Second, he taught me there is no shortcut to great achievement and the only way is through hard-work and self-discipline. I really enjoyed the great discussion with him about research progress, especially during the weekend in a one-to-one manner. Moreover, despite the fact that he has received numerous titles and accomplished tremendous scientific achievements, he is always very humble and willing to accept new ideas. He never showed contempt in my new research ideas and always encouraged me to pursue them.

My sincere appreciation also goes to my second mentor Dr. Rhonda Bassel-Duby. I received systematic training from her in academic writing and presentation. She is always willing to address any issues I encountered during my Ph.D. study with great patience and devotion. She also plays a pivotal role in setting up collaboration with other research groups, which made my Ph.D. study smoother and easier. Most importantly, besides academic

mentorship, she also cared about me in my daily life and provided tremendous emotional support to me, which I shall never forget.

I am also really grateful to Dr. Chengzu Long for his initial mentorship in CRISPR/Cas genome editing. I knew nothing about CRISPR when I first joined the Olson lab. It was he who introduced this revolutionary technology to me and provided endless insights and inspirations to me. His scope in this broad research area with attention to scientific detail still influences me in my daily practice of scientific research.

I would also like to thank my thesis committee members, Drs. Elizabeth Chen, Joshua Mendell and Hesham Sadek for their support in my Ph.D. study. Elizabeth served as chair of my Ph.D. candidacy exam and helped me a lot in my exam proposal preparation. Josh was very generous to share critical research materials for my thesis project. Hesham also provided many insights in mitochondrial respiration and oxygen consumption topics.

It is impossible to finish my Ph.D. research on time without the help from members of the Myoediting team in the Olson lab. John McAnally is an expert in mouse embryo microinjection. His impeccable injection significantly increased my thesis research progress and helped me to publish my first research article during the first year of my Ph.D. study. Dr. Hui Li is mastered in all aspects of research techniques. I was fortunate enough to work with her during my four years of graduate study since she made my research much easier and faster. Dr. Yi-Li Min was a talented Ph.D. student when I joined the Olson lab and she generously provided a valuable DMD mouse model and taught me AAV injection to support my Ph.D. thesis research. I thank Dr. Efrain Sanchez-Ortiz for his excellent support in Western blot

analysis and immunohistochemistry. I thank other members of the Myoediting team, including Dr. Jian Huang, John Shelton, Alex Mireault, and Cristina Rodriguez for their great help.

In addition, I would like to express my gratitude to all graduate students and postdoctoral researchers of the Olson lab. I want to give special thanks to Dr. Stephen Li, who was my benchmate for four years and best friend outside the lab, for hanging out with me during leisure time.

I stayed in Minnesota for my undergraduate and I was fortunate enough to meet many friends there. I want to first express my sincere appreciation to my undergraduate mentor Dr. Michael Kyba for providing me such a great opportunity to do research in his lab. Without his strong recommendation letter, it was impossible for me to be accepted by University of Texas Southwestern Medical Center. I would also like to thank Drs. Yinan Kan, Hongliang Xu, Xueyang Pan and Chao Li for their memorable friendship during my stay in University of Minnesota.

Moreover, I am also grateful to all members of Family Baptist Church in Minnesota, especially to Pastor Lee Ormiston, Pam Ormiston, David and Susan Mowen, Jim and Peggy Murschel, Jerome and Karen Hime, Perry and Beverly Larson, and Jim Amundson for their spiritual support and encouragement.

At last, I want to express my deepest gratitude to my family members in China. I am grateful to my parents for giving me freedom and confidence to chase my dream. I want to thank my dearest brother Yi-Ci Zhao and grandpa Cai-Xiu Yu for showing me the essence of love.

PREVENTION OF DUCHENNE MUSCULAR DYSTROPHY BY CRISPR/CAS THERAPEUTIC GENOME EDITING

by

YU ZHANG

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

May, 2020

PREVENTION OF DUCHENNE MUSCULAR DYSTROPHY BY CRISPR/CAS THERAPEUTIC GENOME EDITING

Publication No. _____

Yu Zhang, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2020

Supervising Professor: Eric N. Olson, Ph.D.

Skeletal muscle is one of the largest tissues in the human body and hence muscle diseases caused by genetic mutations have a profound and systemic impact on human health. Duchenne muscular dystrophy (DMD) is a lethal neuromuscular disorder, caused by mutations in the *DMD* gene on the X chromosome, which consists of 79 exons encoding dystrophin protein. Patients with DMD develop progressive muscle weakness and cardiomyopathy, and ultimately succumb to respiratory and cardiac failure in their mid-20s. The dystrophin gene was identified three decades ago and mutations in the *DMD* gene are well-characterized. However, there is no effective treatment for this debilitating disease. The CRISPR/Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins)

was first discovered as an adaptive immune system in bacteria and archaea for defending against phage infection. Recently, the CRISPR/Cas system has been applied for mammalian genome editing because it provides site-specific DNA double-stranded breaks with simplicity and precision. In this study, I demonstrate the feasibility of using CRISPR/Cpf1 to correct a DMD exon 48-50 out-of-frame deletion mutation in cardiomyocytes derived from patient induced pluripotent stem cells by exon skipping and exon reframing strategies. Next, I precisely correct a *Dmd* exon 23 nonsense mutation in *mdx* mouse by CRISPR/Cpf1-mediated germline editing. Furthermore, I apply CRISPR/Cas9-mediated post-natal genome editing to correct a *Dmd* exon 44 out-of-frame deletion mutation in a DMD mouse model. Finally, I develop an effective strategy to improve CRISPR/Cas9-mediated in vivo genome editing by packaging Cas9 nuclease in conventional single-stranded AAV and CRISPR single guide RNAs in double-stranded self-complementary AAV. This strategy significantly reduces the amount of AAV vector needed for therapeutic genome editing and enhances dystrophin restoration after delivery into a mouse model of DMD harboring an exon 44 deletion. These findings represent an important advancement toward therapeutic translation of genome editing technology for permeant correction of Duchenne muscular dystrophy.

TABLE OF CONTENTS

TITLE i

DEDICATION ii

ACKNOWLEDGEMENTS iii

ABSTRACT viii

TABLE OF CONTENTS x

PRIOR PUBLICATIONS xiv

LIST OF FIGURES xvi

LIST OF TABLES xviii

LIST OF DEFINITIONS xix

CHAPTER ONE 1

DUCHENNE MUSCULAR DYSTROPHY AND CRISPR/CAS THERAPEUTIC GENOME EDITING 1

SKELETAL MUSCLE AND MUSCULAR DYSTROPHY 1

Skeletal Muscle Structure 1

Skeletal Muscle Regeneration and Satellite Cells 2

Muscular Dystrophy 3

Duchenne muscular dystrophy (DMD) gene structure and mutations 4

Animal models of DMD 7

GENOME EDITING 9

History of Genome Editing 9

Meganuclease 10
Zinc-finger nucleases (ZFNs) 11
Transcription activator-like effector nucleases (TALENs) 13
CRISPR/Cas 15
CRISPR/Cas-Mediated Genome Editing 19
Classical non-homologous end joining (C-NHEJ) 20
Homology-directed repair (HDR) 22
Microhomology-mediated end joining (MMEJ) 23
MYOEDITING: PREVENTION OF DUCHENNE MUSCULAR DYSTROPHY 24
Introduction of Myoediting 24
Strategies of CRISPR/Cas-Mediated DMD Correction 26
Exon deletion 26
Exon skipping 30
Exon reframing 31
Exon knock-in 32
IN VIVO GENOME EDITING OF ANIMAL MODELS 33
Delivery of Genome Editing Components by Nonviral Vectors 34
Delivery of Genome Editing Components by Viral Vectors 37

CHAPTER TWO 39
CRISPR/CPF1 CORRECTION OF MUSCULAR DYSTROPHY MUTATIONS IN HUMAN CARDIOMYOCYTES AND MICE 39

Abstract 39

Introduction 40

Results 42

Strategies for CRISPR/Cpf1-mediated genome editing of *DMD* exon 51 42

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes by CRISPR/Cpf1-mediated exon reframing 45

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes by CRISPR/Cpf1-mediated exon skipping 49

Restoration of dystrophin in *mdx* mice by CRISPR/Cpf1-mediated correction 52

Correction of muscular dystrophy in *mdx* mice by CRISPR/LbCpf1-mediated HDR or NHEJ 53

Discussion 60

Materials and Methods 62

CHAPTER THREE 69

ENHANCED CRISPR/CAS9 CORRECTION OF DUCHENNE MUSCULAR DYSTROPHY IN MICE BY A SELF-COMPLEMENTARY AAV DELIVERY SYSTEM 69

Abstract 69

Introduction 70

Results 73

Strategies for CRISPR/Cas9-mediated genome editing of *Dmd* exon 45 73

In vitro genome editing using ssAAV or scAAV-packaged sgRNA 75
Systemic delivery of scAAV-packaged sgRNAs restores dystrophin expression in ΔEx44 mice 77
Systemic delivery of scAAV-packaged sgRNAs restores muscle integrity and improves muscle-function in ΔEx44 mice 83
The scAAV system induces significant INDELs within *Dmd* exon45 and maintains higher copies of viral genome in vivo 89
Discussion 94
Materials and Methods 98

CHAPTER FOUR 106
CONCLUDING REMARKS 106
Challenges of Therapeutic Genome Editing 106
Immunogenicity 106
Off-target effects 108
Long-term effects and benefits of postnatal genome editing 110
Conclusions and Future Perspectives 110

BIBLIOGRAPHY 112

PRIOR PUBLICATIONS

1. **Zhang**, **Y**., Li, H., Min, Y.L., Sanchez-Ortiz, E., Huang, J., Mireault, A.A., Shelton, J.M., Kim, J., Mammen, P.P.A., Bassel-Duby, R., and Olson, E.N. (2020). Enhanced CRISPR/Cas9 Correction of Duchenne Muscular Dystrophy In Mice By A Self-Complementary AAV Delivery System. ***Science Advances*** 2.
2. Amoasii, L., Li, H., **Zhang**, **Y**., Min, Y.L., Sanchez-Ortiz, E., Shelton, J.M., Long, C., Mireault, A.A., Bhattacharyya, S., McAnally, J.R., Bassel-Duby, R., and Olson, E.N. (2019). In vivo non-invasive monitoring of dystrophin correction in a new Duchenne muscular dystrophy reporter mouse. ***Nature Communications*** 10, 4537.
3. Long, C., Li, H., Tiburcy, M., Rodriguez-Caycedo, C., Kyrychenko, V., Zhou, H., **Zhang**, **Y**., Min, Y.L., Shelton, J.M., Mammen, P.P.A., Liaw, N.Y., Zimmermann, W-H, Bassel-Duby, R., and Olson, E.N. (2018). Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. ***Science Advances*** 4, eaap9004.
4. **Zhang**, **Y**., Long, C., Bassel-Duby, R., and Olson, E.N. (2018). Myoediting: Toward Prevention of Muscular Dystrophy by Therapeutic Genome Editing. ***Physiological Reviews*** 98, 1205-1240.
5. **Zhang**, **Y**., Long, C., Li, H., McAnally, J.R., Baskin, K.K., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2017). CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. ***Science Advances*** 3, e1602814.

6. **Zhang**, **Y**., Lee, J.K., Toso, E.A., Lee, J.S., Choi, S.H., Slattery, M., Aihara, H., and Kyba, M. (2016). DNA-binding sequence specificity of DUX4. ***Skeletal muscle*** 6, 8.
7. Arpke, R.W., Darabi, R., Mader, T.L., **Zhang**, **Y**., Toyama, A., Lonetree, C.L., Nash, N., Lowe, D.A., Perlingeiro, R.C., and Kyba, M. (2013). A new immuno-, dystrophin-deficient model, the NSG-mdx(4Cv) mouse, provides evidence for functional improvement following allogeneic satellite cell transplantation. ***Stem Cells*** 31, 1611-1620.

LIST OF FIGURES

FIGURE 1.1 2

FIGURE 1.2 5

FIGURE 1.3 12

FIGURE 1.4 14

FIGURE 1.5 16

FIGURE 1.6 18

FIGURE 1.7 21

FIGURE 1.8 27

FIGURE 1.9 35

FIGURE 2.1 43

FIGURE 2.2 44

FIGURE 2.3 46

FIGURE 2.4 48

FIGURE 2.5 50

FIGURE 2.6 51

FIGURE 2.7 53

FIGURE 2.8 54

FIGURE 2.9 56

FIGURE 2.10 57

FIGURE 2.11 58

FIGURE 3.1 74

FIGURE 3.2 76

FIGURE 3.3 78

FIGURE 3.4 79

FIGURE 3.5 81

FIGURE 3.6 82

FIGURE 3.7 84

FIGURE 3.8 85

FIGURE 3.9 87

FIGURE 3.10 88

FIGURE 3.11 89

FIGURE 3.12 91

FIGURE 3.13 92

FIGURE 3.14 93

LIST OF TABLES

TABLE 2.1 58

TABLE 2.2 59

TABLE 2.3 68

TABLE 3.1 90

TABLE 3.2 104

LIST OF DEFINITIONS

AAV – adeno-associated virus

AAV9 – adeno-associated viruses serotype 9

alt-NHEJ – alternative nonhomologous end joining

AON – anti-sense oligonucleotide

*As*Cpf1 – Cpf1 from *Acidaminococcus sp. BV3L6*

BMD – Becker muscular dystrophy

bp – base pair

CK – creatine kinase

CMs – cardiomyocytes

C-NHEJ – classical nonhomologous end joining

CRISPR – Clustered regularly interspaced short palindromic repeats

Cascade – CRISPR-associated complex for antiviral defense

Cas9 – CRISPR-associated protein 9

Cpf1 – CRISPR from *Prevotella* and *Francisella 1*

crRNA – CRISPR RNA

DMD – Duchenne Muscular Dystrophy

DSB – double-stranded break

EDL – extensor digitorum longus muscle

FDA – Food and Drug Administration

HDR – homology-directed repair

H&E – Hematoxylin and eosin

HITI – homology-independent targeted integration

HR – homologous recombination

IM – intramuscular

INDELs – insertions and deletions

IP – intraperitoneal

iPSCs – induced pluripotent stem cells

ITR – inverted terminal repeat

KDa – kilodalton

Knt – kilonucleotides

*Lb*Cpf1 – Cpf1 from *Lachnospiraceae bacterium ND2006*

MMEJ – microhomology-mediated end joining

NHEJ – nonhomologous end joining

nt – nucleotides

ORF – Open reading frame

PAM – protospacer adjacent motif

PITCh – Precise Integration into Target Chromosome

rAAV – recombinant adeno-associated virus

RFLP – restriction fragment length polymorphism

scAAV – self-complementary adeno-associated virus

sgRNA – single guide RNA

*Sa*Cas9 – Cas9 from *Staphylococcus aureus*

*Sp*Cas9 – Cas9 from *Streptococcus pyogenes*

ssAAV – single-stranded adeno-associated virus

TA – tibialis anterior

TALENs – transcription activator-like effector nucleases

tracrRNA – trans-activating CRISPR RNA

T7E1 – T7 endonuclease I

ZFNs – Zinc finger nucleases

CHAPTER ONE

DUCHENNE MUSCULAR DYSTROPHY AND CRISPR/CAS THERAPEUTIC GENOME EDITING

Acknowledgement

Parts of this chapter, including figures, have been reproduced, with or without modifications, from my previously published work (Zhang et al., 2018).

SKELETAL MUSCLE AND MUSCULAR DYSTROPHY

Skeletal Muscle Structure

From intense body movement in Greco-Roman wrestling to delicate vocal control in coloratura soprano, skeletal muscle supports a remarkably wide range of human activities. As one of the largest tissues, skeletal muscle accounts for approximately 40% of human body weight and is essential for physical support, locomotion, energy expenditure and metabolism.

Skeletal muscle is a highly organized tissue and is composed of thousands of multinucleated myofibers, which are formed by fusion of mono-nucleated myoblasts during development and regeneration. Bundles of myofibers form a muscle fascicle and groups of fascicles contribute to the structure of a skeletal muscle (Figure 1.1, panel A). The functional unit of a myofiber is the sarcomere, which is comprised of actin thin filaments and myosin thick filaments. Sliding of the thin and thick filaments past each other generates a muscle contraction.

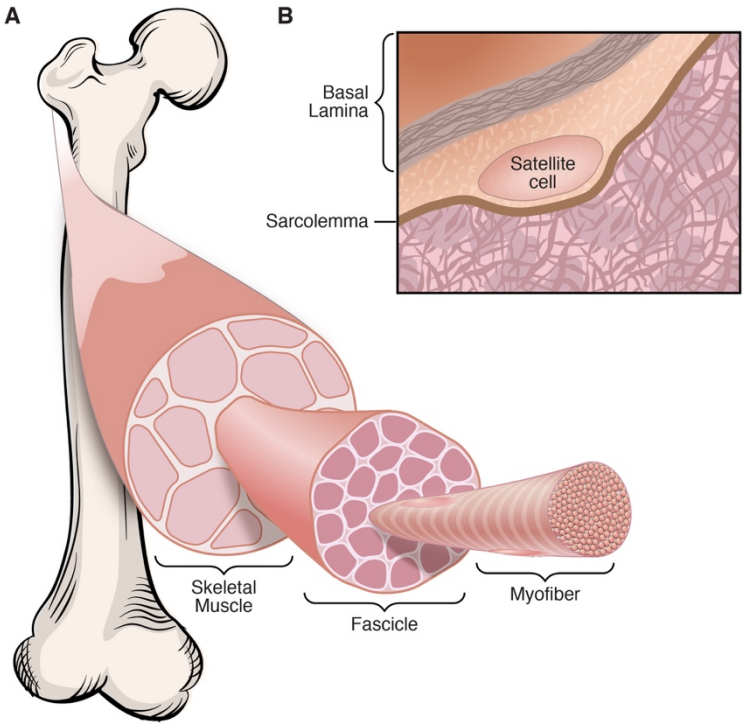

Figure 1.1 Skeletal muscle structure. (**A**) Skeletal muscle is composed of thousands of multinucleated myofibers. Bundles of myofibers form muscle fascicles and groups of fascicles contribute to skeletal muscle structure. (**B**) Satellite cells are adult skeletal muscle stem cells, which reside between the sarcolemma and basal lamina of myofibers.

Skeletal Muscle Regeneration and Satellite Cells

The adult musculature has a remarkable regenerative capacity, primarily due to the contribution of the skeletal muscle resident stem cells, known as satellite cells (Brack and Rando, 2012; Chang and Rudnicki, 2014; Yin et al., 2013). Satellite cells reside between the sarcolemma and basal lamina of myofibers and are marked by expression of a paired-box transcription factor, Pax7 (Seale et al., 2000) (Figure 1.1, panel B). Upon muscle injury, quiescent satellite cells become activated and undergo proliferation and differentiation, and finally forming multinucleated myofibers by fusion. Activated satellite cells can also undergo asymmetric division, in which one daughter cell maintains a satellite stem cell fate and the other one acquires a myogenic commitment, becoming a satellite cell committed myogenic progenitor (Conboy et al., 2007; Kuang et al., 2007; Shinin et al., 2006). Pax7 is the canonical biomarker for quiescent and activated satellite cells, and is downregulated during myogenic

differentiation. Genetic ablation experiments demonstrated that $Pax7^+$ satellite cells are indispensable for adult skeletal muscle regeneration (Lepper et al., 2011; McCarthy et al., 2011; Murphy et al., 2011; Sambasivan et al., 2011).

Skeletal muscle regeneration is evolutionarily conserved among many bilaterians, requiring involvement of satellite cells or satellite-like cells (Baghdadi and Tajbakhsh, 2017). However, in certain bilateral species such as zebra fish and adult newt, myofiber dedifferentiation is also a unique mechanism for skeletal muscle regeneration. For example, extraocular muscle regeneration in adult zebrafish involves dedifferentiation of residual myocytes, which do not express Pax7 but express Mef2c (Saera-Vila et al., 2015). Similarly, limb muscle regeneration in the adult newt requires dedifferentiation of myocytes to Pax7-negative mononuclear cells (Sandoval-Guzman et al., 2014).

Muscular Dystrophy

Despite the remarkable regenerative capacity of skeletal muscle, muscles are vulnerable to numerous disorders, including congenital myopathies, muscular dystrophies, and inflammatory myopathies. Muscular dystrophies are a large group of genetic disorders characterized by progressive weakness of multiple muscle groups. Owing to the advancement of genome research, the genetic causes of many muscular dystrophies have been identified, with many affecting sarcolemma-associated proteins, extracellular matrix proteins, glycosyltransferase enzymes as well as nuclear proteins (Kanagawa and Toda, 2006; Mercuri and Muntoni, 2013). Depending on the mutation type and disease onset, muscular dystrophies

can significantly impair the quality of life and cause premature death. There is no cure for these debilitating diseases.

Duchenne muscular dystrophy (DMD) is an X-linked recessive muscular dystrophy caused by mutations in the *DMD* gene, which encodes dystrophin (Hoffman et al., 1987). DMD is the most common type of monogenic muscular dystrophy, affecting approximately 1 in every 5,000 boys (Guiraud et al., 2015). DMD patients seem normal at birth, but within a few years they begin having trouble walking and lose ambulation between 7 to 12 years of age. Cardiac and respiratory failure causes premature death in DMD patients, often by their early 30s.

DMD gene structure and mutations

Dystrophin is a key component of the dystrophin glycoprotein complex, which is a large multicomponent protein complex essential for sarcolemma integrity and stability (Figure 1.2, panel A) (Gao and McNally, 2015; Guiraud et al., 2015). The structure of the full-length dystrophin protein can be organized into four major domains: i) the N-terminal region containing an actin-binding domain; ii) the central region containing a stretch of 24 spectrin-like repeats, forming the rod domain, which is interrupted by 4 hinge regions; iii) the cysteine-rich domain which contains several sub-domains, including a WW domain, two EF-hand-like domains and a ZZ domain, which are important for interacting with β-dystroglycan, calmodulin and ankyrin-B; and iv) the C-terminal domain which interacts with dystrobrevin and syntrophins (Ahn and Kunkel, 1993; Gao and McNally, 2015; O'Brien and Kunkel, 2001).

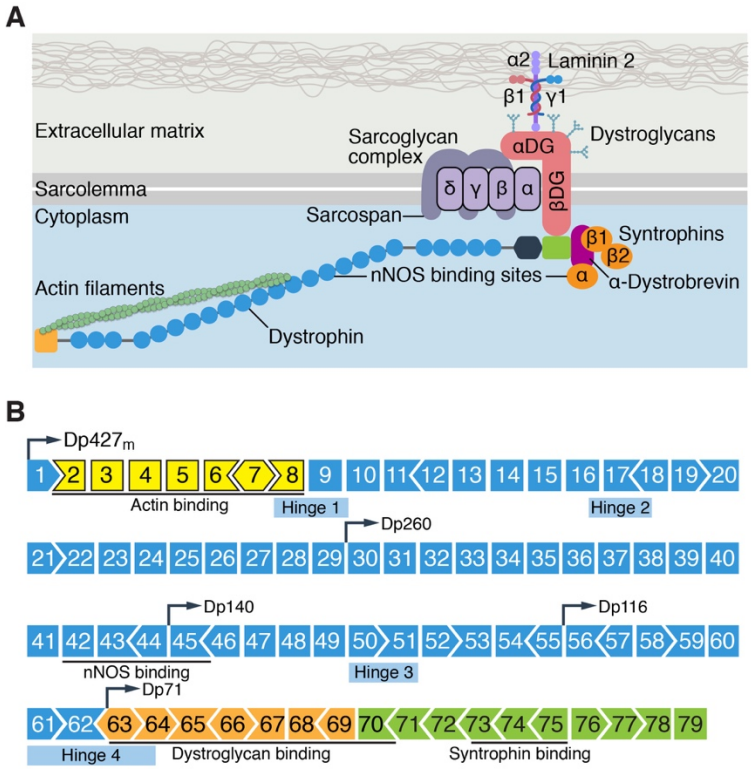

Figure 1.2 Structure of the dystrophin-glycoprotein complex and the dystrophin gene. (**A**) The main components of the dystrophin-glycoprotein complex are the dystroglycan complex, sarcoglycan complex, and dystrophin. The dystrophin-glycoprotein complex provides sarcolemma stability and integrity through interaction with laminin in the basement membrane on the extracellular matrix, and actin in the cytoplasm. Other dystrophin-associated proteins include neuronal nitric oxide synthase (nNOS), dystrobrevins, syntrophins, and sarcospan. Mutations of the main components of the DGC cause muscular dystrophies, such as Duchenne or Becker muscular dystrophy (dystrophin mutation), and limb-girdle muscular dystrophy type 2C, 2D, 2E, and 2F (sarcoglycan mutations). (**B**) The dystrophin gene has 79 exons. Different dystrophin isoforms can be transcribed from various promoters (demarcated as Dp, followed by a numeric number indicating isoform molecular weight in kiloDaltons). The dystrophin protein expressed in skeletal muscle and heart is transcribed from the Dp427m promoter. Domains essential for binding with other DGC components or cytoskeletal proteins are underlined. Exons are color-coded according to the domain they encode. N-terminus (yellow), central rod domain (blue), cysteine-rich domain (orange) and C-terminus (green).

The DMD gene, comprised of 79 exons (Figure 1.2, panel B), gives rise to different isoforms of the dystrophin protein which are expressed in various tissues by tissue-specific promoters and/or alternative splicing (Muntoni et al., 2003). The large 427 kDa cytoskeletal protein which is primarily expressed in skeletal muscle and heart is transcribed from the Dp427m promoter.

More than 7,000 mutations have been identified in the *DMD* gene (Bladen et al., 2015). These mutations can be categorized as deletion (68%), duplication (11%) of single or multiple exons, or small point mutations (20%), such as missense and nonsense substitutions (Aartsma-Rus et al., 2006; Bladen et al., 2015; Monaco et al., 1988). Mutations in the *DMD* gene are not uniformly distributed but cluster into hot spots, which are clustered within exons 2–20 and exons 45–55 (Bladen et al., 2015). Approximately 15% of all exon deletion events and 50% of all exon duplication events are observed within exons 2-20; whereas, 70% of all exon deletion events and 15% of all exon duplication events are observed within exons 45-55 (Bladen et al., 2015; Yang et al., 2013b). In-frame deletion or duplication of exon(s) within the central region of the *DMD* gene retains the protein reading frame and generates either a truncated or extended dystrophin protein. These mutant dystrophin proteins retain their N- and C-termini, which are essential for actin cytoskeleton and dystrophin glycoprotein complex interaction, leading to a milder form of muscular dystrophy, known as Becker muscular dystrophy (BMD) (Aartsma-Rus et al., 2016; Guiraud et al., 2015). In contrast, out-of-frame deletion or duplication of exon(s) either disrupts the protein reading frame or generates a premature termination codon and leads to DMD.

Animal models of DMD

The most commonly used animal model for DMD is the *mdx* mouse, in which a C-to-T transition in exon 23 creates a nonsense mutation, leading to loss of full-length dystrophin expression (Bulfield et al., 1984; Sicinski et al., 1989). The *mdx* mice do not develop severe DMD phenotypes, such as muscle wasting, scoliosis and cardiomyopathy until reaching 15 months of age. In contrast to DMD patients whose lifespan is significantly reduced, the lifespan of *mdx* mice is reduced by only 25% (Chamberlain et al., 2007). Four chemically induced *mdx* strains have also been developed, known as mdx^{2cv}, mdx^{3cv}, mdx^{4cv}, mdx^{5cv}, with a point mutation in intron 42, intron 65, exon 53 or exon 10, respectively (Chapman et al., 1989). In addition to the *mdx* strains with point mutations, four additional DMD mouse models have been established with either exon 2 duplication, exon 45 deletion, exon 50 deletion or exon 52 deletion (Araki et al., 1997; Vulin et al., 2015; Young et al., 2017).

Because dystrophin-deficient mouse models generally do not develop severe pathological phenotypes as seen in DMD patients. Several double knockout (dKO) mouse models were generated, in which the *Dmd* gene was knocked out, along with additional genes required for sarcolemma integrity, stem cell maintenance and muscle homeostasis (Deconinck et al., 1997; Grady et al., 1997; McGreevy et al., 2015; Mourkioti et al., 2013; Sacco et al., 2010). Genome editing technology also played a role in expanding the rodent models of DMD. For example, two DMD rat models were created by TALEN- or CRISPR/Cas-mediated targeting of the *Dmd* exon 23 or exons 3-6, leading to an exon 23 frame shifting or exon 3-6 deletion (Larcher et al., 2014; Nakamura et al., 2014). Most recently, CRISPR/Cas9 was used

to create two mouse models lacks exon 50 or 44, representing two most common “hot-spot” mutations in humans (Amoasii et al., 2017; Min et al., 2019).

In addition to small rodent models, large animal models of DMD have been developed, including dog (Atencia-Fernandez et al., 2015; Kornegay et al., 2012; Schatzberg et al., 1999; Smith et al., 2011; Valentine et al., 1986; Walmsley et al., 2010; Winand et al., 1994), pig (Klymiuk et al., 2013; Selsby et al., 2015; Yu et al., 2016) and non-human primates (Chen et al., 2015). Monkey models of DMD are still at F_0 with mosaicism, which requires additional breeding to generate a pure background (Chen et al., 2015). Disease progression in some porcine models of DMD is so severe that the majority of the affected pigs die within the first week of life, which limits its application in therapeutic translation (Selsby et al., 2015). In contrast, canine DMD models share more similar clinical phenotypes as seen in human patients, including limb muscle fibrosis, joint contracture, hypersalivation and an early cardiac defect (McGreevy et al., 2015). Moreover, canine DMD models have fewer regenerated myofibers than *mdx* mice as indicated by central nucleation, which is histologically similar to human patients (Cozzi et al., 2001; Smith et al., 2011). In addition, canine DMD models develop limb muscle weakness at 2 to 3 months of age and have approximately 75% reduction of lifespan, showing similar disease progression as human patients (Valentine et al., 1988). Therefore, the canine model of DMD seems superior to the other large animal models in regard to current availability, genetic background and speed of disease progression.

GENOME EDITING

History of Genome Editing

Three decades ago the laboratories of Mario Capecchi and Oliver Smithies independently developed methods for homologous recombination (HR)-mediated mammalian gene targeting technology by providing mammalian cells with exogenous plasmid DNA containing sequence homology to the endogenous genome (Doetschman et al., 1987; Mansour et al., 1988; Smithies et al., 1985; Thomas and Capecchi, 1987; Thomas et al., 1986). This HR-mediated technology allows precise gene knockout or correction of genetic mutations. HR-mediated embryonic stem cell gene targeting, together with mouse chimeras and germline transmission technologies developed by the laboratory of Martin Evans (Bradley et al., 1984) paved the way for the generation of "knockin" and "knockout" animal models, which significantly expanded our knowledge of gene function and advanced many fields of biological research. However, because DNA double-strand breaks (DSBs) occur randomly in the genome, the frequency of HR-mediated gene targeting is low (between 10^{-6} and 10^{-4}, depending on the length of sequence homology of the targeting vector) (Deng and Capecchi, 1992). Moreover, screening of correctly targeted clones requires positive-negative selection and/or Southern blot analysis, which is time-consuming and labor-intensive (Capecchi, 2005). Therefore, routine application of the conventional HR-mediated gene targeting technology for studying gene function was not feasible at that time.

In the early '90s, it was discovered that HR-mediated gene targeting efficiency could be enhanced by more than 100-fold, when the DNA DSBs were initiated at the target region by providing mammalian cells with a rare-cutting meganuclease discovered in yeast (Rouet et

al., 1994). This discovery stimulated the development of programmable nucleases for creating site-specific DNA DSBs. Within the past two decades, four major classes of nucleases have been engineered, which are: i) meganucleases (Smith et al., 2006), ii) zinc-finger nucleases (ZFNs) (Miller et al., 2007; Urnov et al., 2005), iii) transcription activator-like effector nucleases (TALENs) (Boch et al., 2009; Christian et al., 2010; Miller et al., 2011; Moscou and Bogdanove, 2009), and iv) CRISPR/Cas endonucleases (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins) (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013b; Zetsche et al., 2015).

Permanent correction of genetic mutations that contribute to monogenic neuromuscular disorders offers the ultimate treatment for these diseases. Early attempts at genome editing for treatment of muscular dystrophies were challenged by low efficiency, cytotoxicity and delivery issues (Gao et al., 2016; Li et al., 2015; Maggio et al., 2016b; Ousterout et al., 2015b; Ousterout et al., 2013; Popplewell et al., 2013; Turan et al., 2016; Xia et al., 2015). The newly discovered CRISPR/Cas system has been effectively used in genome engineering and represents a new approach to therapeutic genome editing.

Meganucleases

Meganucleases are engineered homing endonucleases, which were initially discovered in archaea, bacteria, and unicellular eukaryotic genomes (Stoddard, 2005, 2011). Unlike conventional Type II restriction endonucleases that recognize short 4-8 base pairs (bp) of palindromic DNA sequences (Pingoud et al., 2014), meganucleases require extended DNA recognition sequences (typically 16-18 bp) to generate site-specific DNA DSBs

(Chandrasegaran and Carroll, 2016). Meganucleases have been used to enhance HR-mediated gene targeting efficiency by introducing site-specific DNA DSBs in cultured mammalian cells and plants (Chiurazzi et al., 1996; Choulika et al., 1995; Donoho et al., 1998; Puchta et al., 1996; Rouet et al., 1994), but they have not been widely adopted for genome engineering because the DNA-recognition domain and the nuclease domain overlap (Silva et al., 2011; Smith et al., 2006). This overlap may adversely affect the catalytic activity of the nuclease domain (Arnould et al., 2006), making it very challenging to engineer the DNA recognition domain for specificity in new-sequence binding.

To address this issue, researchers began to focus on the Type IIS restriction enzyme, *Fok*I, which has two separate domains for DNA recognition and cleavage (Kim et al., 1994; Li and Chandrasegaran, 1993; Li et al., 1992; Li et al., 1993). They engineered novel chimeric *Fok*I endonucleases with new DNA sequence specificities by swapping DNA-binding domains from other transcription factors, such as the *Drosophila* Ubx homeodomain (Kim and Chandrasegaran, 1994), yeast Gal4 domain (Kim et al., 1998), zinc-finger protein (Kim et al., 1996), and TAL effector (Christian et al., 2010). The latter two chimeric *Fok*I endonucleases paved the way for the development of ZFNs and TALENs, respectively.

Zinc-finger nucleases (ZFNs)

ZFNs are chimeric endonucleases containing multiple Cys_2-His_2 zinc-finger domains at the amino terminus (N-terminus) for DNA-binding and a *Fok*I nuclease domain at the carboxyl terminus (C-terminus) for DNA cleavage (Figure 1.3) (Kim et al., 1996). Each individual zinc-finger domain contains approximately 30 amino acids folded in a ββα

arrangement and contacts 3 bp of DNA sequence (Pavletich and Pabo, 1991). Each ZFN monomer consists of 3-6 individual zinc-finger domains, and thus can bind to 9-18 bp of DNA sequence.

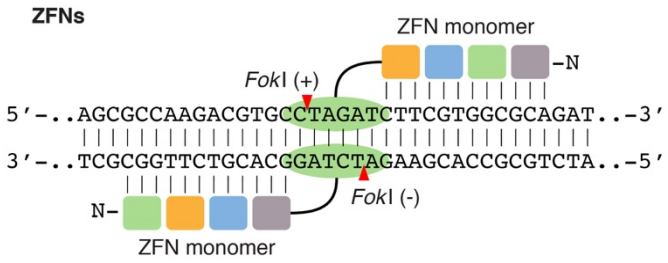

Figure 1.3 Zinc-finger nucleases (ZFNs) used for genome editing. A schematic illustration of a pair ZFN monomers bound to DNA. ZFN is a chimeric endonuclease composed of multiple zinc finger domains (colored boxes) at the N-terminus for DNA binding and a *Fok*1 nuclease domain (green oval) at the C-terminus for DNA cleavage. Dimerization of two *Fok*1 nucleases induces a DNA double-strand break (DSB) with 4 bp of 5' overhang.

Two approaches have been applied to improve genome targeting specificity and expand the targeting range of ZFNs. The first one is to engineer the wild-type (WT) *Fok*I nuclease to reduce the formation of cleavage-competent homodimers (Miller et al., 2007; Szczepek et al., 2007). The engineered ZFNs require hetero-dimerization to form a functional nuclease, in which two monomers are separated by 5-7 bp of spacer (Bitinaite et al., 1998). The second approach is to engineer zinc-finger domains for unique triplet DNA binding specificity by combinatorial library selection and/or oligomerized pool engineering (OPEN) (Bhakta et al., 2013; Dreier et al., 2001; Dreier et al., 2005; Gonzalez et al., 2010; Kim et al., 2011; Maeder et al., 2008; Segal et al., 1999).

These two approaches paved the way for modular assembly of customized ZFNs, which have been used for genome targeting in cultured cells, animals, and plants (Gaj et al., 2013;

Perez-Pinera et al., 2012; Segal and Meckler, 2013; Urnov et al., 2010). However, because of the context-dependent effects between adjacent zinc-finger domains, large-scale assembly of functional ZFNs remains challenging and cytotoxicity caused by off-target effects is also a critical issue (Gabriel et al., 2011; Pattanayak et al., 2011; Sander et al., 2013). Moreover, the genome targeting density of ZFNs is also limited because the engineered zinc-finger domains cannot target all 64 possible triplet DNA sequences, especially 5'-TNN-3' sequences (N represents any nucleotide) (Bhakta and Segal, 2010). These obstacles prevent wide application of ZFNs for genome engineering.

Transcription activator-like effector nucleases (TALENs)

TALENs are chimeric endonucleases that contain multiple DNA-binding domains, known as transcription activator-like effectors (TALEs), at the N-terminus, and a *Fok*I nuclease domain at the C-terminus for DNA cleavage (Figure 1.4) (Christian et al., 2010). Unlike the zinc-finger domain in ZFNs, which binds to a triplet DNA sequence, the TALE domain, consisting of 33-35 amino acids in tandem arrays, recognizes a single base pair (Deng et al., 2012; Mak et al., 2012). The sequence specificity of each TALE repeat is determined by the 12^{th} and 13^{th} amino acids at the TALE domain, known as repeat variable diresidues (RVDs) (Boch et al., 2009; Moscou and Bogdanove, 2009). Similar to ZFNs, functional TALENs require dimerization of the *Fok*I nuclease domain with each TALE arm targeting 15-20 bp of DNA sequence separated by 12-21 bp of spacer. TALENs have been widely used to target genomes of various species including cultured cells, animals and plants (Gaj et al., 2013; Joung and Sander, 2013; Segal and Meckler, 2013; Sun and Zhao, 2013). Although many cloning

methods have been developed for the construction of functional TALENs, such as Type II restriction enzyme-based Golden Gate assembly (Cermak et al., 2011), solid-phase assembly (Briggs et al., 2012; Reyon et al., 2012) and ligation-independent cloning (Schmid-Burgk et al., 2013), modular assembly of customized TALENs is still challenging and time-consuming because each TALEN arm consists of up to 20 highly repetitive TALE arrays.

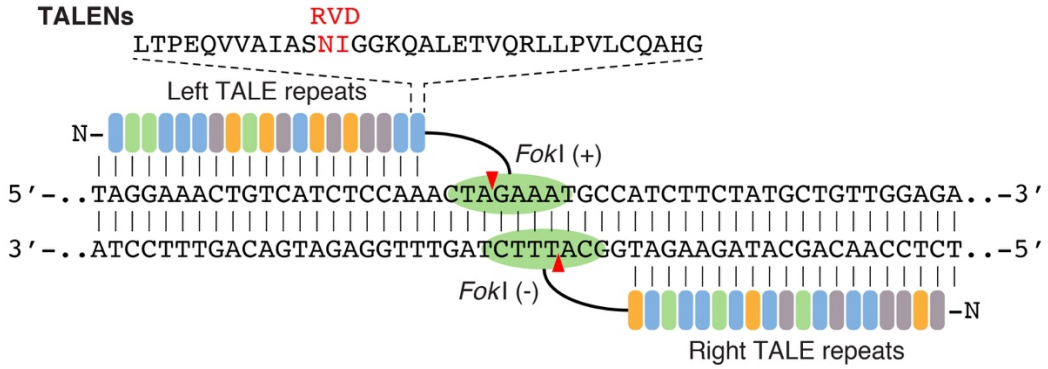

Figure 1.4 Transcription activator-like effector nucleases (TALENs) used for genome editing. A schematic illustration of a pair of TALENs bound to DNA. TALEN is a chimeric endonuclease composed of multiple TALE repeats (colored rectangles) at the N-terminus and a *Fok*1 nuclease domain (green oval) at the C-terminus. Each TALE repeat recognizes 1 bp of DNA and the sequence specificity is determined by repeat-variable diresidues (RVD; shown in red). TALEN-mediated DNA DSBs are induced by dimerization of two *Fok*1 nucleases.

Despite the difficulty of assembling TALE arrays, TALENs still offer many advantages over other programmable nucleases. First, TALENs have the highest genome-targeting density compared to ZFNs and CRISPR/Cas because each TALE array recognizes DNA sequence at single nucleotide resolution (Boch et al., 2009; Moscou and Bogdanove, 2009). Second, TALENs have minimal off-target effects because a functional TALEN requires dimerization of two TALEN pairs, which can bind 30-40 bp of DNA sequence (Deng et al., 2012; Kim and Kim, 2014; Mak et al., 2012). Therefore, TALENs offer benefits for genome engineering.

CRISPR/Cas

The discovery of CRISPR can be dated back to 1987, when a Japanese research group identified a series of directed repeats interspaced with short spacer sequences in the genome of *Escherichia coli*, although the function of these repeats was unknown at that time (Ishino et al., 1987). It was not until the mid-2000s that researchers discovered that these directed repeats are widely present in over 40% of sequenced bacteria and 90% of archaea genomes (Mojica et al., 2000), and found that the short spacer sequences between the directed repeats are of plasmid and viral origin (Bolotin et al., 2005; Mojica et al., 2005; Pourcel et al., 2005). After realizing that the CRISPR locus is actively transcribed and the protein product has potential nuclease and helicase activities, scientists proposed that the CRISPR/Cas system functions as an adaptive immune system in bacteria and archaea to defend against viral infection (Barrangou et al., 2007; Bolotin et al., 2005; Brouns et al., 2008; Haft et al., 2005; Jansen et al., 2002; Makarova et al., 2006; Marraffini and Sontheimer, 2008; Pourcel et al., 2005). The CRISPR/Cas system can be grouped into two classes and six subtypes: the Class 1 system encodes multiple effector proteins forming a Cascade complex (CRISPR-associated complex for antiviral defense) with their corresponding signature proteins, such as Cas3, Cas10 and Csf1 from Type I, III and IV CRISPR systems, respectively (Makarova et al., 2011a; Makarova et al., 2011b; Makarova et al., 2015; Shmakov et al., 2015). The Class 2 system encodes a single Cas protein with multiple functions, including Cas9, Cpf1 and Cas13a/C2c2 from Type II, V and VI CRISPR systems (Abudayyeh et al., 2016; East-Seletsky et al., 2017; Shmakov et al., 2015; Zetsche et al., 2015).

The mechanism of CRISPR immunity in bacteria and archaea varies between different

CRISPR types, but generally can be divided into three elements, which are protospacer acquisition, precursor CRISPR RNA (pre-crRNA) processing, and crRNA-guided cleavage of exogenous nucleic acids (Marraffini, 2015; Wright et al., 2016). Most CRISPR immunity requires a protospacer adjacent motif (PAM) located next to the crRNA target region in the exogenous invading genome. (Hsu et al., 2014; Wright et al., 2016).

Owing to the simplicity of the Class 2 CRISPR system in which only one RNA-guided endonuclease is required for nucleic acid cleavage, scientists engineered Cas9 endonuclease in conjunction with a hybrid crRNA-tracrRNA duplex, known as single guide RNA (sgRNA), for efficient site-specific genome cleavage in eukaryotic cells (Figure 1.5) (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013b). Currently, the most widely used Cas9 endonuclease is from *Streptococcus pyogenes* with 5'-NGG-3' or 5'-NAG-3' PAM preference. Other Cas9 orthologs are also available for genome targeting, including Cas9 endonucleases from *Staphylococcus aureus* (Ran et al., 2015), *Neisseria meningitides* (Hou et al., 2013) and *Streptococcus thermophilus* (Magadan et al., 2012; Muller et al., 2016), although these Cas9 orthologs recognize longer and more complicated PAM sequences.

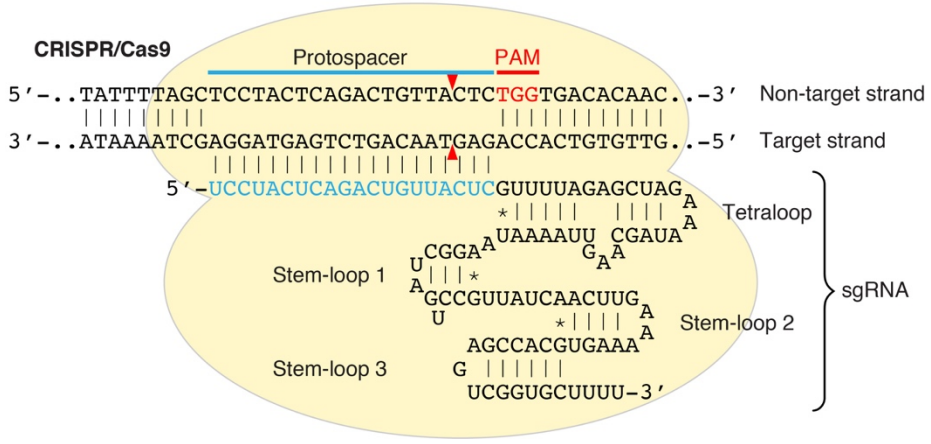

Figure 1.5 Clustered regularly interspaced short palindromic repeats-CRISPR associated protein 9 (CRISPR/Cas9) used for genome editing. A schematic illustration of the engineered CRISPR/Cas9 system from *Streptococcus pyogenes*. In the CRISPR/Cas9 system, target recognition is mediated by DNA hybridization with a single guide RNA (sgRNA), which is an engineered RNA chimera composed of CRISPR RNA (crRNA) and *trans*-activating crRNA (tracrRNA). CRISPR/Cas9-mediated DNA DSB requires a protospacer adjacent motif (PAM; shown in red) and cleavage is induced at the nucleotide 3 bp proximal to the PAM. Red arrowheads indicate cleavage site.

Besides the Type II CRISPR/Cas9 system, the most recently discovered Type V CRISPR effectors including Cpf1 (Zetsche et al., 2015) further expand the range of genome editing and nucleic acid detection. CRISPR/Cpf1, from the Class 2 Type V CRISPR system, is a RNA-guided endonuclease capable of DNA cleavage (Figure 1.6) (Fonfara et al., 2016; Zetsche et al., 2015). Two Cpf1 orthologs, *Lb*Cpf1 (from *Lachnospiraceae bacterium ND2006*) and *As*Cpf1 (from *Acidaminococcus sp. BV3L6*) have been engineered for genome editing in a variety of systems, including mammalian cells, animals, and plants (Endo et al., 2016; Hur et al., 2016; Jiang et al., 2017; Kim et al., 2017; Kim et al., 2016; Ma et al., 2017b; Murovec et al., 2017; Port and Bullock, 2016; Toth et al., 2016; Ungerer and Pakrasi, 2016; Zaidi et al., 2017; Zhang et al., 2017b). The CRISPR/Cpf1 system has many unique features compared to CRISPR/Cas9: (i) Cas9-mediated genome cleavage requires two RNA components consisting of a crRNA and a tracrRNA (which can be engineered as a single sgRNA hybrid), whereas Cpf1-mediated genome cleavage is tracrRNA-independent, so it only requires a short crRNA. (ii) The PAM sequence of Cpf1 is 5'-TTTN-3', located at the 5' end of a protospacer; in contrast, the 5'-NGG-3' or 5'-NAG-3' PAM for SpCas9 is located at the 3' end of a protospacer. (iii) Cas9-mediated DNA DSB is blunt-ended and proximal to the

PAM site, whereas Cpf1-mediated DNA DSBs are cleaved as a staggered cut distal to the PAM site. (iv) The pre-crRNA processing in the CRISPR/Cas9 system is catalyzed by an additional RNase III, whereas Cpf1 has intrinsic RNase activity and can directly process pre-crRNA by itself (Fonfara et al., 2016; Zetsche et al., 2015).

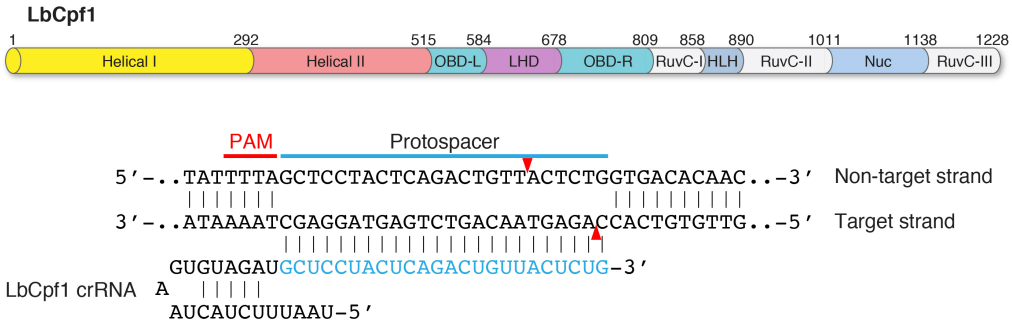

Figure 1.6 CRISPR from *Prevotella* and *Francisella 1* (CRISPR/Cpf1) used for genome editing. A novel Class 2 CRISPR/Cpf1 system has been engineered for RNA-guided DNA cleavage. Domain organization of the *Lb*Cpf1 protein discovered in *Lachnospiraceae bacterium ND2006*. All Cpf1 orthologs have two nuclease domains: i) the RuvC domain which cleaves the non-target DNA strand and ii) the Nuc domain which cleaves the target DNA strand. The *Lb*Cpf1 crRNA is shown hybridizing with its DNA target. The PAM is highlighted in red. Red arrowheads indicate cleavage site.

Because of its T-rich PAM preference, Cpf1 represents an alternative to Cas9 for genome editing at AT-rich loci. In addition to the canonical 5'-TTTN-3' PAM sequence, Cpf1 also recognizes 5'-CTTV-3', 5'-TCTV-3', 5'-TTCV-3' (V represents A, G, or C) as non-canonical PAMs, because the PAM-binding channel of Cpf1 has conformational flexibility (Yamano et al., 2017) that further expands the targeting range of the CRISPR/Cpf1 system. Another advantage of Cpf1 compared to Cas9 is the convenience of multiplex genome editing. CRISPR/Cas9-mediated multiplex genome editing requires multiple sgRNAs transcribed from separate promoters or additional RNA sequences for recognition and cleavage by other

nucleases if multiple sgRNAs are transcribed from a single promoter (Kabadi et al., 2014; Sakuma et al., 2014; Tsai et al., 2014; Xie et al., 2015). However, CRISPR/Cpf1-mediated multiplex genome editing only requires a single promoter for the transcription of multiple crRNAs, because Cpf1 can process polycistronic crRNAs into individual ones using its own RNase activity, which significantly simplifies multiplex genome editing (Zetsche et al., 2017). Therefore, Cpf1 is more than an alternative to Cas9 in terms of genome and epigenome editing but offers a broader range of genomic editing options.

CRISPR/Cas-mediated Genome Editing

On average, each human cell undergoes approximately 50 spontaneous DNA DSBs during each cell cycle (Vilenchik and Knudson, 2003). DNA DSBs occur randomly, so the efficiency of HR-mediated gene targeting in the absence of programmable nucleases is extremely low (Deng and Capecchi, 1992). The RNA-guided CRISPR/Cas system significantly enhances and simplifies genome editing, in which the Cas9-sgRNA ribonucleoprotein complex binds to DNA by base-pairing with sgRNA, generating a site-specific DNA DSB adjacent to the PAM sequence. Depending on the cell cycle stage and repair machinery, the DNA DSBs can be repaired by error-prone non-homologous end joining (NHEJ) or by accurate homology-directed repair (HDR). Additionally, there is a third DNA DSB repair pathway known as microhomology-mediated end joining (MMEJ), which is a sub-type of alternative NHEJ (alt-NHEJ).

Classical Non-homologous end joining (C-NHEJ)

Classical NHEJ (C-NHEJ) DNA repair machinery is triggered when a CRISPR/Cas-induced DNA DSB occurs in the absence of a repair template (Figure 1.7, panel A). Although C-NHEJ is active in all stages of the cell cycle, it occurs preferentially during the G1 phase when the DNA-end resection activity is low (Ira et al., 2004). The end of a DNA DSB is recognized by Ku70/Ku80 heterodimers, which recruit and activate the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs). Depending on the nature of the break, the ends of DNA DSBs can be directly ligated by the DNA ligase IV-XRCC4 complex, or requires additional processing steps, such as end resection by Artmis, WRN, or APLF nucleases and nucleotide synthesis by DNA polymerases μ and λ (Ciccia and Elledge, 2010; Davis and Chen, 2013).

DNA DSBs repaired by C-NHEJ usually generate insertions or deletions (INDELs). Depending on the location of the site-specific cleavage, C-NHEJ has been used for different purposes of genome editing. The most widely used application of C-NHEJ is gene disruption, because INDELs often cause a frameshift of an exon and subsequently disrupt gene function, resulting in a gene knockout. However, C-NHEJ can also cause exon skipping if the INDELs disrupt the splice acceptor site (Li et al., 2015), or an exonic splice enhancer/silencer sequence (Mou et al., 2017); although the outcome of the latter scenario is less predictable. Depending on the reading frame of the skipped exon and adjacent exons, exon skipping can cause gene knockout when the newly spliced adjacent exons are out of frame. Conversely, exon skipping can also produce a truncated protein if the newly spliced adjacent exons are perfectly in frame with each other.

C-NHEJ was generally considered as an error-prone DSB repair pathway. However, some recent studies also demonstrated the precision of C-NHEJ (Auer et al., 2014; Maresca et al., 2013) and used this repair pathway for homology-independent targeted integration (HITI) of DNA fragments into post-mitotic cells and animals, further expanding the application of CRISPR/Cas-mediated C-NHEJ in genome editing (Suzuki et al., 2016).

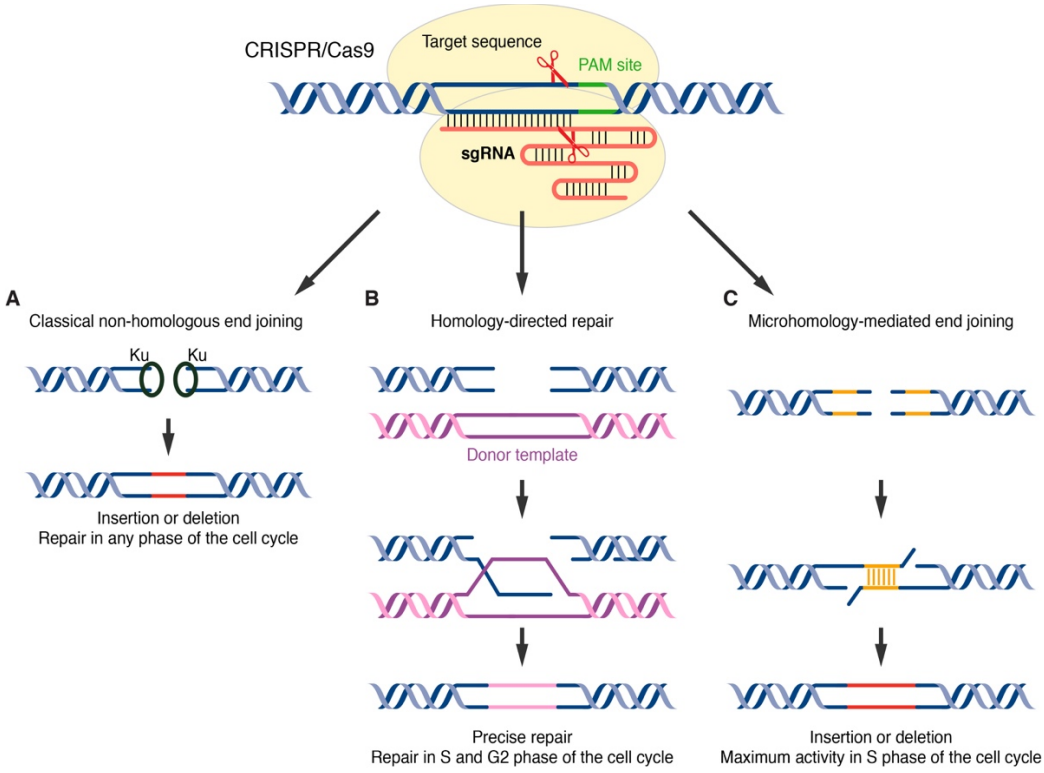

Figure 1.7 DNA repair pathways involved in CRISPR/Cas-induced DNA double-strand break repair. *A.* Classical non-homologous end joining (c-NHEJ) is a Ku-dependent DNA repair pathway that is active throughout the cell cycle. In the absence of a donor template, c-NHEJ generates insertions or deletions (INDELs; shown in red) in the genome. ***B.*** When a DNA double-strand break (DSB) is induced in the S or G2 phase of the cell cycle, homology-directed repair (HDR) can be triggered if a donor template is present (magenta), leading to precise repair of the genome. ***C.*** A Ku-independent microhomology-mediated end joining (MMEJ) pathway can be used for DNA DSB repair if the DNA breakage site shares sequence homology. MMEJ-mediated repair generates INDELs in the genome (red).

Homology-directed repair (HDR)

DNA DSBs can also be repaired by homology-directed repair (HDR) during S and G2 phases of the cell cycle, when sister chromatids can be used as a template for HR (Figure 1.7, panel B) (You and Bailis, 2010). During HDR, the end of the DNA DSB is recognized by the MRE11-RAD50-NBS1 (MRN) complex, which undergoes initial DNA end resection induced by MRE11 (Williams et al., 2007), followed by extensive end resection induced by the EXO1-BLM complex (Bolderson et al., 2010), producing single-stranded DNA (ssDNA). The exposed ssDNA is coated by RPA until RAD51 detects the homology sequence, leading to strand invasion and Holliday junction formation (West, 2003). HDR is completed when the Holliday junction is either dissolved by the BLM/TOPOIII complex or resolved by GEN1 or SLX1/SLX4 nucleases (Wyatt and West, 2014).

Currently, the most widely used repair templates for HDR are double-stranded DNA (circular or linearized plasmid) and single-stranded oligodeoxynucleotides (ssODNs). Before programmable nucleases were employed in HR-mediated gene targeting, the length of sequence homology on the targeting vector for HDR could be up to 14 kb for efficient gene targeting (Deng and Capecchi, 1992). The development of programmable nucleases, especially the CRISPR/Cas system significantly enhanced site-specific DNA DSBs and further reduced the length of sequence homology on the targeting vector to several hundred base pairs (Maruyama et al., 2015; Zhang et al., 2017a).

A ssODN can also serve as a repair template for HDR, especially for introducing small DNA modifications (Long et al., 2014; Richardson et al., 2016; Wu et al., 2013; Zhang et al., 2017b). Interestingly, asymmetrical ssODN complementary to the non-target strand (the DNA

strand that does not base pair with the CRISPR sgRNA) can drive the efficiency of HDR up to 60%, because the Cas9 endonuclease first releases the PAM-distal non-target DNA strand, which is more available for ssODN binding (Richardson et al., 2016). Due to the accuracy of HDR, it is possible that the Cas9 endonuclease will continuously generate DSBs at the target site as long as the PAM and the sgRNA target sequence remain intact, even when HDR is completed. Because of codon degeneracy, introducing a silent mutation at the third nucleotide of the triplet codon for the disruption of the PAM and/or sgRNA target sequence can effectively overcome the re-cleavage event (Long et al., 2014; Zhang et al., 2017b).

Microhomology-mediated end joining (MMEJ)

Microhomology-mediated end joining (MMEJ) is a Ku-independent alt-NHEJ repair pathway for DNA DSBs, which displays maximal activity in S phase (Figure 1.7, panel C) (Truong et al., 2013). Similar to HDR, MMEJ undergoes initial DNA end resection induced by MRE11 but does not require extensive end resection induced by the EXO1-BLM complex (Rass et al., 2009; Truong et al., 2013; Xie et al., 2009). If microhomology is present, the exposed ssDNA ends generated by initial DNA end resection will anneal with each other and the gap between the newly annealed ssDNA will be filled by DNA polymerases θ (Chan et al., 2010; Yu and McVey, 2010) and finally ligated by the LIG3-XRCC1 complex (Audebert et al., 2004; Simsek et al., 2011; Wang et al., 2005a).

In the absence of template DNA, MMEJ is an error-prone DNA repair pathway because of INDEL formation (Sfeir and Symington, 2015; Truong et al., 2013). However, several studies have adopted MMEJ for precise integration of exogenous reporter genes into the

genome after TALEN or CRISPR/Cas9-mediated DNA DSBs (Hisano et al., 2015; Nakade et al., 2014; Sakuma et al., 2016). This method, known as Precise Integration into Target Chromosome (PITCh), requires three DNA DSBs, with one DSB located at the target locus in the genome, and the other two DSBs located at the 5'- and 3'-ends of a reporter cassette (e.g. GFP). The reporter cassette is cloned into a plasmid with 5-25 bp of sequence homology to the target locus, serving as the MMEJ repair template. Therefore, MMEJ provides an alternative method for precise genome editing similar to HDR.

MYOEDITING: PREVENTION OF DUCHENNE MUSCULAR DYSTROPHY

Introduction of Myoediting

To date, more than 800 monogenic neuromuscular disorders with mutations in over 400 different genes have been recorded (Bonne et al., 2017). The discovery and application of programmable nucleases for genome editing paves the way for permanent correction of these genetic diseases (Cox et al., 2015; Porteus, 2015; Prakash et al., 2016). Meganucleases, ZFNs and TALENs have been reported to correct mutations responsible for certain muscular dystrophies including Duchenne muscular dystrophy (DMD), limb-girdle muscular dystrophy (LGMD), and myotonic dystrophy (DM) (Gao et al., 2016; Li et al., 2015; Maggio et al., 2016b; Ousterout et al., 2015b; Ousterout et al., 2013; Popplewell et al., 2013; Turan et al., 2016; Xia et al., 2015). However, these early versions of programmable nucleases were not widely adopted for correcting mutations in various muscular dystrophies because of the low genome targeting density, difficulty of assembly of the functional nuclease domains and cytotoxicity caused by off-target effects. The CRISPR/Cas system revolutionized the genome

editing field and significantly simplified the process of permanent correction of monogenic neuromuscular disorders.

CRISPR/Cas-mediated genome editing in skeletal muscle and heart, which we termed myoediting (Long et al., 2016; Long et al., 2014), can permanently correct various *DMD* mutations and restore dystrophin function. Initially, myoediting was performed in the germline of *mdx* mice, a mouse model of DMD with a nonsense mutation in exon 23. By injecting Cas9 mRNA, a sgRNA targeting the mutated exon 23, and a ssODN repair template into the zygotes of *mdx* mice, it was demonstrated that CRISPR/Cas9-mediated myoediting can successfully correct the *Dmd* mutation by HDR or NHEJ and restore dystrophin expression (Long et al., 2014). However, germline editing in humans is currently not feasible, necessitating alternative strategies for therapeutic genome editing. Therefore, we and other groups used recombinant adeno-associated virus (rAAV) to deliver the CRISPR/Cas9 genome editing components to postnatal *mdx* mice for skipping or deleting the mutated exon *in vivo* (Bengtsson et al., 2017; El Refaey et al., 2017; Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016). The rAAV-delivered CRISPR/Cas9-mediated postnatal genome editing successfully restored dystrophin expression and improved muscle function in *mdx* mice. These studies underscore the therapeutic potential of the CRISPR/Cas9 system for treating devastating muscle diseases.

The CRISPR/Cpf1 system was also used to correct *DMD* mutations in human induced pluripotent stem cells (iPSCs) and in *mdx* mice either by exon skipping or HDR (Zhang et al., 2017b), which further expands the range of CRISPR/Cas-mediated genome editing in AT-rich loci. Due to post-mitotic and multinucleation features, skeletal muscle is ideal for therapeutic CRISPR/Cas9 genome editing because genomic correction of a sub-population of nuclei leads

to steady improvement of muscle function (Bengtsson et al., 2017; El Refaey et al., 2017; Long et al., 2016; Long et al., 2014; Nelson et al., 2016; Tabebordbar et al., 2016; Zhang et al., 2017b). Therefore, CRISPR/Cas-mediated myoediting represents a novel method for DMD treatment. In the following sections, different strategies of applying the CRISPR/Cas system for correcting *DMD* mutations are discussed in detail.

Strategies of CRISPR/Cas-mediated *DMD* Correction

Initial efforts to apply programmable nucleases such as meganuclease, ZFN and TALEN for precise genome editing provided many insights into the permanent correction of *DMD* mutations (Li et al., 2015; Maggio et al., 2016b; Ousterout et al., 2015b; Ousterout et al., 2013; Popplewell et al., 2013). The CRISPR/Cas system significantly simplified the genome editing process. To date, four strategies have been developed for CRISPR/Cas-mediated correction of *DMD* mutations, which are exon deletion, exon skipping, exon reframing and exon knock-in.

Exon deletion

Approximately 80% of mutations found in the *DMD* gene are out-of-frame exon deletions or duplications, leading to reading frame incompatibility between adjacent exons (Figure 1.8, panel A) (Aartsma-Rus et al., 2006; Bladen et al., 2015). The most traditional strategy to permanently restore the *DMD* open reading frame (ORF) is in-frame exon deletion, in which a pair of sgRNAs is used to generate two simultaneous DNA DSBs within the intron regions flanking the out-of-frame exon, leading to complete removal of a single exon or

multiple exons to generate a compatible reading frame outcome with the adjacent exon (Figure 1.8, panel B). CRISPR/Cas-mediated exon deletion is best suited for correcting *DMD* mutations caused by exon duplication and has been reported with high efficiency in human DMD myoblasts with exon 2 or exon 18-30 duplications (Lattanzi et al., 2017; Wojtal et al., 2016). Removal of the duplicated exons, restores the *DMD* ORF and produces full-length dystrophin protein that is indistinguishable from wild type or normal dystrophin, although small INDELs can be observed at the genomic level.

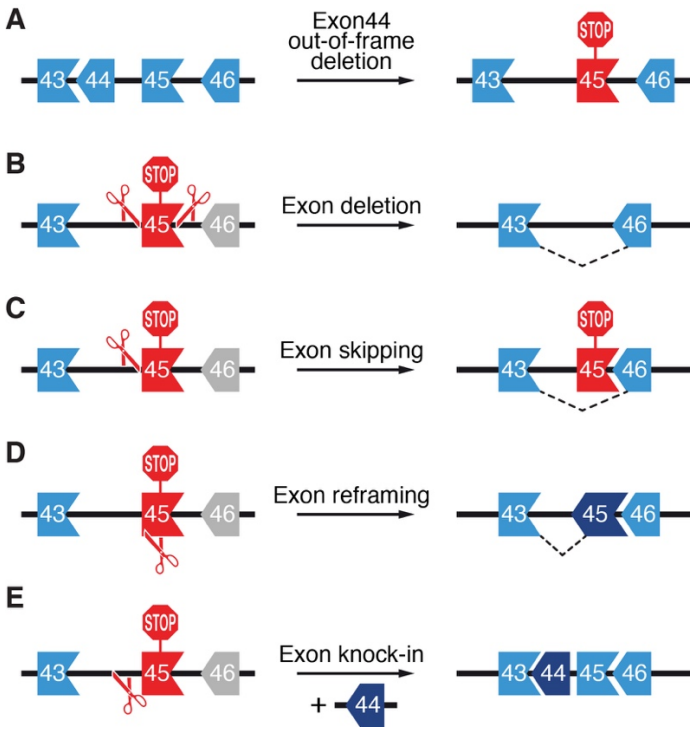

Figure 1.8 Strategies for CRISPR/Cas-mediated correction of *DMD* mutations. (**A**) A schematic illustration showing arrangement of exons 43-46 of the *DMD* gene in terms of their reading frame compatibility. This genomic region is used here as an example to highlight the strategies for CRISPR/Cas9 correction of *DMD* mutations. An out-of-frame deletion of *DMD* exon 44 results in splicing of exon 43 to exon 45. This creates a premature stop codon in exon 45 (red STOP sign). (**B**) Exon deletion is used to restore the *DMD* reading frame. Two sgRNAs targeting introns 44 and 45 will generate two DNA DSBs flanking exon 45. This leads to excision

of exon 45 and subsequent splicing of exon 43 to exon 46. (**C**) Exon skipping is mediated by a single sgRNA which targets the splice acceptor site of exon 45. The INDELs generated by NHEJ-mediated repair disrupt the splice acceptor site of exon 45, leading to splicing of exon 43 to exon 46. (**D**) Exon reframing is mediated by a single sgRNA targeting exon 45. The INDELs in exon 45 generated by NHEJ-mediated repair may restore the reading frame compatibility of exon 45 with exons 43 and 46. (**E**) Exon knock-in relies on HDR DNA repair pathway in the presence of a donor template. A single sgRNA targeting intron 44 will generate a DNA DSB and repaired by HDR when exon 44 is used as a donor template, leading to exon 44 knock-in between exons 43 and 45.

DMD mutations caused by an out-of-frame exon deletion can be corrected by an in-frame exon deletion, producing a truncated dystrophin protein with internal deletions. For example, cultured myoblasts from DMD patients with an out-of-frame deletion of exons 48-50 have an incompatible reading frame when exon 47 is spliced with exon 51. A pair of sgRNAs targeting intron 50 and 51 was used to delete exon 51, restoring the reading frame between exon 47 and 52 (Ousterout et al., 2015a). Similarly, the dystrophin reading frame incompatibility caused by an out-of-frame deletion of exons 45-52 has been corrected by exon 53 deletion, leading to splicing of exon 44 to exon 54 and subsequently restoring the dystrophin reading frame (Maggio et al., 2016a; Maggio et al., 2016b). Multiple exons can also be deleted to restore the dystrophin reading frame. By using a pair of sgRNAs targeting intron 44 and 55, a large deletion extending from exon 45 to 55 was generated, leading to reading frame restoration of exon 44 to exon 56 (Ousterout et al., 2015a; Young et al., 2016). Similarly, a large deletion extending from exons 44 to 54 was generated, by using a pair of sgRNAs targeting intron 43 and 54, leading to reading frame restoration between exon 43 and 55 (Maggio et al., 2016a; Maggio et al., 2016b). Several *in vivo* studies in postnatal *mdx* or *mdx*^{4cv} mice used exon deletion strategies to remove a single or multiple exons with a point mutation

and thereby restored dystrophin expression and muscle function (Bengtsson et al., 2017; El Refaey et al., 2017; Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016; Xu et al., 2016).

Exon deletion is a promising strategy to correct mutations clustered in the second hot spot region (exons 45-55) because the spectrin-like repeats within the central rod domain are tolerant of large in-frame deletions (Gao and McNally, 2015; Guiraud et al., 2015). However, special consideration should be given to mutations at the N- and C-termini of dystrophin because these regions encode many essential domains known to interact with the actin cytoskeleton and dystrophin glycoprotein complex. For example, three different exon deletion strategies were applied to correct the *DMD* mutation caused by an out-of-frame deletion of exons 8-9 and different outcomes were observed in regard to dystrophin protein stability and function (Kyrychenko et al., 2017). Specifically, an in-frame deletion of exons 7-11 retained all three actin binding sites but this truncated dystrophin was structurally unstable and showed minimal recovery of cardiomyocyte function *in vitro*. In contrast, in-frame deletion of exons 3-9 only retained actin binding site 1 but was the most effective strategy to restore functionality of human iPSC-derived cardiomyocytes. Reading frame restoration does not guarantee functional recovery and hence additional empirical analysis should be performed to further evaluate different correction strategies.

Exon skipping

Exon skipping has been achieved using anti-sense oligonucleotide (AON)-based therapy (Aartsma-Rus, 2012). However, AON-based exon skipping corrects at the mRNA level, while retaining the mutant *DMD* in the genome. Thus, this approach requires life-long treatment. In contrast, CRISPR/Cas-based exon skipping is achieved by NHEJ-mediated disruption of the splice acceptor or donor sequence at the genomic level, leading to permanent exon skipping and completely eliminating the source of the mutation. For example, human iPSCs derived from DMD patients with exon 44 deletion have an incompatible reading frame between exon 43 and 45 (Figure 1.8, panel A). A single sgRNA was designed to specifically target the intron 44 and exon 45 boundary, thereby inducing a DNA DSB at the splice acceptor site of exon 45 (Figure 1.8, panel C). The INDELs generated by NHEJ-based DSB repair disrupted the splice acceptor sequence of exon 45, leading to exon 45 skipping during mRNA splicing (Li et al., 2015). Similarly, *DMD* mutations caused by an out-of-frame deletion of exons 48-50 or exons 45-52 have been corrected by skipping exon 51 or exon 53, respectively (Maggio et al., 2016a; Maggio et al., 2016b; Zhang et al., 2017b). Recently, exon skipping has also been used to correct *Dmd* in a mouse model representing the most commonly deleted hot spot mutation in humans (Amoasii et al., 2017). This mouse model has an out-of-frame deletion of exon 50, which generates a premature stop codon in exon 51. A single sgRNA was designed to target the exon 51 splice acceptor site, leading to exon 51 skipping. Therefore, using a single sgRNA-mediated exon skipping strategy, which abolishes either the splice site acceptor or splice site donor or allows for reframing, overcomes the necessity of double sgRNA-based exon deletion.

Usually, the single sgRNA-mediated exon skipping strategy generates a relatively small INDEL at the intron/exon boundary, destroying the exon splice acceptor or donor site but retaining the residual part of the exon sequence in the genome. If "AG" nucleotides are present in the residual part of the exon, they can serve as a pseudo-splice acceptor sequence, rendering the single sgRNA-mediated exon skipping ineffective. Therefore, additional experimental studies, such as reverse transcription polymerase chain reaction (RT-PCR) or Western blot analysis should be performed to confirm exon skipping at the RNA and protein level.

Exon reframing

A NHEJ-based reframing strategy can also be applied to restore the dystrophin ORF, in which a single or a pair of sgRNAs are used to generate DNA DSBs within the exon region, leading to a targeted frameshift, since in theory, one-third of INDELs created by NHEJ should be in-frame (Figure 1.8, panel D). Several studies have applied this strategy to restore the dystrophin reading frame by inducing targeted frameshifts in exons with an incompatible reading frame in regard to the adjacent exon, including exons 23, 45, 50, 51, 53 and 54 (Bengtsson et al., 2017; Iyombe-Engembe et al., 2016; Li et al., 2015; Long et al., 2014; Maggio et al., 2016a; Maggio et al., 2016b; Ousterout et al., 2015a; Zhang et al., 2017b). Unlike the exon deletion strategy, which excises a single or multiple exons, exon reframing only creates small INDELs and hence, minimizes the length of the genomic deletion.

Both exon skipping and exon reframing strategies require using one sgRNA-mediated single cut in the genome. These two strategies are considered more efficient than using two

sgRNA-mediated double cuts in the genome. This is because exon deletion by excision using two sgRNAs requires two cooperative DNA DSBs. However, two DNA DSBs do not always occur simultaneously since there is a possibility that a single DNA DSB can be rapidly rejoined by NHEJ-mediated DNA repair, leaving the second intronic DSB ineffective. In this situation, exon deletion cannot be achieved because of the latency between the two DNA DSBs. In contrast, one sgRNA-mediated single cut near the splice acceptor site can be sufficient to restore the *DMD* ORF. For example, if the INDEL disrupts the splice acceptor sequence, this could lead to exon skipping. Alternatively, if the INDEL does not disrupt the splice acceptor sequence, there is still a possibility that one-third of the INDELs within the exon could be in-frame, leading to exon reframing.

Exon knock-in

In general, *DMD* mutations corrected by exon deletion, skipping or reframing strategies will generate truncated dystrophin proteins with internal deletions. In principle, *DMD* mutations can also be corrected by exon knock-in, leading to expression of full-length dystrophin protein (Figure 1.8, panel E). Exon knock-in requires a DNA donor template and active cell cycle to induce HDR-mediated precise editing in the S and G2 phases. This repair strategy has been used in DMD patient-derived iPSCs to correct a mutation caused by an out-of-frame deletion of exon 45 (Li et al., 2015). In addition, point mutations in mouse *Dmd* exon 23 and 53 have also been corrected by HDR-mediated precise editing (Bengtsson et al., 2017; Long et al., 2014; Zhang et al., 2017b; Zhu et al., 2017).

Mutations at specific regions in the N- and C-termini of dystrophin generally are not feasible for exon deletion or skipping-based correction because essential domains known to interact with cytoskeletal actin or the sarcoglycan complex are encoded within these regions. Therefore, exon knock-in is required to correct these types of mutations. However, due to the post-mitotic nature of mature striated muscle and cardiomyocytes, HDR efficiency remains low in CRISPR/Cas-mediated postnatal genome editing (Bengtsson et al., 2017). Recently, precise genome editing in post-mitotic cells and animals with high efficiency was reported (Suzuki et al., 2016). This technology, which was termed homology-independent targeted integration (HITI), only relies on the NHEJ pathway and can be used to precisely integrate DNA fragments into the mammalian genome, regardless of the cell cycle state, which may provide opportunities to correct certain *DMD* mutations by exon knock-in.

IN VIVO GENOME EDITING OF ANIMAL MODELS.

The ultimate goal of therapeutic application of genome editing is to permanently correct mutations that contribute to human monogenic diseases. Programmable nucleases such as the CRISPR/Cas-system have been demonstrated to be effective in precise correction of pathogenic mutations found in the human embryo (Kang et al., 2016b; Liang et al., 2015; Ma et al., 2017a; Tang et al., 2017). However, ethical issues, as well as public policies, restrict therapeutic application of human germline editing, leaving postnatal genome editing as the means to achieve the same goal.

Therapeutic genome editing requires delivering programmable nucleases and other genome editing components to target cells, which can be achieved ex vivo or in vivo. Ex vivo

editing requires in vitro editing of the cellular genome, followed by transplantation of the targeted cell population to the original host. CRISPR/Cas-mediated ex vivo editing has been shown to be effective in the hematopoietic system (DeWitt et al., 2016; Genovese et al., 2014; Gundry et al., 2016; Sather et al., 2015; Wang et al., 2015). In theory, ex vivo genome editing can be applied to the skeletal muscle system because satellite cells serve as adult stem cells of skeletal muscle and are capable of surviving manipulation in vitro. However, skeletal muscle is the largest tissue in the human body, comprising approximately 40% of body weight. Transplantation of edited satellite cells or progenitor cells may improve local muscle function but systemic functional recovery remains questionable (Zhu et al., 2017). In addition to skeletal muscle, the heart is also affected in a variety of muscular dystrophies, but does not have stem cells or stem cell-like cell populations capable of transplantation and regeneration (Balsam et al., 2004; Ellison et al., 2013; Murry et al., 2004; Orlic et al., 2001; Sultana et al., 2015; van Berlo et al., 2014). Taken together, ex vivo transplantation-based approaches are generally not feasible to treat muscular dystrophies, especially for those affecting both skeletal muscle and the heart. Therefore, in vivo postnatal genome editing turns out to be a more practical approach to permanently correct genetic mutations causing muscular dystrophies.

Delivery of Genome Editing Components by Non-Viral Vectors

Achieving in vivo postnatal genome editing requires an efficient and effective delivery system. Genome editing components can be physically or chemically delivered to target cells by non-viral vectors (Figure 1.9). For example, microinjection and electroporation have been demonstrated as effective methods to deliver CRISPR/Cas genome editing components to

target cells (Chen et al., 2016; Kimura et al., 2014; Meca-Cortes et al., 2017; Niu et al., 2014; Yang et al., 2013a). However, these physical approaches are widely used for germline editing or embryo manipulation but are not feasible for systemic delivery in the postnatal host. In contrast, hydrodynamic intravenous injection can achieve systemic delivery in multiple tissues, including liver, kidney, lung, skeletal muscle and heart (Bonamassa et al., 2011; Suda and Liu, 2007). Some studies have applied this technology to deliver CRISPR/Cas genome editing components to postnatal mice for mutating cancer genes in the mouse liver, correcting a tyrosinemia mutation and disrupting hepatitis B virus (Lin et al., 2014a; Xue et al., 2014; Yin et al., 2014b; Zhen et al., 2015). However, hydrodynamic injection requires a large injection volume and high pressure, which may damage tissues or organs.

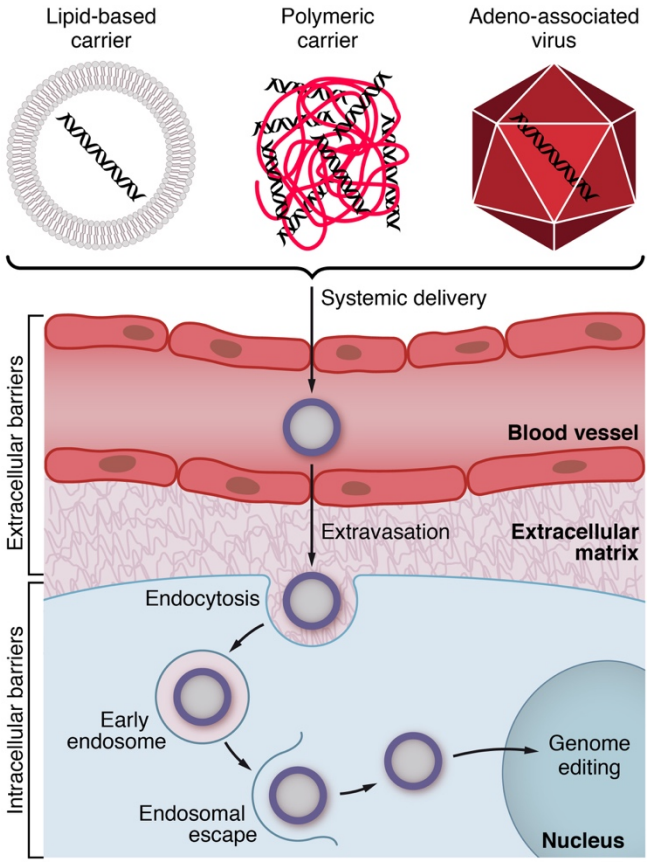

Figure 1.9 Systemic delivery of genome editing components. Non-viral (lipid-based carriers or polymeric carriers) and viral (rAAV) vector-based delivery of genome editing components into target tissue. Following administration, the delivery vectors pass through the blood vessel by extravasation to reach their target tissues. In the tissue, they undergo cytoplasmic trafficking, endosomal escape and nuclear entry to perform genome editing in the nucleus.

In addition to physical vectors, chemical vectors, such as polymeric carriers and lipid-based carriers, are also widely used for in vivo delivery. Polymeric carriers are cationic polymers that can condense negatively charged DNA or RNA, whereas lipid-based carriers can spontaneously assemble into liposomes consisting of nucleic acids and cationic or neutral lipids (Yin et al., 2014a). The CRISPR/Cas system and other genome editing components can be chemically delivered into animals in the form of plasmid DNA, mRNA or ribonucleoprotein complexes. Chemical vectors can protect nucleic acids or ribonucleoproteins from degradation by endonucleases or proteases in physiological fluids and extracellular space, improving stability and half-life. For example, chemically modified CRISPR/Cpf1 mRNA and crRNA have been reported to enhance genome editing efficiency (Li et al., 2017). However, many challenges still remain, including efficient delivery to the tissue of interest, cellular internalization and protection from the lysosomal degradation pathway (Wang et al., 2017; Yin et al., 2014a).

It has been reported that plasmid DNA can be retained in skeletal muscle for over a year after intramuscular injection, suggesting the likelihood of long term gene expression and retention in post-mitotic tissue (Wolff et al., 1992). In contrast, localized non-viral delivery such as intramuscular injection cannot generate a systemic effect, which is not ideal for treating muscular dystrophies since skeletal muscle is one of the largest tissues in the human body and multiple muscle groups can be affected. Systemic non-viral delivery, such as intravenous or intraperitoneal injection, have extended the targeting range but may cause gene expression in non-targeted tissues. Therefore, delivery of genome editing components to animals by non-viral vectors still requires improvement and optimization before clinical translation.

Delivery of Genome Editing Components by Viral Vectors

Many viral vectors have been used for gene therapy, including lentivirus, retrovirus, herpes simplex virus, poxvirus, adenovirus, adeno-associated virus, baculovirus, and Epstein-Barr virus (Nayerossadat et al., 2012; Wang et al., 2016). Among these, adeno-associated virus (AAV) is the most promising viral vector for the delivery of genome editing components to specific tissues, such as muscle and heart. AAV is an non-enveloped DNA virus with an approximately 5 kb linear ssDNA genome (Srivastava et al., 1983). The genome of wild-type AAV has two major ORFs flanked by two inverted terminal repeats (ITRs), while in the recombinant AAV (rAAV), the viral ORFs encoding the replication and capsid proteins are replaced by the customized gene expression cassette (Flotte and Berns, 2005). Both wild-type and rAAVs are nonpathogenic in humans or animals, and their propagation requires a helper virus, making them a safe delivery system for therapeutic genome editing (Kotterman and Schaffer, 2014; Samulski and Muzyczka, 2014; Sun et al., 2003).

The rAAVs have many appealing features, including broad spectrum tissue tropism with minimal integration risk and long-term transgene expression from the episomal genome after viral transduction (Duan et al., 1998; Nonnenmacher and Weber, 2012; Schnepp et al., 2005). Currently, 13 AAV serotypes are widely available for gene delivery and each of them shows different tissue tropism. AAV serotypes 1, 6, 8 and 9 have high tropism in skeletal muscle and heart (Blankinship et al., 2004; Chao et al., 2000; Gao et al., 2002; Gregorevic et al., 2004; Inagaki et al., 2006; Pan et al., 2015; Wang et al., 2005b; Yue et al., 2011; Zincarelli et al., 2010). In addition, tissue tropism and transduction efficiency can also be improved by pseudo-typing. For example, AAV2 genomes pseudo-packaged into AAV5 capsids can

enhance gene delivery to skeletal muscle, whereas improved cardiomyocyte transduction has been observed by pseudo-packaging AAV2 genomes into AAV6 capsids (Duan et al., 2001; Sipo et al., 2007). rAAV has been successfully used as a delivery system to administer CRISPR/Cas9 and other genome editing components to postnatal *mdx* or *mdx*^{4cv} mice and correct *Dmd* mutations by exon deletion or reframing strategies (Bengtsson et al., 2017; El Refaey et al., 2017; Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016) (Figure 1.9). In addition, several other studies reported rAAV-delivered CRISPR/Cas9-mediated *in vivo* genome editing in mouse models of human Huntington disease and congenital muscular dystrophy (Kemaladewi et al., 2017; Monteys et al., 2017; Yang et al., 2017). These studies demonstrated that the combination of a rAAV-based delivery system with CRISPR/Cas9-mediated postnatal genome editing is a compelling strategy to permanently correct mutations responsible for monogenic neuromuscular disorders. However, long term benefits and effects in animal models still need to be examined to prepare for future clinical trials.

CHAPTER TWO

CRISPR/CPF1 CORRECTION OF MUSCULAR DYSTROPHY MUTATIONS IN HUMAN CARDIOMYOCYTES AND MICE

Acknowledgement

Parts of this chapter, including figures, have been reproduced, with or without modifications, from my previously published work (Zhang et al., 2017b).

Abstract

Duchenne muscular dystrophy (DMD), caused by mutations in the x-linked dystrophin gene (*DMD*), is characterized by fatal degeneration of striated muscles. Dilated cardiomyopathy is one of the most common lethal features of the disease. We deployed Cpf1, a unique class 2 CRISPR effector, to correct *DMD* mutations in patient-derived induced pluripotent stem cells (iPSCs) and *mdx* mice, an animal model of DMD. Cpf1-mediated genomic editing of human iPSCs, either by skipping of an out-of-frame *DMD* exon or correcting a nonsense mutation, fully restored dystrophin expression after differentiation to cardiomyocytes and enhanced contractile function. Similarly, pathophysiological hallmarks of muscular dystrophy were corrected in *mdx* mice following Cpf1-medated germline editing. These findings are the first to show the efficiency of Cpf1-mediated correction of genetic mutations in human cells and in an animal disease model, and represent a significant step toward therapeutic translation of gene editing for correction of DMD.

Introduction

Duchenne muscular dystrophy (DMD) is an X-linked recessive disease caused by mutations in the gene coding for dystrophin, which is a large cytoskeletal protein essential for integrity of muscle cell membranes (Hoffman et al., 1987; Koenig et al., 1987). DMD causes progressive muscle weakness, culminating in premature death by the age of 30, generally from cardiomyopathy. There is no effective treatment for this disease. Numerous approaches to rescue dystrophin expression in DMD have been attempted, including delivery of truncated dystrophin or utrophin by recombinant adeno-associated virus (rAAV) (Bostick et al., 2011; Hollinger and Chamberlain, 2015) and skipping of mutant exons with anti-sense oligonucleotides and small molecules (Shimizu-Motohashi et al., 2016). However, these approaches cannot correct *DMD* mutations or permanently restore dystrophin expression.

The CRISPR (clustered regularly interspaced short palindromic repeats) system functions as an adaptive immune system in bacteria and archaea that defends against phage infection (Mojica et al., 2005). In this system, an endonuclease is guided to specific genomic sequences by a single guide RNA (sgRNA), resulting in DNA cutting near a proto-spacer adjacent motif (PAM) sequence. The CRISPR-Cas (CRISPR-associated proteins) system represents a promising approach for correction of diverse genetic defects. However, many challenges remain to be addressed. For example, *Streptococcus pyogenes* Cas9 (*Sp*Cas9), currently the most widely used Cas9 endonuclease, has a G-rich PAM requirement (NGG) that excludes genome editing of AT-rich regions (Ran et al., 2015). Additionally, the large size of *Sp*Cas9 reduces the efficiency of packaging and delivery in low-capacity viral vectors, such as adeno-associated virus (AAV) vectors. The Cas9 endonuclease from *Staphylococcus aureus*

(*Sa*Cas9), although smaller in size than *Sp*Cas9, has a PAM sequence (NNGRRT) that is longer and more complex, thus limiting the range of its genomic targets (Ran et al., 2015). Smaller CRISPR enzymes with greater flexibility in recognition sequence and comparable cutting efficiency would facilitate precision gene editing, especially for translational applications.

Recently, a new RNA-guided endonuclease, named Cpf1 (CRISPR from *Prevotella* and *Francisella 1*), was shown to be effective in mammalian genome cleavage (Zetsche et al., 2015). Cpf1 has several unique features that expand its genome editing potential: (1) Cpf1-mediated cleavage is guided by a single and short crRNA (abbreviated as gRNA), whereas Cas9-mediated cleavage is guided by a hybrid of CRISPR RNA (crRNA) and a long trans-activating crRNA (tracrRNA) (Jinek et al., 2012). (2) Cpf1 prefers a T-rich PAM at the 5'-end of a protospacer, while Cas9 requires a G-rich PAM at the 3' end of the target sequence. (3) Cpf1-mediated cleavage produces a sticky end distal to the PAM site, which activates DNA repair machinery, while Cas9 cutting generates a blunt end. (4) Cpf1 also has RNase activity, which can process precursor crRNAs to mature crRNAs (Fonfara et al., 2016; Zetsche et al., 2015). Like Cas9, Cpf1 binds to a targeted genomic site and generates a double-stranded break (DSB), which is then repaired either by non-homologous end-joining (NHEJ) or by homology-directed repair (HDR) if an exogenous template is provided. Although Cpf1 has been shown to be active in mammalian genome editing (Kim et al., 2016; Zetsche et al., 2015), its potential usefulness for correction of genetic mutations in mammalian cells and animal models of disease has yet to be demonstrated.

Previously, we and others used CRISPR-Cas9 to correct the DMD mutation in mice and human cells (Amoasii et al., 2018; Amoasii et al., 2017; Bengtsson et al., 2017; Hakim et

al., 2018; Long et al., 2016; Min et al., 2019; Nelson et al., 2016; Nelson et al., 2019; Tabebordbar et al., 2016). In this chapter I show that Cpf1 provides a robust and efficient RNA-guided genome editing system that can be used to permanently correct *DMD* mutations by different strategies, thereby restoring dystrophin expression and preventing progression of the disease. These findings provide a new approach for the permanent correction of human genetic mutations.

Results

Strategies for CRISPR/Cpf1-mediated genome editing of DMD exon 51

Exon deletions preceding exon 51 of the human *DMD* gene, which disrupt the open reading frame (ORF) by juxtaposing out of frame exons, represent the most common type of human DMD mutation (Aartsma-Rus et al., 2009). Skipping of exon 51 can, in principle, restore the DMD ORF in 13% of DMD patients with exon deletions (Cirak et al., 2011). To test the potential of Cpf1 to correct this type of "hot-spot" mutation, we used DMD patient fibroblast-derived iPSCs (Riken HPS0164, abbreviated as Riken51), which harbor a deletion of exons 48 to 50, introducing a premature termination codon within exon 51 (Figure 2.1, panel A).

The splice acceptor region is generally pyrimidine rich (Padgett, 2012), which creates an ideal PAM sequence for genome editing by Cpf1 endonuclease (Figure 2.1, panel B). To rescue dystrophin expression in Riken51 iPSCs, we used a Cpf1 gRNA to target exon 51, introducing small insertions and deletions (INDELs) in exon 51 by NHEJ and subsequently reframing the dystrophin ORF, theoretically, in one-third of corrected genes, a process we refer

to as “reframing” (Figure 2.1, panel A). We also compared two Cpf1 orthologues, *Lb*Cpf1 (from *Lachnospiraceae bacterium ND2006*) and *As*Cpf1 (from *Acidaminococcus sp. BV3L6*), which use the same PAM sequences for genome cleavage.

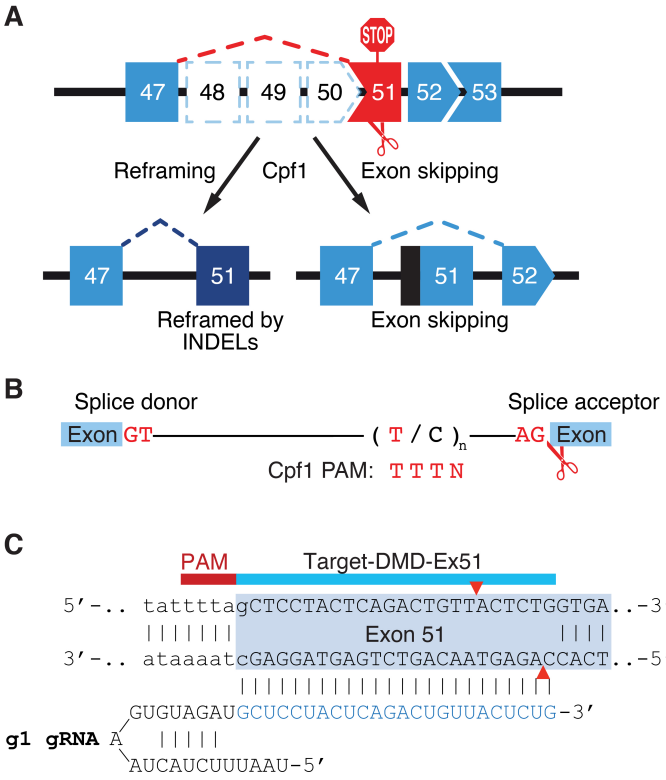

Figure 2.1 Strategy for CRISPR/Cpf1-meidated correction of *DMD* exons 48-50 out-of-frame deletion mutation. (**A**) A *DMD* deletion of exons 48-50 results in splicing of exon 47 to 51, generating an out-of-frame mutation of dystrophin. Two strategies were used for the restoration of dystrophin expression by Cpf1. In the “reframing” strategy, small INDELs in exon 51 restore the protein reading frame of dystrophin. The “exon skipping” strategy is achieved by disruption of the splice acceptor of exon 51, which results in splicing of exon 47 to 52 and restoration of the protein reading frame. (**B**) The 3’ end of an intron is T-rich, which generates Cpf1 PAM sequences enabling genome cleavage by Cpf1. (**C**) Illustration of Cpf1 gRNA targeting DMD exon 51. The T-rich PAM (red line) is located upstream of exon 51 near the splice acceptor site. The sequence of the Cpf1 g1 gRNA targeting exon 51 is shown, highlighting the complementary nucleotides in blue. Cpf1 cleavage produces

a staggered-end distal to the PAM site (demarcated by red arrowheads). The 5' region of exon 51 is shaded in light blue. Exon sequence is upper case. Intron sequence is lower case.

Cpf1 cleavage was targeted to the T-rich splice acceptor site of exon 51 using a guide RNA (designated g1) (Figure 2.1, panel C), which was cloned into plasmids p*Lb*Cpf1-2A-GFP and p*As*Cpf1-2A-GFP (Figure 2.2, panel A). These plasmids express human codon optimized *Lb*Cpf1 or *As*Cpf1, plus GFP; enabling fluorescence activated cell sorting (FACS) of Cpf1-expressing cells (Figure 2.2, panel A). Initially, we evaluated the cleavage efficiency of Cpf1-editing with g1 in human 293T cells. Both *Lb*Cpf1 and *As*Cpf1 efficiently induced DNA cleavage with g1, as detected using a T7E1 assay that recognizes and cleaves non-perfectly matched DNA (Figure 2.2, panel B).

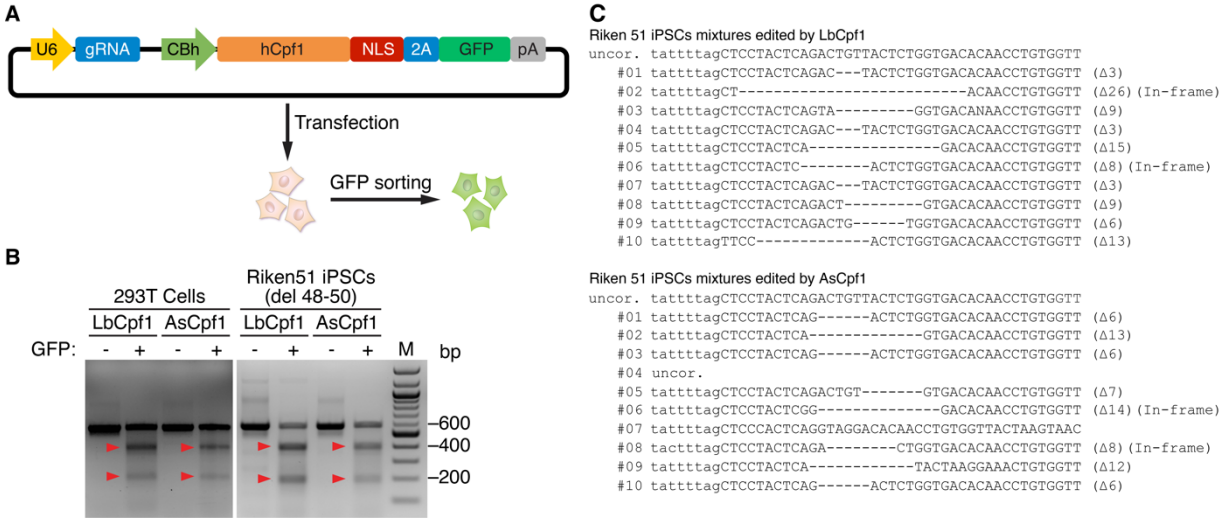

Figure 2.2 in vitro genome editing in human 293T cells and iPSCs by CRISPR/*Lb*Cpf1 or CRISPR/*As*Cpf1. (**A**) Illustration of a plasmid encoding human codon-optimized Cpf1 (hCpf1) with a nuclear localization signal (NLS) and 2A-GFP. The plasmid also encodes a Cpf1 gRNA driven by the U6 promoter. Cells transfected with this plasmid express GFP, allowing for selection of Cpf1-expressing cells by FACS. (**B**) T7E1 assays using human 293T cells or DMD iPSCs (RIKEN51) transfected with plasmid expressing *Lb*Cpf1 or *As*Cpf1, gRNA and GFP

show genome cleavage at *DMD* exon 51. Red arrowheads point to cleavage products. M, marker. (**C**) DNA sequencing of *DMD* exon 51 from a mixture of RIKEN 51 iPSCs edited by *Lb*Cpf1 or *As*Cpf1 using g1 gRNA.

Next, we used *Lb*Cpf1 and *As*Cpf1 with g1 to edit Riken51 iPSCs, and by the T7E1 assay we observed genome cleavage at *DMD* exon 51. Genomic PCR products from the Cpf1-edited *DMD* exon 51 were cloned and sequenced (Figure 2.2, panel C). We observed INDELs near the exon 51 splice acceptor site in both *Lb*Cpf1- and *As*Cpf1-edited Riken51 iPSCs. Out of 20 clones, we observed four clones with reframed *DMD* exon 51, which restored the ORF (Figure 2.2, panel C).

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes after CRISPR/Cpf1-mediated reframing

Riken51 iPSCs edited by CRISPR-Cpf1 using the reframing strategy were induced to differentiate into cardiomyocytes (Figure 2.3, panel A). Cardiomyocytes with the reframed *DMD* gene were identified by RT-PCR using a forward primer targeting exon 47 and a reverse primer targeting exon 52 and PCR products were sequenced (Figure 2.3, panel B and C). Uncorrected iPSC-derived cardiomyocytes have a premature termination codon following the first 8 amino acids encoded by exon 51, which creates a premature stop codon (Figure 2.3, panel C). Cardiomyocytes differentiated from Cpf1-edited Riken51 iPSCs showed restoration of the *DMD* ORF as seen by sequencing of the RT-PCR products from amplification of exons 47 to 52 (Figure 2.3, panel C). We also confirmed restoration of dystrophin protein expression by Western blot analysis and immunocytochemistry using dystrophin antibody (Figure 2.3, panel D and E). Surprisingly, even without clonal selection and expansion, cardiomyocytes

differentiated from Cpf1-edited iPSC mixtures showed levels of dystrophin protein comparable to WT cardiomyocytes (Figure 2.3, panel D).

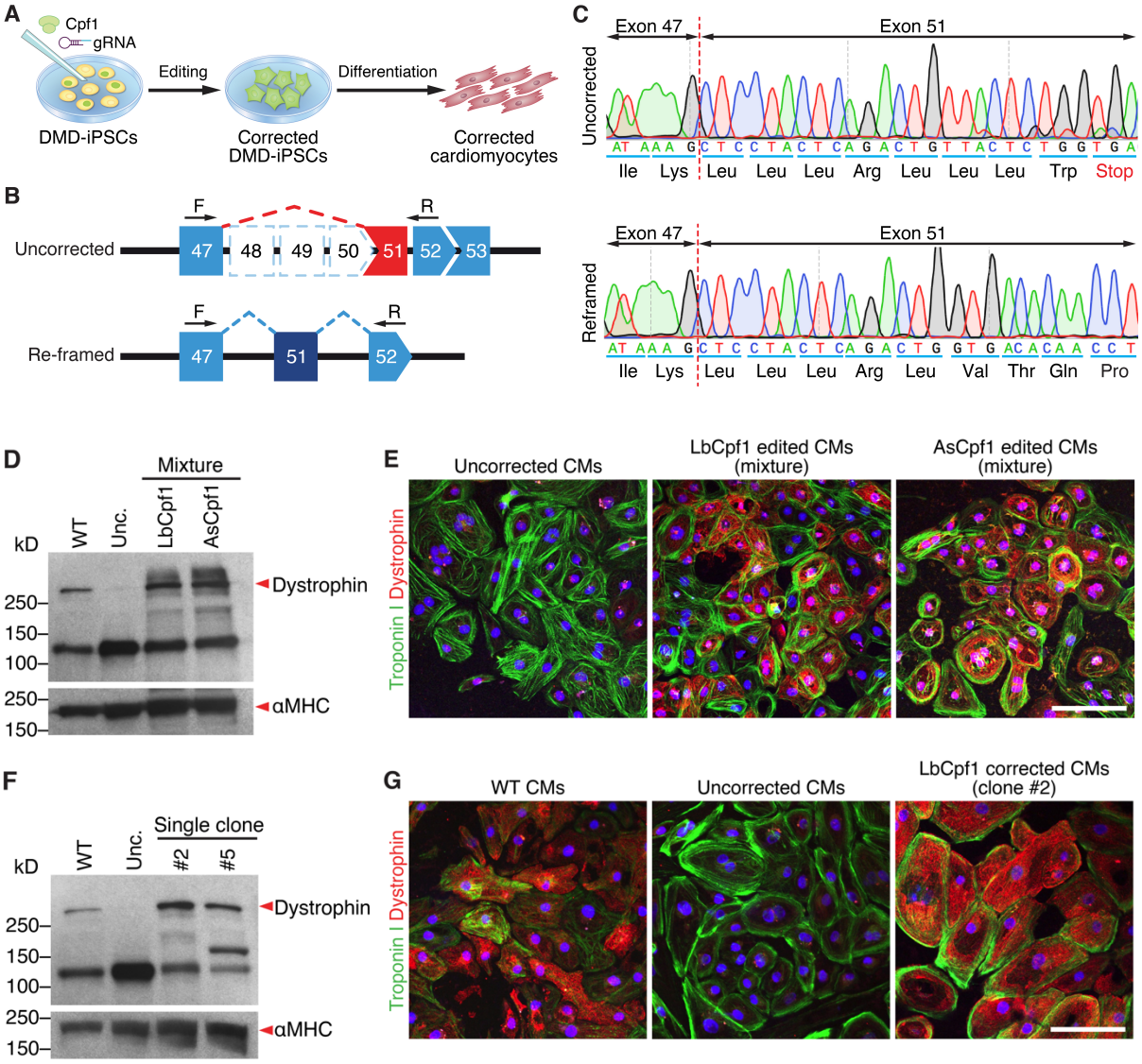

Figure 2.3 DMD iPSC-derived cardiomyocytes express dystrophin after Cpf1-mediated genome editing by reframing. (**A**) DMD skin fibroblast-derived iPSCs were edited by Cpf1 using gRNA (corrected DMD-iPSCs) and then differentiated into cardiomyocytes (corrected cardiomyocytes) for analysis of genetic correction of the DMD mutation. (**B**) A *DMD* deletion of exons 48-50 results in splicing of exon 47 to 51, generating an out-of-

frame mutation of dystrophin. Forward primer (F) targeting exon 47 and reverse primer (R) targeting exon 52 were used in RT-PCR to confirm the reframing strategy by Cpf1-meditated genome editing in cardiomyocytes. Uncorrected cardiomyocytes lack exons 48-50. In contrast, after reframing, exon 51 is placed back in-frame with exon 47. (**C**) Sequencing of representative RT-PCR products shows that uncorrected DMD iPSC-derived cardiomyocytes have a premature stop codon in exon 51, which creates a nonsense mutation. After Cpf1-mediated reframing, the ORF of dystrophin is restored. Dashed red line denotes exon boundary. (**D**) Western blot analysis shows dystrophin expression in a mixture of DMD iPSC-derived cardiomyocytes edited by reframing with *Lb*Cpf1 or *As*Lpf1 and g1 gRNA. Even without clonal selection, Cpf1-mediated reframing is efficient and sufficient to restore dystrophin expression in the cardiomyocyte mixture. αMHC is loading control. (**E**) Immunocytochemistry shows dystrophin expression in iPSC-derived cardiomyocyte (CM) mixtures following *Lb*Cpf1- or *As*Cpf1-mediated reframing. Dystrophin staining (red); Troponin I staining (green). Scale bar = 100 microns. (**F**) Western blot analysis shows dystrophin expression in single clones (#2 and #5) of iPSC-derived cardiomyocytes following clonal selection after *Lb*Cpf1-mediated reframing. αMHC is loading control. (**G**) Immunocytochemistry showing dystrophin expression in clone #2 *Lb*Cpf1-edited iPSC-derived cardiomyocytes. Scale bar = 100 microns.

From mixtures of LbCpf1-edited Riken51 iPSCs, we picked two clones (clone #2 and #5) with in-frame INDELs of different sizes and differentiated the clones into cardiomyocytes. Clone #2 had an 8 bp deletion at the 5'-end of exon 51, together with an endogenous deletion of exons 48-50. The total 405 bp deletion restored the *DMD* ORF and allowed for the production of a truncated dystrophin protein with a 135 amino acid deletion. Clone #5 had a 17 bp deletion in exon 51 and produced dystrophin protein with a 138 amino acid deletion. Although there is high efficiency of cleavage by Cpf1, the amount of DNA inserted or deleted at the cleavage site varies. Additionally, INDELs can generate extra codons at the edited locus, causing changes of the ORF. The dystrophin protein expressed by clone #2 cardiomyocytes generated four additional amino acids (Leu-Leu-Leu-Arg) between exon 47 and exon 51,

whereas dystrophin protein expressed by clone #5 cardiomyocytes generated only one additional amino acid (Leu). From both clones #2 and #5, we observed restored dystrophin protein by Western blot analysis and immunocytochemistry (Figure 2.3, panel F and G). Due to the large size of dystrophin, the internally-deleted forms migrated similarly to WT dystrophin on SDS-PAGE (Figure 2.3, panel F)

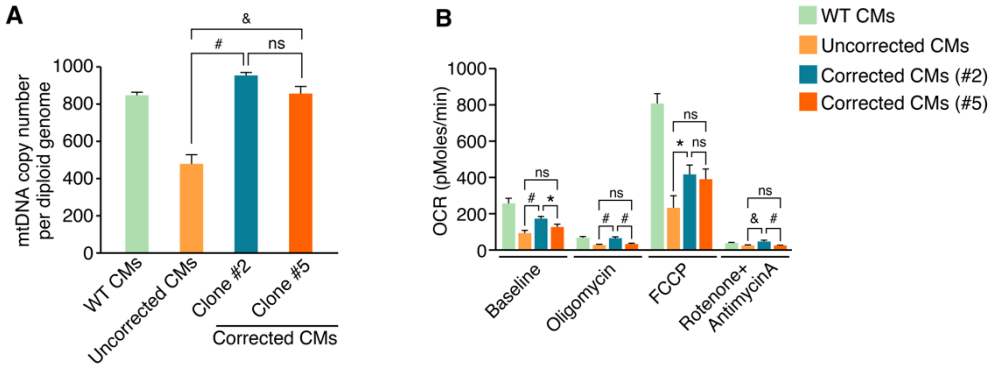

Figure 2.4 DMD iPSC-derived cardiomyocytes restore mitochondria number and improve oxygen consumption rate after Cpf1-mediated genome editing by reframing. (**A**) Quantification of mitochondrial DNA copy number in single clones (#2 and #5) of LbCpf1-edited iPSC-derived cardiomyocytes. Data are represented as mean ± SEM (n = 3). (&) $P < 0.01$; (#) $P < 0.005$; (ns) not significant. (**B**) Basal oxygen consumption rate (OCR) of single clones (#2 and #5) of LbCpf1-edited iPSC-derived cardiomyocytes, and OCR in response to oligomycin, FCCP, and Rotenone and Antimycin A, normalized to cell number. Data are represented as mean ± SEM (n = 5). (*) $P < 0.05$; (&) $P < 0.01$; (#) $P < 0.005$; (ns) not significant.

We also performed functional analysis of DMD iPSC-derived cardiomyocytes by measuring mitochondrial DNA copy number and cellular respiration rates. Uncorrected DMD iPSC-derived cardiomyocytes had significantly fewer mitochondria than the LbCpf1-corrected cardiomyocytes (Figure 2.4, panel A). After *Lb*Cpf1-mediated reframing, both corrected clones restored mitochondrial number to a level comparable to that of WT cardiomyocytes (Figure 2.4, panel A). Clone #2 iPSC-derived cardiomyocytes also showed an increase in

oxygen consumption rate (OCR) compared to uncorrected iPSC-derived cardiomyocytes at baseline (Figure 2.4, panel B). OCR was inhibited by oligomycin in all iPSC-derived cardiomyocytes, and treatment with the uncoupling agent FCCP enhanced OCR. Finally, treatment with rotenone and antimycin A further inhibited OCR in all cardiomyocytes. These results demonstrate that Cpf1-mediated *DMD* correction improved respiratory capacity of mitochondria in corrected iPSC-cardiomyocytes. Our findings show that Cpf1-mediated reframing is a highly efficient strategy to rescue DMD phenotypes in human cardiomyocytes.

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes by CRISPR/Cpf1-mediated exon skipping

In contrast to the single gRNA-mediated reframing method, which introduces small INDELs, exon skipping uses two gRNAs to disrupt splice sites and generates a large deletion (Figure 2.5, panel A). As an independent strategy to restore dystrophin expression in the Riken51 iPSCs, we designed two *Lb*Cpf1 gRNAs (g2 and g3) that target the 3'-end of intron 50 and tested the cleavage efficiency in human 293T cells. T7E1 assay showed that g2 had higher cleavage efficiency within intron 50 compared to g3 (Figure 2.5, panel B). Therefore, we co-delivered *Lb*Cpf1, g2 and g1 (g1 targets the 5' region of exon 51) into Riken51 iPSCs with the aim of disrupting the splice acceptor site of exon 51. Genomic PCR showed a lower band in LbCpf1-edited iPSCs (Figure 2.5, panel C) and sequencing data confirmed the presence of a deletion of ~200 bp between intron 50 and exon 51, which disrupted the conserved splice acceptor site (Figure 2.5, panel D).

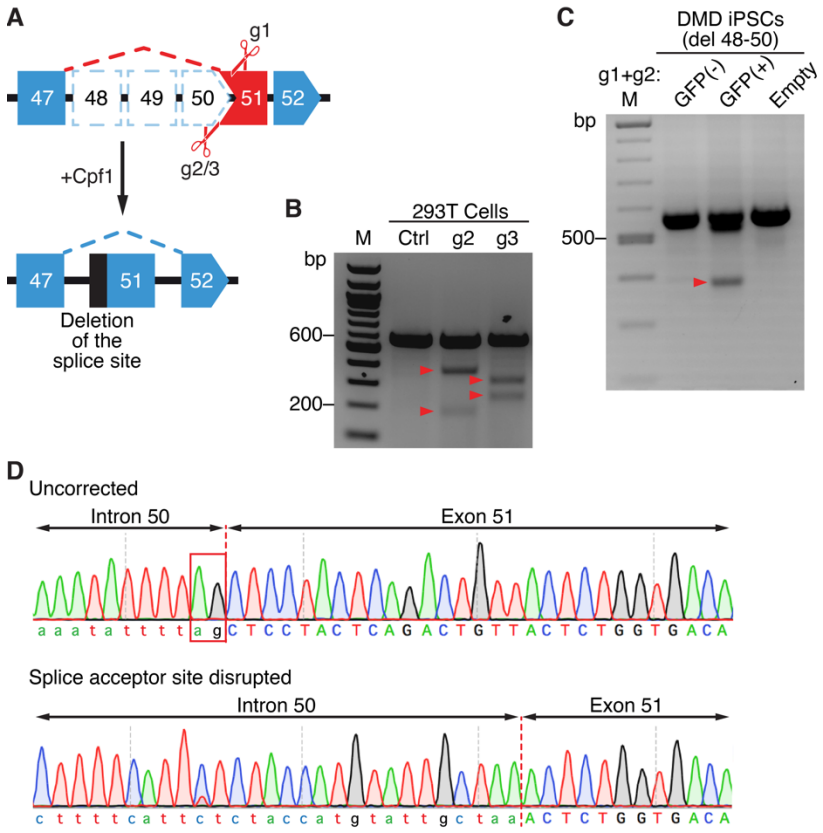

Figure 2.5 Strategy for Cpf1-mediated *DMD* exon 51 skipping in iPSC-derived cardiomyocytes. (**A**) Two gRNAs, either gRNA (g2 or g3), which target intron 50, and the other (g1), which targets exon 51, were used to direct Cpf1-mediated removal of the exon 51 splice acceptor site. (**B**) T7E1 assay using 293T cells transfected with *Lb*Cpf1 and gRNA2 (g2) or gRNA3 (g3) shows cleavage of the DMD locus at intron 50. Red arrowheads denote cleavage products. M, marker. (**C**) PCR products of genomic DNA isolated from DMD-iPSCs transfected with a plasmid expressing LbCpf1, g1 + g2 and GFP. The lower band (red arrowhead) indicates removal of the exon 51 splice acceptor site. (**D**) Sequence of the lower PCR band from panel c shows a 200-bp deletion, spanning from the 3'-end of intron 50 to the 5'-end of exon 51. This confirms removal of the "ag" splice acceptor of exon 51. The sequence of the uncorrected allele is shown above that of the LbCpf1-edited allele.

Riken51 iPSCs edited by the exon skipping strategy with g1 and g2 were differentiated into cardiomyocytes. Cells harboring the edited *DMD* allele were identified by RT-PCR using a forward primer targeting exon 47 and a reverse primer targeting exon 52; showing deletion

of the exon 51 splice acceptor site which allows skipping of exon 51 (Figure 2.6, panel A). Sequencing of the RT-PCR products confirmed that exon 47 was spliced to exon 52, which restored the *DMD* ORF (Figure 2.6, panel B). Western blot analysis and immunocytochemistry confirmed the restoration of dystrophin protein expression in a mixture of *Lb*Cpf1-edited cardiomyocytes with g1 and g2 (Figure 2.6, panel C and D). Thus, Cpf1-editing by the exon skipping strategy is highly efficient in rescuing the DMD phenotype in human cardiomyocytes.

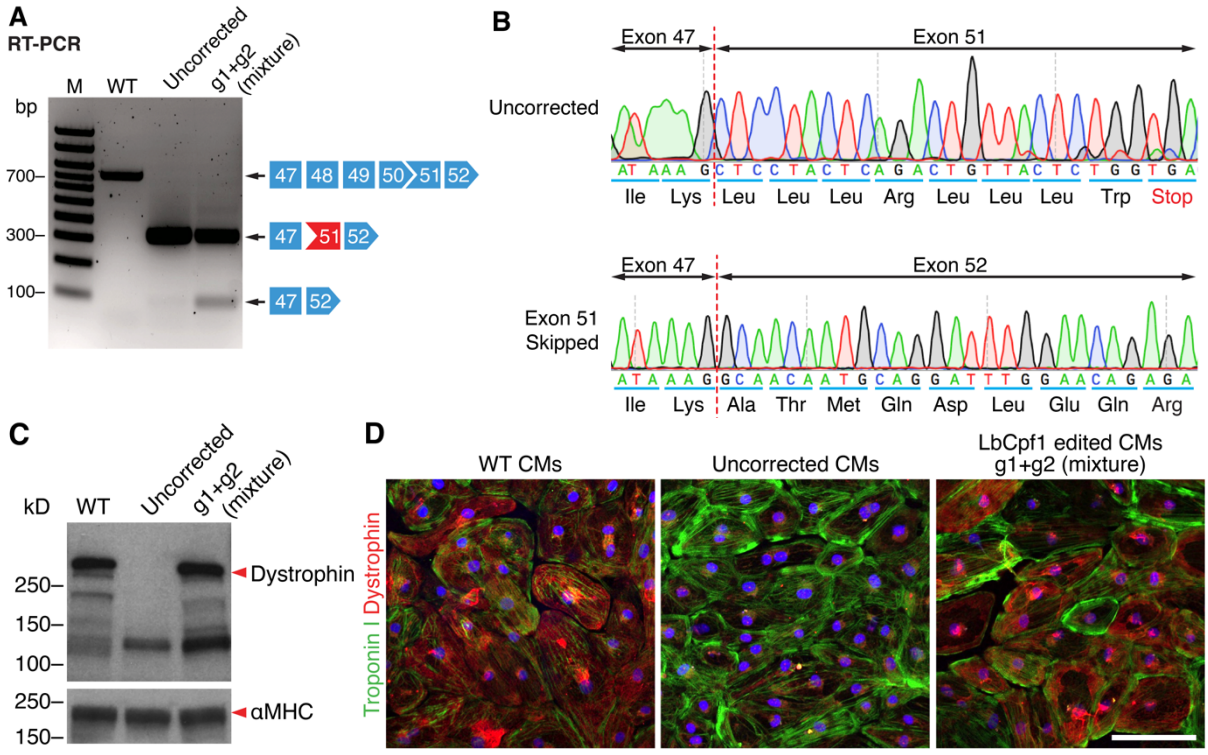

Figure 2.6 DMD iPSC-derived cardiomyocytes express dystrophin after Cpf1-mediated exon skipping. (**A**) RT-PCR of iPSC-derived cardiomyocytes using primer sets described in Fig. 2B. The 700-bp band in the WT lane is the dystrophin transcript from exon 47-52; the 300-bp band in the uncorrected lane is the dystrophin transcript from exon 47-52 with exon 48-50 deletion; and the lower band in the g1+g2 mixture lane (edited by LbCpf1) shows exon 51 skipping. (**B**) Sequence of the lower band from panel e (g1+g2 mixture lane) confirms skipping of exon 51, which reframed the DMD ORF. (**C**) Western blot analysis shows dystrophin protein

expression in iPSC-derived cardiomyocyte mixtures after exon 51 skipping by *Lb*Cpf1 with g1 + g2. αMHC is loading control. (**D**) Immunocytochemistry shows dystrophin expression in iPSC-derived cardiomyocyte mixtures (CMs) following Cpf1-mediated exon skipping with g1 + g2 gRNA compared to WT and uncorrected CMs. Dystrophin staining (red). Troponin I staining (green). Scale bar = 100 microns.

Restoration of dystrophin in mdx mice by Cpf1-mediated correction

To further evaluate the potential of Cpf1-mediated *Dmd* correction *in vivo*, we used *Lb*Cpf1 to permanently correct the mutation in germline of *mdx* mice by HDR-mediated correction or NHEJ-mediated reframing. *mdx* mice carry a nonsense mutation in exon 23 of the *Dmd* gene, due to a C to T transition (Figure 2.7, panel A). Three gRNAs (g1, g2 and g3) that target exon 23 were screened and tested in mouse 10T1/2 fibroblasts for cleavage efficiency (Figure 2.7, panel B). The T7E1 assay revealed that *Lb*Cpf1 and *As*Cpf1 had different cleavage efficiencies at *Dmd* exon 23 (Figure 2.7, panel C).

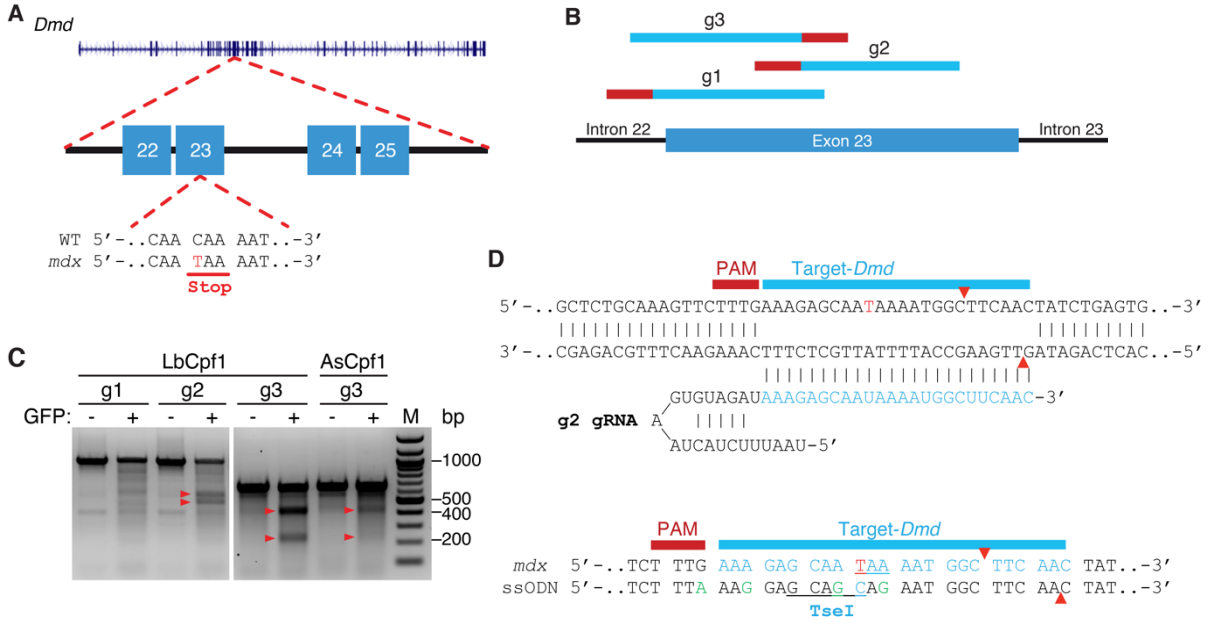

Figure 2.7 CRISPR/*Lb*Cpf1-mediated editing of *Dmd* exon 23 of the *mdx* mouse. (**A**) Illustration of mouse *Dmd* locus highlighting the mutation at exon 23. Sequence shows the nonsense mutation caused by C to T transition, which creates a premature stop codon. (**B**) Illustration showing the targeting location of gRNAs (g1, g2 and g3) (shown in light blue) on exon 23 of the *Dmd* gene. Red line represents *Lb*Cpf1 PAM. (**C**) T7E1 assay using mouse 10T1/2 cells transfected with *Lb*Cpf1 or *As*Cpf1 with different gRNAs (g1, g2 or g3) targeting exon 23 shows that *Lb*Cpf1 and *As*Cpf1 have different cleavage efficiency at the *Dmd* exon 23 locus. Red arrowheads show cleavage products of genome editing. M, marker. (**D**) Illustration of *Lb*Cpf1-mediated gRNA (g2) targeting of *Dmd* exon 23. Red arrowheads indicate the cleavage site. The ssODN HDR template contains the *mdx* correction, four silent mutations (green) and a TseI restriction site (underlined).

*Lb*Cpf1-editing with g2 recognizes a PAM sequence 9 bps upstream of the mutation site and creates a staggered double-stranded DNA cut 8 bps downstream of the mutation site (Figure 2.7, panel D). To obtain HDR genome editing, we used a 180 bp single-stranded oligodeoxynucleotide (ssODN) in combination with *Lb*Cpf1 and g2 since it has been shown that ssODNs are more efficient in introducing genomic modification than double-stranded donor plasmids (Long et al., 2014; Wu et al., 2013). We generated a ssODN containing 90 bp of homology sequence flanking the cleavage site, including, four silent mutations and a TseI restriction site to facilitate genotyping as previously described. This ssODN was designed to be used with *Lb*Cpf1 and g2 to correct the C to T mutation within *Dmd* exon 23 and to restore dystrophin in *mdx* mice by HDR.

Correction of muscular dystrophy in mdx mice by CRISPR/LbCpf1-mediated HDR or NHEJ

mdx zygotes were co-injected with in vitro transcribed *Lb*Cpf1 mRNA, in vitro transcribed g2 gRNA and 180 bp ssODN and re-implanted into pseudo-pregnant females

(Figure 2.8, panel A). Three litters of *Lb*Cpf1-edited *mdx* mice were analyzed by T7E1 assay and TseI RFLP (restriction fragment length polymorphism) (Figure 2.8, panel B and C). Out of 24 pups born, 12 were T7E1 positive and 5 carried corrected alleles (*mdx* C1-C5), as detected by TseI RFLP and sequencing (Figure 2.8, panel C and D).

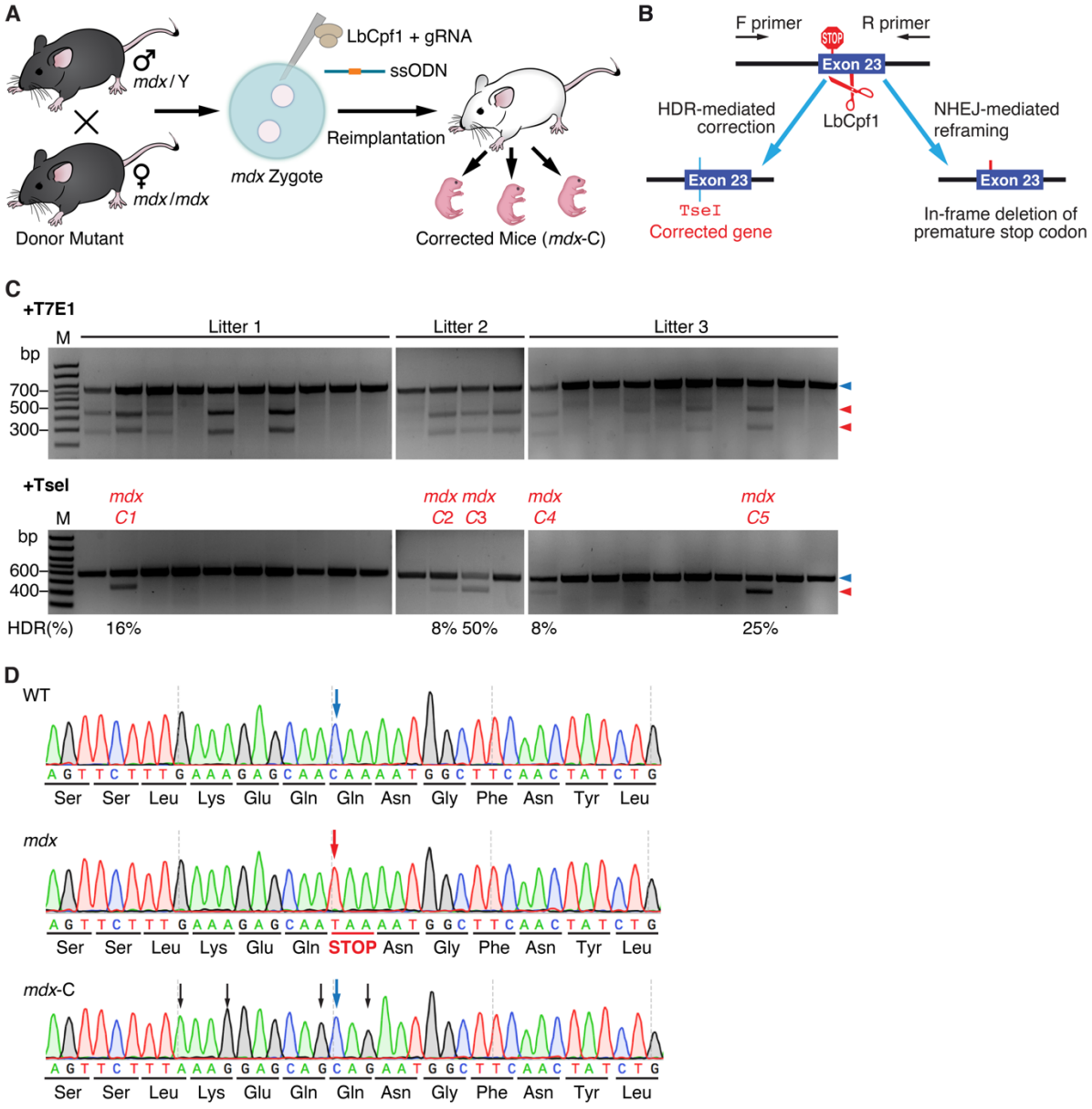

Figure 2.8 CRISPR/*Lb*Cpf1-mediated *Dmd* exon 23 correction in *mdx* mice. (**A**) Strategy of gene correction in *mdx* mice by *Lb*Cpf1-mediated germline editing. Zygotes from intercrosses of *mdx* parents were injected with gene editing components (*Lb*Cpf1 mRNA, g2 gRNA and ssODN) and reimplanted into pseudo-pregnant mothers, which gave rise to pups with gene correction (*mdx*-C). (**B**) Illustration showing *Lb*Cpf1 correction of *mdx* allele by HDR or NHEJ. (**C**) Genotyping results of LbCpf1-edited *mdx* mice. Top panel shows T7E1 assay. Blue arrowhead denotes uncleaved DNA and red arrowhead shows T7E1 cleaved DNA. Bottom panel shows TseI RFLP assay. Blue arrowhead denotes uncorrected DNA. Red arrowhead points to TseI cleavage indicating HDR correction. *mdx*-C1-C5 denotes *Lb*Cpf1-edited *mdx* mice. (**D**) Top panel shows sequence of WT *Dmd* exon 23. Middle panel shows sequence of *mdx Dmd* exon 23 with C to T mutation, which generates a STOP codon. Bottom panel shows sequence of *Dmd* exon 23 with HDR correction by *Lb*Cpf1-mediated editing. Black arrow points to silent mutations introduced by the ssODN HDR template.

Additionally, we performed immunohistochemistry and hematoxylin and eosin (H&E) staining of different tissues from multiple *mdx*-C mice with correction rates from 8%-50%. All *Lb*Cpf1-corrected *mdx*-C mice showed restored dystrophin expression in multiple tissues, including skeletal muscles, heart and brain (Figure 2.9). Importantly, muscles of *mdx*-C mice with 50% genomic correction displayed no sign of fibrosis or inflammatory infiltration (Figure 2.10, panel A). Western blot analysis showed expression of dystrophin protein in multiple skeletal muscle groups, heart and brain (Figure 2.10, panel B), consistent with percentages of dystrophin-positive fibers seen with immunohistochemistry (Figure 2.9).

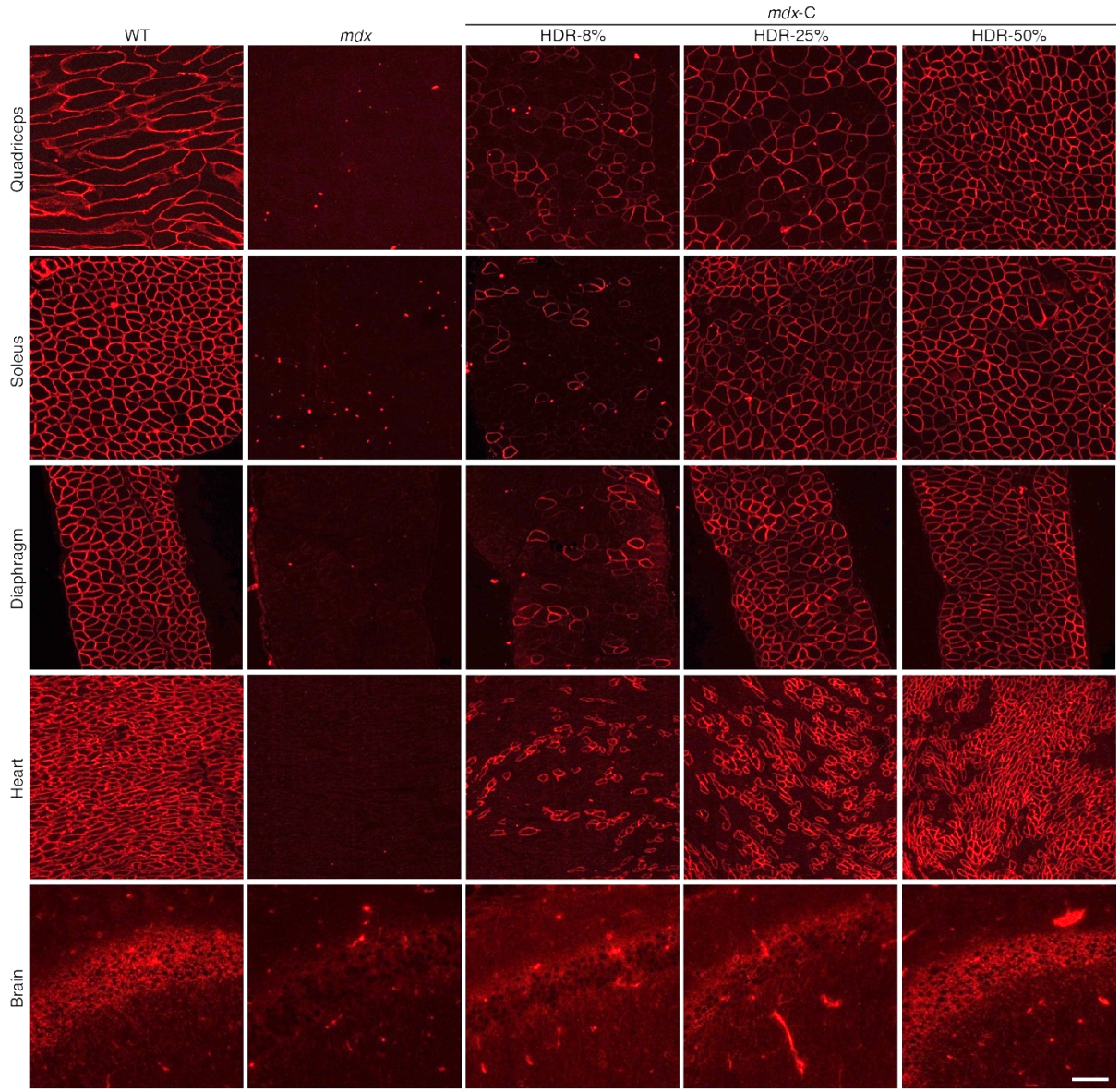

Figure 2.9 Immunohistochemistry of skeletal muscles, heart and brain of WT, *mdx* and *Lb*Cpf1-edited mice (*mdx*-C). Immunohistochemistry of quadriceps, soleus, diaphragm, heart and brain from WT, *mdx,* and *Lb*Cpf1-edited *mdx*-C mice (HDR-8%, HDR-25%, HDR-50% corrected allele) using antibody to dystrophin (red) shows restored dystrophin expression after *Lb*Cpf1-mediated correction. Scale bar = 100 microns.

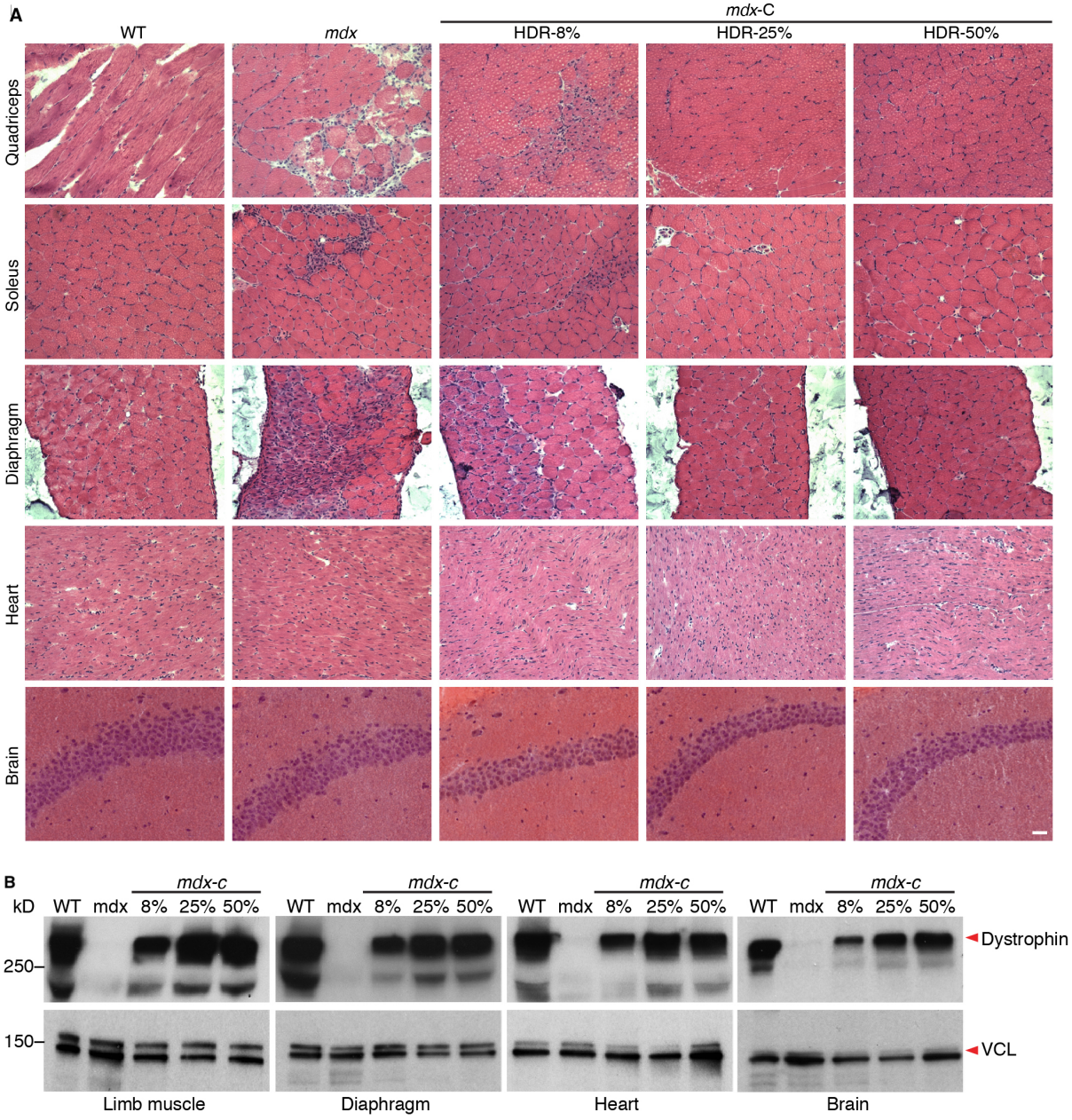

Figure 2.10 H&E staining and Western blot analysis of skeletal muscles, heart and brain of WT, *mdx* and *Lb*Cpf1-edited mice (*mdx*-C). (**A**) H&E staining of quadriceps, soleus, diaphragm, heart and brain of WT, *mdx,* and LbCpf1-edited *mdx*-C mice (HDR-8%, HDR-25%, HDR-50% corrected allele) shows reduced fibrosis and inflammatory infiltration after *Lb*Cpf1-mediated correction. Scale bar = 100 microns. (**B**) Western blot analysis of limb muscles, diaphragm, heart and brain of WT, *mdx,* and LbCpf1-edited *mdx*-C mice (HDR-8%, HDR-25%,

HDR-50% corrected allele) shows restored dystrophin expression after *Lb*Cpf1-mediated correction. Vinculin (VCL) is loading control.

*Lb*Cpf1-mediated correction of the *Dmd* mutation in germ cells was evaluated in eggs and sperm of *Lb*Cpf1-corrected *mdx*-C mice by T7E1 assays and TseI RFLP. All *Lb*Cpf1-corrected *mdx*-C mice carried a corrected allele in their germ cells (Figure 2.11). Genome editing efficiency of Cas9 and *Lb*Cpf1 was compared at the *Dmd* exon 23 locus and no significant difference was observed (Table 2.1).

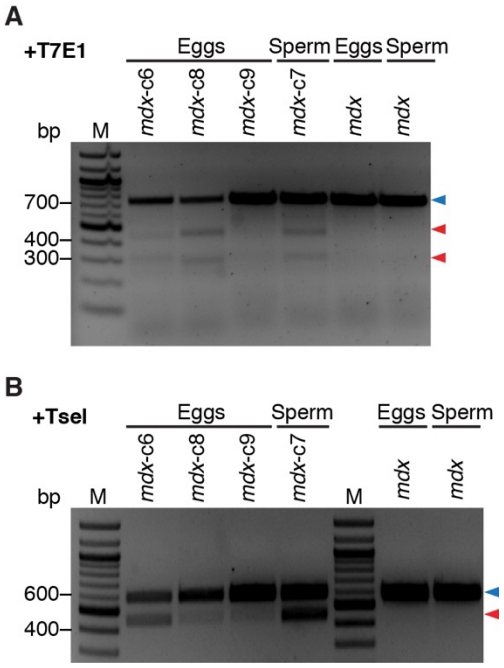

Figure 2.11 T7E1 and TseI RFLP analysis of germ cells from *Lb*Cpf1-edited mice (*mdx*-C) and uncorrected *mdx* mice. (**A**) T7E1 assay of germ cells from *Lb*Cpf1-edited mice (*mdx*-C) and uncorrected *mdx* mice. Blue arrowhead denotes uncleaved DNA and red arrowhead shows T7E1 cleaved DNA. (**B**) TseI RFLP assay of germ cells from *Lb*Cpf1-edited mice (*mdx*-C) and uncorrected *mdx* mice. Blue arrowhead denotes uncorrected DNA. Red arrowhead points to TseI cleavage indicating HDR correction in germline. *mdx*-C6-C9 denotes *Lb*Cpf1-edited *mdx* mice.

Table 2.1 Comparison of CRISPR/Cas9 and CRISPR/Cpf1 mediated HDR correction in *mdx* mice.

Strain	Dose of Cas9/sgRNA/ssODN (ng/ul)	Dose of Cpf1/sgRNA/ssODN (ng/ul)	Injection Method	No. of Transferred Zygotes	No. of Pups/zygotes (%)	No. of Mutant Founders/Pups (%)	No. of HDR/Pups (%)
mdx	10/10/10	-	Nuc	103	29 (28%)	4 (14%)	1 (3.4%)
		-	Nuc+cyt	150	58 (39%)	7 (12%)	4 (6.9%)
mdx	50/20/10	-	Nuc	30	14 (47%)	2 (6.7%)	0
		-	Nuc+cyt	120	23 (19%)	3 (39%)	2 (8.9%)
mdx	-	50/20/10	Nuc+cyt	270	118 (44%)	28 (23.7)	11 (9.3%)

Physiological and functional analyses were also performed in wild-type, *mdx* and *Lb*Cpf1-corrected *mdx*-C mice. Serum creatine kinase levels were decreased substantially in *mdx*-C mice and were inversely correlated with the percentage of genome correction (Table 2.2). Importantly, the forelimb grip strength test indicated that *Lb*Cpf1-corrected *mdx*-C mice had improved muscle strength compared to *mdx* mice (Table 2.2).

Table 2.2 Serum CK measurement and forelimb grip strength of WT, *mdx* and *Lb*Cpf1-corrected *mdx*-C mice. M, male; F, female.

Table 1. Serum CK measurement and forelimb grip strength of WT, *mdx*, and LbCpf1-corrected *mdx*-C mice. M, male; F, female.

Mouse no.	Percent correction by HDR	Sex	Weight (g)	CK (U/liter)	Forelimb grip strength (grams of force)						
					Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Average ± SD
WT-1	—	M	16.6	455	103	106	97	71	88	82	91.2 ± 13.4
WT-2	—	M	19.5	220	87	95	73	73	74	78	80.0 ± 9.1
WT-3	—	M	18.7	306	79	97	74	78	84	84	82.7 ± 8.0
WT-4	—	F	15.5	184	86	97	85	83	85	88	87.3 ± 5.0
WT-5	—	F	15.4	175	88	85	100	96	88	86	90.5 ± 6.1
WT-6	—	F	12.6	157	76	75	73	78	64	61	71.2 ± 7.0
mdx-1	0	M	18.8	8579	76	92	86	33	32	29	58.0 ± 29.9
mdx-2	0	M	21.0	9440	62	58	54	45	45	47	51.8 ± 7.3
mdx-3	0	M	21.5	5936	77	54	69	57	56	61	62.3 ± 8.9
mdx-4	0	F	16.1	6306	69	63	69	61	67	60	64.8 ± 4.0
mdx-5	0	F	15.8	6349	83	88	85	59	55	54	70.7 ± 16.2
mdx-6	0	F	16.2	4168	69	71	53	59	59	57	61.3 ± 7.1
mdx-C1	16%	F	21.7	1233	112	111	112	115	101	109	110.0 ± 4.8
mdx-C2	8%	F	19.7	4920	119	109	108	95	86	85	100.3 ± 13.8
mdx-C3	50%	F	20.2	248	110	115	114	112	112	104	111.2 ± 3.9
mdx-C4	8%	M	22.7	6607	49	45	42	30	25	21	35.3 ± 11.5
mdx-C5	25%	F	17.7	3239	92	110	96	91	110	95	99.0 ± 8.7

Discussion

In this study, we show that the newly discovered CRISPR-Cpf1 nuclease can efficiently correct *DMD* mutations in patient-derived iPSCs and *mdx* mice, allowing for restoration of dystrophin expression. Lack of dystrophin in DMD has been show to disrupt integrity of the sarcolemma, causing mitochondria dysfunction and oxidative stress (Millay et al., 2008; Mourkioti et al., 2013). We found increased mitochondrial DNA and higher oxygen consumption rates in LbCpf1-corrected iPSC-derived cardiomyocytes compared to uncorrected DMD iPSC-derived cardiomyocytes. Metabolic abnormalities of human DMD iPSC-derived cardiomyocytes were also rescued by Cpf1-mediated genomic editing. Our findings also demonstrated the robustness and efficiency of Cpf1 in mouse genome editing. By using HDR-mediated correction, the ORF of the mouse *Dmd* gene was completely restored and pathophysiological hallmarks of the dystrophic phenotype such as fibrosis and inflammatory infiltration were also rescued.

Two different strategies, “exon reframing” and “exon skipping”, were applied to restore the ORF of the *DMD* gene using *Lb*Cpf1-mediated genome editing. Reframing creates small INDELs and restores the ORF by placing out-of-frame codons in-frame. Only one gRNA is required for reframing. Although we did not observe any differences in subcellular localization between WT dystrophin protein and reframed dystrophin protein by immunocytochemistry, we observed differences in dystrophin expression level, mitochondrial DNA quantity, and oxygen consumption rate in different edited clones, suggesting that reframed dystrophin may not be structurally or functionally identical to WT dystrophin.

Various issues should be considered with respect to the use of one or two gRNAs with Cpf1-editing. Here we show that two gRNAs are more effective than one gRNA for disruption of the splice acceptor site compared to reframing. When using two gRNAs, Cpf1-editing excises the intervening region and not only removes the splice acceptor site but can be designed to remove deleterious "AG" nucleotides, eliminating the possibility of generating a pseudo-splice acceptor site. However, with two gRNAs there is the necessity that both gRNAs cleave simultaneously, which may not occur. If only one of the two gRNAs cleaves, the desired deletion will not be generated. However, there remains the possibility that cleavage with one of the two gRNAs will generate INDELS at the targeted exon region, reframing the ORF, since in theory, one third of the INDELS will be in-frame. Using one gRNA to disrupt the splice acceptor site seems more efficient because it eliminates the need for two simultaneous cuts to occur. However, there is uncertainty with respect to the length of the INDEL generated by one gRNA-mediated editing. More importantly, with one gRNA there remains the possibility of leaving exonic "AG" nucleotides near the cleavage site, which can serve as an alternative pseudo-splice acceptor site.

With its unique T-rich PAM sequence, Cpf1 further expands the genome editing range of the CRISPR family, which is important for potential correction of other disease-related mutations since not all mutation sites contain G-rich PAM sequences for *Sp*Cas9 or PAMs for other Cas9 orthologues. Moreover, the staggered cut generated by Cpf1 may be also advantageous for NHEJ-mediated genome editing (Maresca et al., 2013). Finally, the *Lb*Cpf1 used in this study is 140-amino-acids smaller than the most widely used *Sp*Cas9, which would enhance packaging and delivery by AAV.

Our findings show that Cpf1 is highly efficient in correcting human and mouse *Dmd* mutations *in vitro* and *in vivo*. CRISPR-Cpf1-mediated genome editing represents a new and powerful approach to permanently eliminate genetic mutations and rescue abnormalities associated with DMD and other disorders.

Materials and Methods

Generation of pLbCpf1-2A-GFP and pAsCpf1-2A-GFP plasmids

Human codon-optimized *Lb*Cpf1 and *As*Cpf1 were PCR amplified from pY016 plasmid (pcDNA3.1-h*Lb*Cpf1), a gift from Feng Zhang (Addgene plasmid # 69988) and pY010 plasmid (pcDNA3.1-h*As*Cpf1), a gift from Feng Zhang (Addgene plasmid # 69982), respectively. Cpf1 cDNA and T2A-GFP DNA fragment were cloned into the backbone of the p*Sp*Cas9(BB)-2A-GFP (PX458) plasmid, a gift from Feng Zhang (Addgene plasmid # 48138) that was cut with AgeI/EcoRI to remove *Sp*Cas9(BB)-2A-GFP. In-Fusion HD cloning kit (Takara Bio) was used. Cpf1 guide RNAs (gRNAs) targeting the human *DMD* or the mouse *Dmd* locus were sub-cloned into a newly generated p*Lb*Cpf1-2A-GFP plasmid and p*As*Cpf1-2A-GFP plasmid using BbsI digestion and T4 ligation. Detailed primer sequences can be found in supplementary materials (Table 2.3).

Human iPSC maintenance, nucleofection and differentiation

Human iPSCs (RBRC-HPS0164) were purchased from Cell Bank RIKEN BioResource Center. Human iPSCs were cultured in mTeSR^{TM}1 media (STEMCELL Technologies) and

passaged approximately every 4 days (1:18 split ratio). One hour before nucleofection, iPSCs were treated with 10 uM ROCK inhibitor (Y-27632) and dissociated into single cells using Accutase (Innovative Cell Technologies, Inc.). $1x10^6$ iPS cells were mixed with 5 ug of p*Lb*Cpf1-2A-GFP or p*As*Cpf1-2A-GFP plasmid and nucleofected using the P3 Primary Cell 4D-Nucleofector X kit (Lonza) according to manufacturer's protocol. After nucleofection, iPSCs were cultured in mTeSR™1 media supplemented with 10 uM ROCK inhibitor, penicillin-streptomycin (1:100) (ThermoFisher Scientific) and 100 ug/ml Primosin (InvivoGen). Three days post-nucleofection, GFP(+) and (-) cells were sorted by FACS and subjected to T7E1 assay. Single clones derived from GFP(+) iPSCs were picked and sequenced. iPSCs were induced to differentiate into cardiomyocytes, as previously described (Burridge et al., 2015).

Genomic DNA isolation

Genomic DNA of mouse 10T1/2 fibroblasts and human iPSCs was isolated using Quick-gDNA MiniPrep kit (Zymo Research) according to manufacturer's protocol.

Reverse-transcription PCR

RNA was isolated using TRIzol (ThermoFisher Scientific), according to manufacturer's protocol. cDNA was synthesized using iScript Reverse Transcription Supermix (Bio-Rad Laboratories) according to manufacturer's protocol. RT-PCR was performed using primers flanking *DMD* exon 47 and 52 (forward: 5'-CCCAGAAGAGCAAGATAAACTTGAA-3'; reverse: 5'-CTCTGTTCCAAATCCTGCTTGT-3'). RT-PCR products amplified from WT

cardiomyocytes, uncorrected cardiomyocytes and exon 51 skipped cardiomyocytes were 717 bps, 320 bps and 87 bps, respectively.

Dystrophin Western blot analysis

Western blot analysis for human iPSC-derived cardiomyocytes was performed as previously described (Long et al., 2014) using rabbit anti-dystrophin antibody (Abcam, ab15277) and mouse anti-cardiac myosin heavy chain antibody (Abcam, ab50967). For mouse skeletal muscles, heart and brain, the Western blot was performed as previously described (Long et al., 2014) using mouse anti-dystrophin antibody (Sigma-Aldrich, D8168) and mouse anti-vinculin antibody (Sigma-Aldrich, V9131).

Dystrophin immunocytochemistry and immunohistochemistry

iPSC-derived cardiomyocytes fixed with acetone, blocked with serum cocktail (2% normal horse serum/2% normal donkey serum/0.2% bovine serum albumin (BSA)/PBS), and incubated with dystrophin antibody (MANDYS8, 1:800, Sigma-Aldrich) and troponin-I antibody (H170, 1:200, Santa Cruz Biotechnology) in 0.2% BSA/PBS. Following overnight incubation at 4°C, they were incubated with secondary antibodies (biotinylated horse anti-mouse IgG, 1:200, Vector Laboratories, fluorescein-conjugated donkey anti-rabbit IgG, 1:50, Jackson Immunoresearch) for one-hour. Nuclei were counterstained with Hoechst 33342 (Molecular Probes). Immunohistochemisty of skeletal muscles, heart and brain were performed as previously described (Long et al., 2014) using dystrophin antibody (MANDYS8, 1:800, Sigma-Aldrich). Nuclei were counterstained with propidium iodide (Molecular Probes).

Mitochondrial DNA copy number quantification

Genomic and mitochondrial DNA were isolated using Trizol, followed by back extraction as previously described (Zechner et al., 2010). KAPA SYBR FAST qPCR kit (Kapa Biosystems) was used to perform real-time PCR to quantitatively determine mitochondrial DNA copy number. Human mitochondrial *ND1* gene was amplified using primers (forward: 5'-CGCCACATCT-ACCATCACCCTC -3'; reverse: 5'- CGGCTAGGCTAGAGGTGGCTA -3'). Human genomic *LPL* gene was amplified using primers (forward: 5'- GAGTATGCAGA-AGCCCCGAGTC -3'; reverse: 5'- TCAACATGCCCAACTGGTTTCTGG -3'). mtDNA copy number per diploid genome was calculated using formula: ΔC_T = (mtND1 C_T – LPL C_T) and mtDNA copy number per diploid genome = $2 \times 2^{-\Delta CT}$

Cellular respiration rates

Oxygen consumption rates (OCR) were determined in human iPSC-derived cardiomyocytes using the XF24 Extracellular Flux Analyzer (Seahorse Bioscience) following the manufacturer's protocol as previously described (Baskin et al., 2014).

In vitro transcription of LbCpf1 mRNA and gRNA

Human codon-optimized *Lb*Cpf1 was PCR amplified from p*Lb*Cpf1-2A-GFP to include the T7 promoter sequence (Table 2.3). The PCR product was transcribed using mMESSAGE mMACHINE T7 transcription kit (ThermoFisher Scientific) according to manufacturer's protocol. Synthesized *Lb*Cpf1 mRNA were poly-A tailed with *E. coli* Poly(A) Polymerase (New England Biolabs) and purified using NucAway spin columns (ThermoFisher Scientific).

The template for *Lb*Cpf1 gRNA in vitro transcription was PCR amplified from p*Lb*Cpf1-2A-GFP plasmid and purified using Wizard SV gel and PCR clean-up system (Promega). The *Lb*Cpf1 gRNA was synthesized using MEGAshortscript T7 transcription kit (ThermoFisher Scientific) according to manufacturer's protocol. Synthesized *Lb*Cpf1 gRNA were purified using NucAway spin columns (ThermoFisher Scientific).

Single-stranded oligodeoxynucleotide (ssODN)

ssODN was used as HDR template and synthesized by Integrated DNA Technologies as 4nM Ultramer Oligonucleotides. ssODN was mixed with LbCpf1 mRNA and gRNA directly without purification. The sequence of ssODN is:

5’TGATATGAATGAAACTCATCAAATATGCGTGTTAGTGTAAATGAACTTC TATTTAATTTTGAGGCTCTGCAAAGTTCTTTAAAGGAGCAGCAGAATGGCTTCAA CTATCTGAGTGACACTGTGAAGGAGATGGCCAAGAAAGCACCTTCAGAAATATG CCAGAAATATCTGTCAGAATTT-3’

CRISPR/Cpf1-mediated genome editing by one-cell embryo injection

All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Texas Southwestern Medical Center. Detailed injection procedures were performed as described previously (Long et al., 2014). The only modification was replacing Cas9 mRNA and Cas9 gRNAs with *Lb*Cpf1 mRNA and *Lb*Cpf1 gRNAs.

Mouse forelimb grip strength test and serum creatine kinase (CK) measurement

Grip strength test and serum CK measurement were performed as previously described (Long et al., 2014) by the by the Neuro-Models Core Facility and the Metabolic Phenotyping Core at UT Southwestern Medical Center, respectively.

PCR amplification of genomic DNA, T7E1 assay, and TseI RFLP analysis

All these protocols were performed as previously published (Long et al., 2014).

Statistical analysis

Statistical analysis was assessed by two-tailed Student's t-test. Data are shown as mean ± SEM. A $P<0.05$ value was considered statistically significant

Table 2.3 Primers used in this study.

	Primer Name		Primer Sequence
Cloning primers for pCpf1-2A-GFP	AgeI-nLbCpf1-F1	F	tttttttcaggttGGaccggtgccaccATGAGCAAGCTGGA
	nLbCpf1-R1	R	TGGGGTTATAGTAGGCCATCCACTTC
	nLbCpf1-F2	F	GATGGCCTACTATAACCCCAGCG
	nLbCpf1-R2	R	GGCATAGTCGGGGACATCATATG
	AgeI-nAsCpf1-F1	F	tttttttcaggttGGaccggtgccaccATGACACAGTTCGAG
	nAsCpf1-R1	R	TCCTTCTCAGGATTGTTCAGGTCGTA
	nAsCpf1-F2	F	CTGAACAATCCTGAGAAGGAGCC
	nAsCpf1-R2	R	GGCATAGTCGGGGACATCATATG
	nCpf1-2A-GFP-F	F	ATGATGTCCCCGACTATGCCgaattcGGCAGTGGAGAGGG
	nCpf1-2A-GFP-R	R	AGCGAGCTCTAGttagaattcCTTGTACAG
In vitro transcription of LbCpf1 mRNA	T7-Scaffold-F	F	CACCAGCGCTGCTTAATACGACTCACTATAGGGAAAT
	T7-Scaffold-R	R	AGTAGCGCTTCTAGACCCTCACTTCCTACTCAG
	T7-nLb-F1	F	AGAAGAAATATAAGACTCGAGgccaccATGAGCAAGCTGGAGAAGTTTAC
	T7-nLb-R1	R	TGGGGTTATAGTAGGCCATCC
	T7-nLB-NLS-F2	F	GATGGCCTACTATAACCCCAGCG
	T7-nLB-NLS-R2	R	CCCGCAGAAGGCAGCGTCGACTTAGGCATAGTCGGGGACATCATATG
	T7-nAs-F1	F	AGAAGAAATATAAGACTCGAGgccaccATGACACAGTTCGAGGGCTTTAC
	T7-nAs-R1	R	TCCTTCTCAGGATTGTTCAGGTCGTA
	T7-nAs-NLS-F2	F	CTGAACAATCCTGAGAAGGAGCC
	T7-nAs-NLS-R2	R	CCCGCAGAAGGCAGCGTCGACTTAGGCATAGTCGGGGACATCATATG
Human *DMD* Exon 51 gRNA	nLb-DMD-E51-g1-Top	F	CACCGTAATTTCTACTAAGTGTAGATgCTCCTACTCAGACTGTTACTCTGTTTTTTT
	nLb-DMD-E51-g1-Bot	R	AAACAAAAAAAACAGAGTAACAGTCTGAGTAGGAGcATCTACACTTAGTAGAAATTAC
	nLb-DMD-E51-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATtaccatgtattgctaaacaaagtaTTTTTTT
	nLb-DMD-E51-g2-Bot	R	AAACAAAAAAAtactttgtttagcaatacatggtaATCTACACTTAGTAGAAATTAC
	nLb-DMD-E51-g3-Top	F	CACCGTAATTTCTACTAAGTGTAGATattgaagagtaacaatttgagccaTTTTTTT
	nLb-DMD-E51-g3-Bot	R	AAACAAAAAAAtggctcaaattgttactcttcaatATCTACACTTAGTAGAAATTAC
	nAs-DMD-E51-g1-Top	F	CACCGTAATTTCTACTCTTGTAGATgCTCCTACTCAGACTGTTACTCTGTTTTTTT
	nAs-DMD-E51-g1-Bot	R	AAACAAAAAAAACAGAGTAACAGTCTGAGTAGGAGcATCTACAAGAGTAGAAATTAC
Human *DMD* Exon 51 T7E1	DMD-E51-T7E1-F1	F	ttccctggcaaggtctga
	DMD-E51-T7E1-R1	R	ATCCTCAAGGTCACCCACC
Human cardiomyocytes	Riken51-RT-PCR-F1	F	CCCAGAAGAGCAAGATAAACTTGAA
	Riken51-RT-PCR-R1	R	CTCTGTTCCAAATCCTGCATTGT
Human cardiomyocytes mtDNA copy number qPCR	hmtND1-qF1	F	CGCCACATCTACCATCACCCTC
	hmtND1-qR1	R	CGGCTAGGCTAGAGGTGGCTA
	hLPL-qF1	F	GAGTATGCAGAAGCCCCGAGTC
	hLPL-qR1	R	TCAACATGCCCAACTGGTTTCTGG
Mouse *Dmd* Exon 23 gRNA	nLb-dmd-E23-g1-Top	F	CACCGTAATTTCTACTAAGTGTAGATaggctctgcaaagttctTTGAAAGTTTTTTT
	nLb-dmd-E23-g1-Bot	R	AAACAAAAAAAACTTTCAAagaactttgcagagcctATCTACACTTAGTAGAAATTAC
	nLb-dmd-E23-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAGCAACAAAATGGCttcaacTTTTTTT
	nLb-dmd-E23-g2-Bot	R	AAACAAAAAAAgttgaaGCCATTTTGTTGCTCTTTATCTACACTTAGTAGAAATTAC
	nLb-mdmd-E23-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAGCAATAAAATGGCttcaacTTTTTTT
	nLb-mdmd-E23-g2-Bot	R	AAACAAAAAAAgttgaaGCCATTTTATTGCTCTTTATCTACACTTAGTAGAAATTAC
	nLb-dmd-E23-g3-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAACTTTGCAGAGCctcaaaaTTTTTTT
	nLb-dmd-E23-g3-Bot	R	AAACAAAAAAAttttgagGCTCTGCAAAGTTCTTTATCTACACTTAGTAGAAATTAC
	nLb-dmd-I22-g1-Top	F	CACCGTAATTTCTACTAAGTGTAGATctgaatatctatgcattaataactTTTTTTT
	nLb-dmd-I22-g1-Bot	R	AAACAAAAAAAagttattaatgcatagatattcagATCTACACTTAGTAGAAATTAC
	nLb-dmd-I22-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATtattatattacagggcatattataTTTTTTT
	nLb-dmd-I22-g2-Bot	R	AAACAAAAAAAtataatatgccctgtaatataataATCTACACTTAGTAGAAATTAC
	nLb-dmd-I23-g3-Top	F	CACCGTAATTTCTACTAAGTGTAGATAGgtaagccgaggtttggcctttaTTTTTTT
	nLb-dmd-I23-g3-Bot	R	AAACAAAAAAAtaaaggccaaacctcggcttacCTATCTACACTTAGTAGAAATTAC
	nLb-dmd-I23-g4-Top	F	CACCGTAATTTCTACTAAGTGTAGATcccagagtccttcaaagatattgaTTTTTTT
	nLb-dmd-I23-g4-Bot	R	AAACAAAAAAAtcaatatctttgaaggactctgggATCTACACTTAGTAGAAATTAC
In vitro transcription of LbCpf1 gRNA	T7-Lb-dmd-E23-uF	F	GAATTGTAATACGACTCACTATAGGGTAATTTCTACTAAGTGTAGAT
	T7-Lb-dmd-E23-g1-R	R	CTTTCAAagaactttgcagagcctATCTACACTTAGTAGAAATTA
	T7-Lb-dmd-E23-mg2-R	R	GttgaaGCCATTTTATTGCTCTTTATCTACACTTAGTAGAAATTA
	T7-Lb-dmd-E23-g3-R	R	ttttgagGCTCTGCAAAGTTCTTTATCTACACTTAGTAGAAATTA
	T7-Lb-dmd-I22-g2-R	R	tataatatgccctgtaatataataATCTACACTTAGTAGAAATTACCCTATAGTGAG
	T7-Lb-dmd-I22-g4-R	R	tcaatatctttgaaggactctgggATCTACACTTAGTAGAAATTACCCTATAGTGAG
Mouse *Dmd* Exon 23 T7E1	Dmd-E23-T7E1-F729	F	Gagaaacttctgtgatgtgaggacata
	Dmd-E23-T7E1-R1	R	CAAACCTCGGCTTACCTGAAAT
	Dmd-E23-T7E1-R729	R	caatatctttgaaggactctgggtaaa
	Dmd-E23-T7E1-R3	R	aattaatagaagtcaatgtagggaagg

CHAPTER THREE

ENHANCED CRISPR/CAS9 CORRECTION OF DUCHENNE MUSCULAR DYSTROPHY IN MICE BY A SELF-COMPLEMENTARY AAV DELIVERY SYSTEM

Acknowledgement

Parts of this chapter, including figures, have been reproduced, with or without modifications, from my previously published work (Zhang et al., 2020).

Abstract

Duchenne muscular dystrophy (DMD) is a lethal neuromuscular disease, caused by mutations in the dystrophin gene (DMD). Mutations that delete exon 44, thereby disrupting the DMD open reading frame, are one of the most common causes of DMD. We have previously applied CRISPR/Cas9-mediated "single-cut" gene editing to correct diverse genetic mutations in animal models of DMD. However, high doses of AAV are required for efficient in vivo genome editing, posing challenges for clinical application. In this study, we packaged Cas9 nuclease in conventional single-stranded AAV (ssAAV) and CRISPR single guide RNAs in double-stranded self-complementary AAV (scAAV) and delivered this dual AAV system into a mouse model of DMD harboring an exon 44 deletion. The doses of scAAV required for efficient gene editing were at least 20-fold lower than with ssAAV. Mice receiving systemic treatment showed restoration of dystrophin expression in all skeletal muscle groups and the heart, reduced DMD pathological phenotypes, and improved muscle contractility. These

findings are the first to show that the efficiency of CRISPR/Cas9-mediated genome editing can be significantly improved by using the scAAV system and represent an important advancement toward therapeutic translation of genome editing for treating neuromuscular diseases.

Introduction

Duchenne muscular dystrophy (DMD) is an X-linked monogenic neuromuscular disease caused by mutations in the *DMD* gene, which encodes dystrophin (Hoffman et al., 1987; Koenig et al., 1987). Dystrophin, together with dystroglycans and sarcoglycans, maintains sarcolemma integrity and stability by interacting with intracellular actin and extracellular laminin (Campbell and Kahl, 1989; Gao and McNally, 2015; Guiraud et al., 2015). More than 7,000 mutations have been identified in DMD patients, including single- and multi-exon deletions or duplications, and small missense or nonsense substitutions (Aartsma-Rus et al., 2006; Bladen et al., 2015). Patients with DMD manifest progressive muscle weakness and ultimately develop fatal respiratory and cardiac failure in their mid-20s.

To date, two clinical therapies are available for DMD treatment, including steroid supplementation and morpholino antisense oligomer injection (Bushby et al., 2004; Charleston et al., 2018). Long-term corticosteroid supplement partially alleviates DMD pathological phenotypes but cannot restore dystrophin expression. Morpholino antisense oligomers allow skipping of mutant *DMD* exons, but less than 1% of normal levels of dystrophin protein can be restored by this treatment (Charleston et al., 2018). In addition, several clinical trials are currently evaluating the therapeutic benefits of truncated versions of dystrophin delivered by adeno-associated virus (AAV) (Duan, 2018). However, these gene replacement therapies

cannot restore the expression of endogenous dystrophin protein and are dependent on the expression pattern of the exogenous promoters within the AAV, as well as the longevity of AAV expression. Thus, developing a strategy for permanent and efficient correction of mutations in the endogenous *DMD* gene may provide an ultimate cure for this lethal neuromuscular disorder.

Application of the CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins) system for engineering site-specific DNA double-stranded breaks (DSB) provides simplicity and precision in mammalian genome editing (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013b). We and others showed that the CRISPR-Cas system can be used to efficiently correct missense mutations in mouse models of DMD by homology directed repair (HDR)-based germline editing or non-homologous end joining (NHEJ)-based postnatal editing (Bengtsson et al., 2017; Hakim et al., 2018; Long et al., 2016; Long et al., 2014; Nelson et al., 2016; Nelson et al., 2019; Tabebordbar et al., 2016; Zhang et al., 2017b; Zhu et al., 2017). While missense and nonsense substitutions only account for ~20% of *DMD* mutations, single- or multi-exon deletions are more prevalent (~68%) in DMD populations (Aartsma-Rus et al., 2006; Bladen et al., 2015). We recently reported the successful rescue of DMD phenotypes in mice and dogs harboring exon 44 or 50 deletions by injecting recombinant AAV9-packaged Cas9 nuclease and single guide RNAs (sgRNAs) (Amoasii et al., 2018; Amoasii et al., 2017; Min et al., 2019). These studies demonstrated that the CRISPR-Cas system can be deployed to correct diverse genetic mutations that cause DMD and offer the prospect of a potential gene therapy for the permanent correction of DMD.

Currently the most widely used delivery vector for gene therapy is recombinant AAV, which is a nonenveloped virus with a single-stranded linear DNA viral genome (Samulski and Muzyczka, 2014). As the largest tissue in the human body, skeletal muscle accounts for ~40% of body weight. Therefore, a high dose of AAV (5.5×10^{14} to 1.8×10^{15} vg/kg) is required to achieve long-term, efficient genome editing in animal models of DMD (Bengtsson et al., 2017; Hakim et al., 2018; Min et al., 2019; Nelson et al., 2019). However, several studies in large animals reported that systemic administration of high doses of AAV ($\geq 1.5 \times 10^{14}$ vg/kg) may cause acute liver toxicity (Hinderer et al., 2018; Kornegay et al., 2010). In addition, in our previous study, we found that the efficiency of in vivo CRISPR/Cas9-mediated genome editing was highly dose-dependent, and that elevating the dose of sgRNA AAV relative to Cas9 AAV enhanced the efficiency of genome editing (Min et al., 2019). Moreover, it has been suggested that the sgRNA AAV genome is preferentially depleted after systemic delivery of CRISPR-Cas9 genome editing components (Hakim et al., 2018). Therefore, systemic delivery of CRISPR-Cas9 genome editing components by a high dose of single-stranded AAV (ssAAV) for the treatment of DMD remains challenging.

In order to reduce the viral dose used for gene therapy without compromising genome editing efficiency and to prevent preferential depletion of the sgRNA AAV genome, we packaged a CRISPR sgRNA expression cassette into a double-stranded AAV vector. A double-stranded AAV genome can be generated by mutating the terminal resolution site sequence on one side of the inverted terminal repeats (ITR), leading to production of self-complementary AAV (scAAV) (McCarty et al., 2003; Wang et al., 2003). Unlike conventional ssAAV, scAAV can bypass the second-strand synthesis, which is a rate-limiting step for gene expression

(Ferrari et al., 1996; Fisher et al., 1996). Moreover, double-stranded scAAV is less prone to DNA degradation after viral transduction, thereby increasing the number of copies of stable episomes (McCarty, 2008; Ren et al., 2005). The scAAV system has been used in several gene replacement clinical trials for the treatment of spinal muscular atrophy and limb-girdle muscular dystrophy (Mendell et al., 2017; Pozsgai et al., 2017).

In this study, we performed in vivo genome editing in mice with a deletion of *Dmd* exon 44 (ΔEx44) by coupling ssAAV-packaged *Sp*Cas9 nuclease with scAAV-expressed sgRNAs. This dual AAV delivery system provided significant improvements in viral transduction efficiency, genome editing, and functional recovery in skeletal muscles and heart. Of note, at least 20-fold less scAAV was required to achieve these improvements compared to the ssAAV treated cohort. Thus, the scAAV system represents a promising strategy for delivering CRISPR-Cas9 genome editing components and represents an important advancement toward potential therapeutic translation.

Results

Strategies for CRISPR/Cas9-mediated genome editing of Dmd exon 45

Deletion of exon 44 of the human *DMD* gene generates a premature stop codon in exon 45 and represents one of most common mutations of DMD. As a strategy to correct exon 44 out-of-frame deletion mutations, we designed a sgRNA to target the splice acceptor region of exon 45 (Figure 3.1, panel A). This sgRNA recognizes a 5'-TGG-3' protospacer adjacent motif (PAM) in exon 45 and generates insertions and deletions (INDELs) 7 base pairs (bp)

downstream of the 5'-AG-3' splice acceptor site (Figure 3.1, panel B). Depending on the size of INDELs, two types of NHEJ-mediated DNA repair events can restore the open reading frame (ORF) of the *Dmd* gene. These include exon 45 skipping, if the INDEL is large enough to delete the 5'-AG-3' splice acceptor sequence in exon 45, or reframing of exon 45 through INDELs that either insert one nucleotide (3n+1) or delete two nucleotides (3n-2) (Figure 3.1, panel A).

To test whether double-stranded scAAV is capable of packaging sgRNAs, we cloned the sgRNA expression cassette into a scAAV vector and the conventional ssAAV vector as a control (Figure 3.1, panel C). Alkaline denaturing gel electrophoresis was performed to confirm the integrity of both AAVs (Figure 3.1, panel D). The size of ssAAV-sgRNA is 3.9 kilo-nucleotides (knt) and remains unchanged after alkaline gel electrophoresis. The size of scAAV-sgRNA is 1.4 kilo-base pairs and is doubled to 2.8 knts under denaturing conditions, indicative of the double-stranded viral genome.

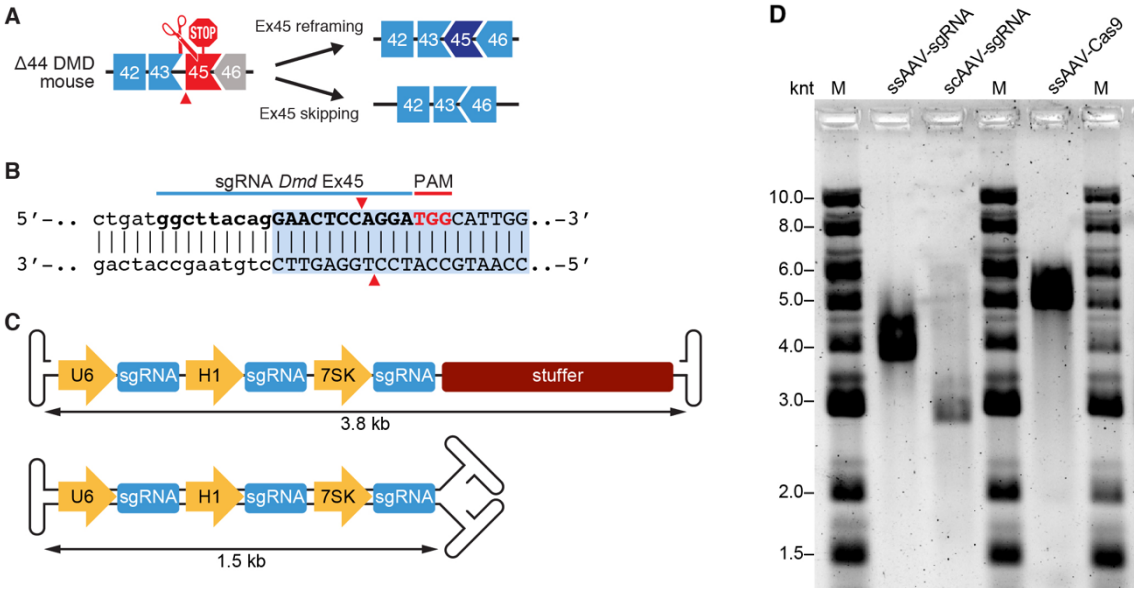

Figure 3.1. Strategies for CRISPR/Cas9-mediated genome editing in *Dmd* ΔEx44 mice. (**A**) An out-of-frame deletion of Dmd exon 44 results in splicing of exon 43 to 45, generating a premature stop codon in exon 45. A CRISPR/Cas9-mediated "single-cut" strategy was designed to restore the open reading frame (ORF) of the Dmd gene. If the genomic insertions and deletions (INDELs) result in one nucleotide insertion (3n+1) or two nucleotides deletion (3n-2), exon 45 will be reframed with adjacent exon 43 and 46. If the INDEL is large enough to delete the 5'-AG-3' splice acceptor sequence, exon 45 will be skipped, resulting in splicing of exon 43 to exon 46. (**B**) Illustration of sgRNA targeting Dmd exon 45. This sgRNA recognizes a 5'-TGG-3' PAM in exon 45 and generates a cut 7 base pairs downstream of the 5'-AG-3' splice acceptor site. (**C**) Illustration of AAV vectors used to deliver the sgRNA expression cassette. Three copies of the same sgRNA are driven by three RNA polymerase III promoters, U6, H1, and 7SK. The top vector produces ssAAV. A 2.3k stuffer sequence was cloned into the ssAAV vector for optimal packaging. The bottom vector produces double-stranded scAAV. (**D**) The viral genomes of the Cas9 vector and sgRNA vectors were analyzed by gel electrophoresis under alkaline denaturing conditions. The size of ssAAV-sgRNA and ssAAV-Cas9 is 3.9 and 5.1 knt, respectively, and remains unchanged after alkaline gel electrophoresis. The size of scAAV-sgRNA is 1.4 kilo-base pair and is doubled to 2.8 knts under denaturing conditions, indicating its double-stranded viral genome. M, marker; knt, kilo-nucleotides.

In vitro genome editing using ssAAV or scAAV-packaged sgRNA

To compare the efficiency of ssAAV and scAAV-packaged sgRNAs in vitro, we differentiated *Sp*Cas9-expressing C2C12 mouse myoblasts for 5 days to myotubes and transduced the myotubes with each of the AAVs. One week after viral transduction, we performed Tracking of INDELs by Decomposition (TIDE) analysis to detect INDELs within the *Dmd* exon 45 region. We found that the total INDELs exhibited a dose-dependent curve for both ssAAV and scAAV-expressed sgRNA after one week post viral transduction (Figure 3.2, panel A). Specifically, to reach a level of 10% INDELs required 5×10^8 vg/mL of scAAV and 1×10^{10} vg/mL of ssAAV, representing a 20-fold increase in efficiency of scAAV. To

reach an intermediate level of INDELS of ~22%, 40-fold less scAAV (1.8×10^9 vg/mL) was required compared to ssAAV (7.8×10^{10} vg/mL).

Furthermore, a high level of INDELs (over 40%) was achieved by 7.2×10^9 vg/mL of scAAV, whereas ssAAV required 5×10^{12} vg/mL, representing a 70-fold improvement in efficiency with scAAV. We also analyzed the INDEL composition in myotubes transduced with ssAAV or scAAV and found that ~50% of total INDEL events contained a +1 nt insertion, which can bring exon 45 in-frame with exon 43 (Figure 3.2, panel B). Therefore, scAAV-expressed sgRNA demonstrated enhanced efficiency by in vitro genome editing at *Dmd* exon 45 compared to the conventional ssAAV-expressed sgRNA. Moreover, the majority of the INDEL events (over 50%) contained a single nt insertion, which is able to restore the *Dmd* exon 45 ORF.

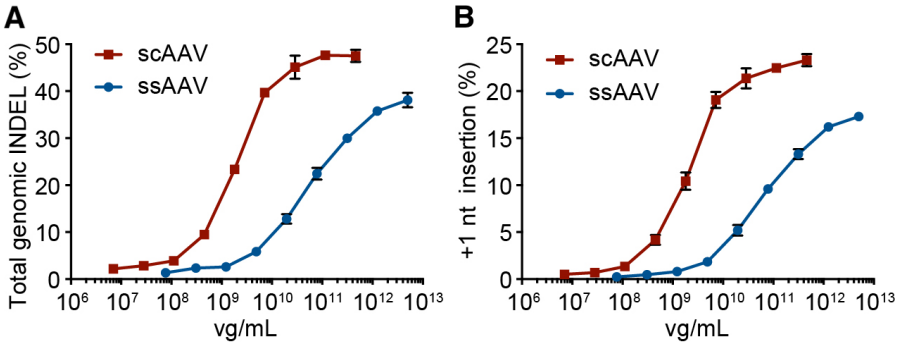

Figure 3.2. In vitro genome editing using ssAAV or scAAV-packaged sgRNA. (**A**) Analysis of total INDEL event in 5-day differentiated myotubes transduced with scAAV or ssAAV-packaged sgRNA at multiple doses. Data are represented as mean ± SEM (n = 3). (**B**) Analysis of +1 nt insertion event in 5-day differentiated myotubes transduced with scAAV or ssAAV-packaged sgRNA at multiple doses. Data are represented as mean ± SEM (n = 3).

Systemic delivery of scAAV-packaged sgRNAs restores dystrophin expression in ΔEx44 mice

To further evaluate the efficacy of the scAAV system by in vivo genome editing, we delivered ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA systemically in ΔEx44 mice through intraperitoneal (IP) injection. The AAV9 serotype was chosen because of its tropism to skeletal muscle and heart (Inagaki et al., 2006). Moreover, *Sp*Cas9 expression was driven by a muscle specific promoter containing key regulatory elements derived from creatine kinase promoter and enhancer, restricting its expression to striated muscles (Himeda et al., 2011). Recent studies demonstrate that AAV-packaged sgRNA is the rate limiting factor for in vivo genome editing in dystrophic mouse models (Hakim et al., 2018; Min et al., 2019). Therefore, we kept ssAAV-packaged *Sp*Cas9 at a constant dose of 8×10^{13} vg/kg while titrating scAAV or ssAAV-packaged sgRNA at multiple doses.

Four weeks after systemic AAV delivery, skeletal muscles and heart of CRISPR/Cas9-edited ΔEx44 mice were harvested for analysis. By immunohistochemistry, we found that dystrophin restoration in skeletal muscles was dose-dependent (Figure 3.3, panel A and Figure 3.4, panel A). Mice receiving the lowest dose of scAAV-packaged sgRNA (4×10^{12} vg/kg) showed 40% and 32% dystrophin-positive myofibers in tibialis anterior (TA) and triceps, respectively; diaphragm and heart showed higher percentages of dystrophin-positive myocytes, reaching 95% (Figure 3.3, panel A and Figure 3.4, panel A). In contrast to the scAAV-treated cohort, ΔEx44 mice receiving lowest dose of ssAAV-packaged sgRNA (4×10^{12} vg/kg) showed less than 5% dystrophin-positive myofibers in TA and triceps; diaphragm and heart showed 52% and 61% dystrophin-positive myocytes, respectively (Figure 3.3, panel B and Figure 3.4, panel B).

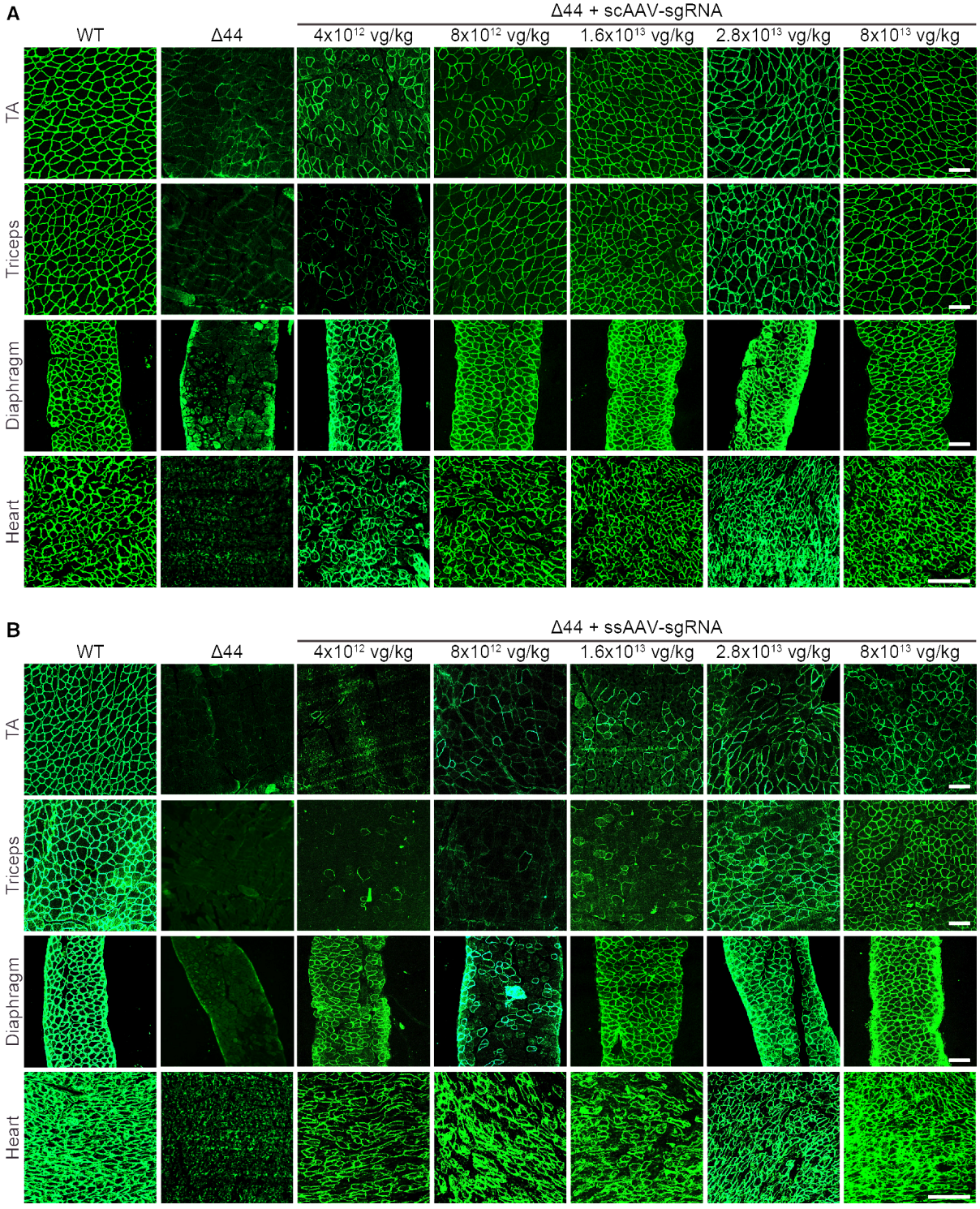
A
Δ44 + scAAV-sgRNA
WT
Δ44
4x10^12 vg/kg
8x10^12 vg/kg
1.6x10^13 vg/kg
2.8x10^13 vg/kg
8x10^13 vg/kg
TA
Triceps
Diaphragm
Heart
B
Δ44 + ssAAV-sgRNA
WT
Δ44
4x10^12 vg/kg
8x10^12 vg/kg
1.6x10^13 vg/kg
2.8x10^13 vg/kg
8x10^13 vg/kg
TA
Triceps
Diaphragm
Heart

Figure 3.3 Systemic AAV delivery of CRISPR/Cas9 genome editing components to ΔEx44 mice rescues dystrophin expression. (**A** and **B**) Immunohistochemistry shows restoration of dystrophin in tibialis anterior (TA), triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV-packaged sgRNA (**A**) or ssAAV-packaged sgRNA (**B**). *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Dystrophin is shown in green. n = 5 for each muscle group. Scale bar, 100 μm.

When the dose of scAAV-packaged sgRNA was increased to 1.6×10^{13} vg/kg, virtually all myofibers and cardiomyocytes were dystrophin-positive (Figure 3.3, panel A and Figure 3.4, panel A). For the ssAAV-treated cohort (1.6×10^{13} vg/kg), diaphragm and heart showed over 75% of dystrophin-positive myocytes; however, dystrophin-positive myofibers in TA and triceps were still below 18% (Figure 3.3, panel B and Figure 3.4, panel B).

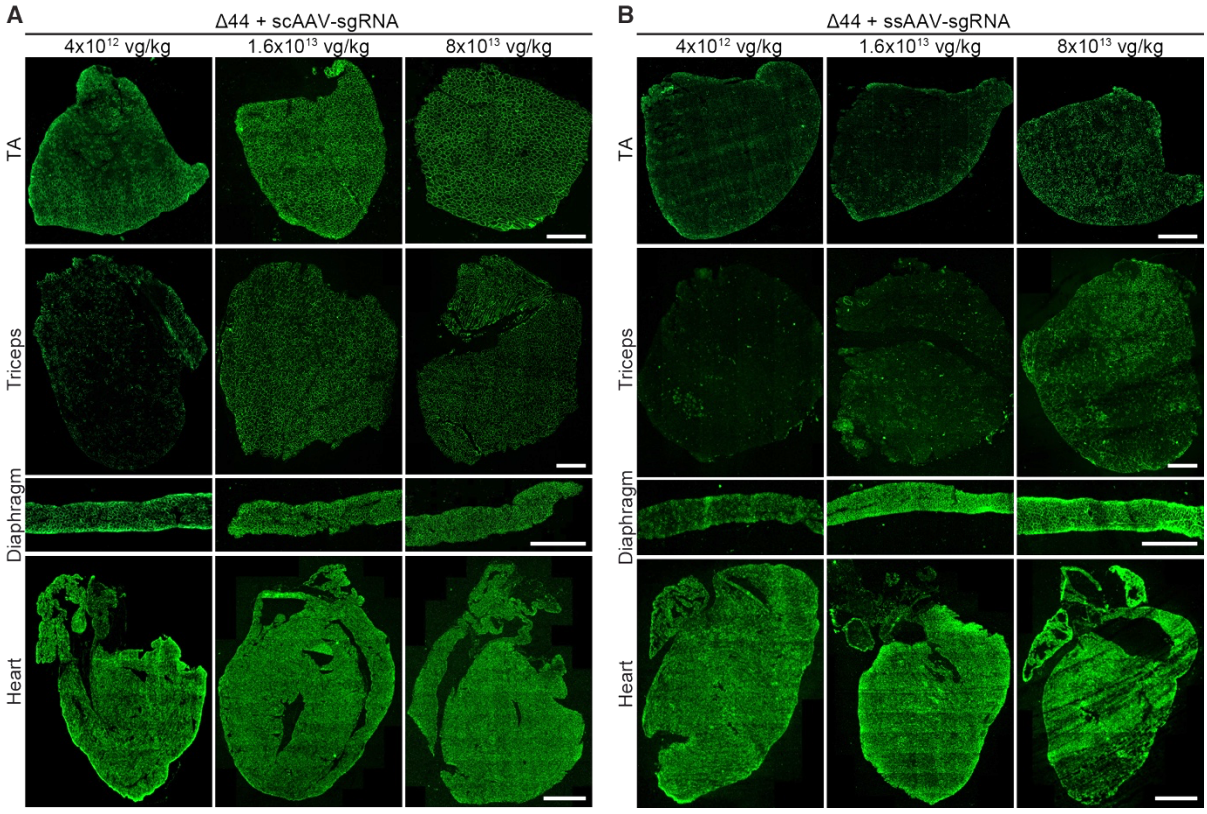

Figure 3.4 Whole muscle scanning of immunohistochemistry of TA, triceps, diaphragm, and heart of CRISPR/Cas9-corrected ΔEx44 mice. (**A** and **B**) Whole muscle scanning of TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV-packaged sgRNA (**A**) or ssAAV-packaged sgRNA (**B**). *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Dystrophin is shown in green. n = 5 for each muscle group. Scale bar in TA, triceps, diaphragm is 500 μm, in heart is 1.5mm.

Next, we performed Western blot analysis to quantitatively detect dystrophin restoration in skeletal muscles and heart after systematic delivery of scAAV or ssAAV-packaged sgRNA. The lowest dose of scAAV-packaged sgRNA (4×10^{12} vg/kg) restored 18%, 14% and 50% of dystrophin protein in TA, triceps and diaphragm, respectively (Figure 3.5, panel A and B). When the dose of scAAV-packaged sgRNA was increased to 1.6×10^{13} vg/kg, dystrophin protein restoration in each skeletal muscle group was greater than 50% (Figure 3.5, panel A and B). Of note, saturation was observed in heart because at every dose of scAAV tested, dystrophin protein restoration exceeded 70% (Figure 3.5, panel A and B). Interestingly, although ssAAV-packaged *Sp*Cas9 was injected at a constant dose (8×10^{13} vg/kg), ΔEx44 mice receiving a higher dose of scAAV-packaged sgRNA showed elevated Cas9 protein expression in skeletal muscles and heart (Figure 3.5, panel A and C).

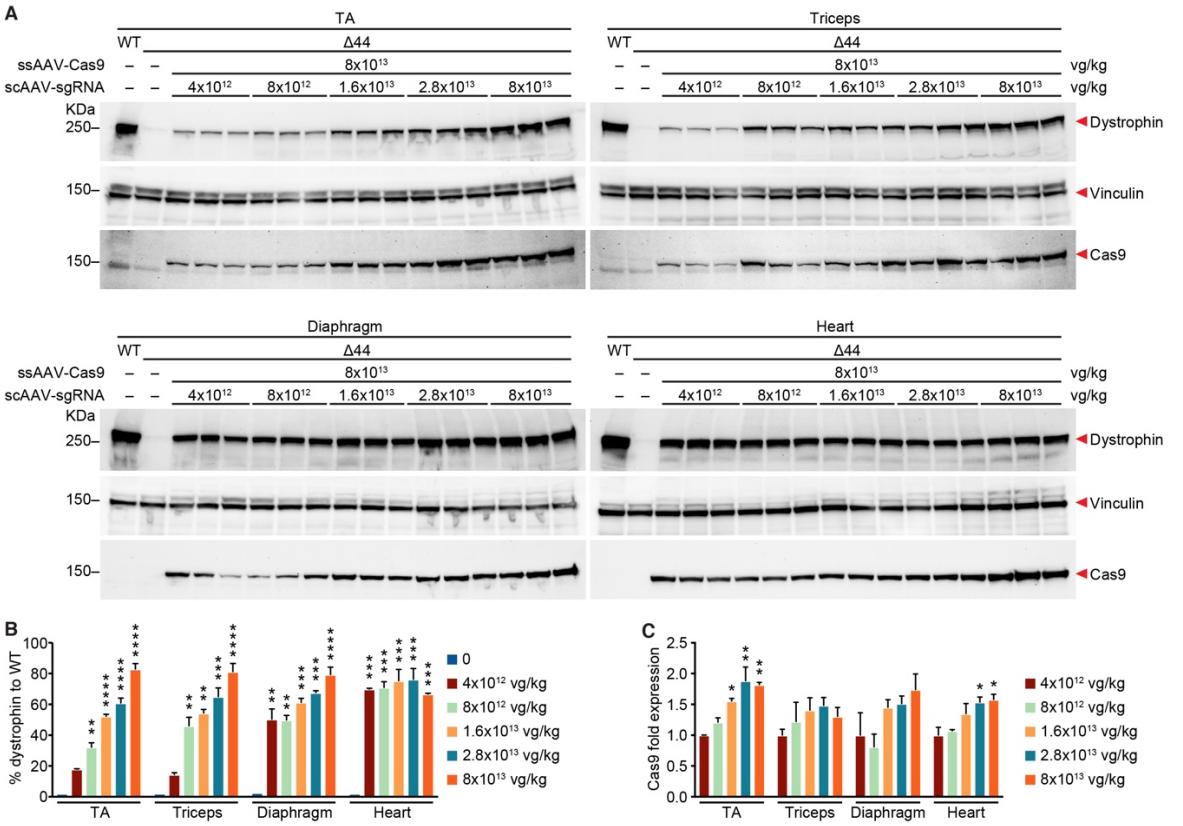

Figure 3.5 Western blot analysis of skeletal muscles and heart of ΔEx44 mice receiving systemic AAV delivery of CRISPR/Cas9 genome editing components. (**A**) Western blot analysis shows restoration of dystrophin expression in the TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Vinculin was used as the loading control (n = 3). (**B**) Quantification of dystrophin expression in TA, triceps, diaphragm, and heart. Relative dystrophin intensity was calibrated with vinculin internal control before normalizing to the WT control. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. **P<0.005, ***P<0.001, ****P<0.0001 (n=3). (**C**) Quantification of Cas9 expression in TA, triceps, diaphragm, and heart. Relative Cas9 intensity was calibrated with vinculin internal control before normalizing to the group treated with lowest dose of scAAV-packaged sgRNA (4×10^{12} vg/kg). Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, **P<0.005 (n=3).

In contrast to the scAAV-treated cohort, ΔEx44 mice receiving ssAAV-packaged sgRNA showed significant lower efficiency in dystrophin restoration by Western blot quantification (Figure 3.6). Specifically, mice receiving the highest dose of ssAAV (8×10^{13} vg/kg) showed only 13%, 16% and 30% of normal dystrophin protein levels in TA, triceps and diaphragm, respectively (Figure 3.6). This was inadequate compared to mice treated with scAAV because more than 80% of dystrophin protein was restored after receiving the same dose of scAAV-packaged sgRNA (Figure 3.5, panel A and B). Therefore, scAAV-expressed sgRNA demonstrated greater efficiency in in vivo genome editing compared to the conventional ssAAV-expressed sgRNA.

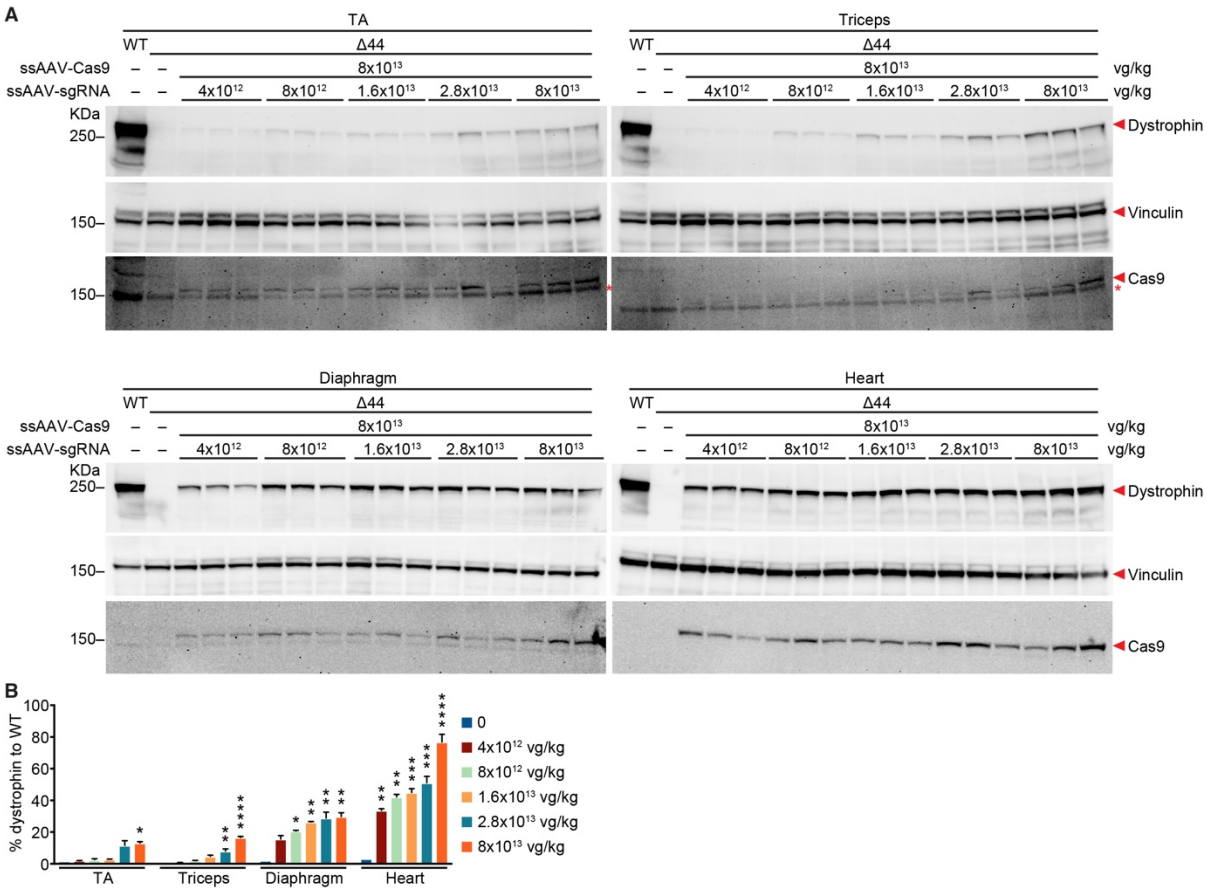

Figure 3.6 Western blot analysis of skeletal muscles and heart of ΔEx44 mice treated with ssAAV-packaged CRISPR/Cas9 genome editing components. (**A**) Western blot analysis shows restoration of dystrophin expression in the TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and ssAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Vinculin was used as the loading control (n = 3). Red asterisks indicates non-specific band. (**B**) Quantification of dystrophin expression in TA, triceps, diaphragm, and heart. Relative dystrophin intensity was calibrated with vinculin internal control before normalizing to the WT control. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, **P<0.005, ***P<0.001, ****P<0.0001 (n=3).

Systemic delivery of scAAV-packaged sgRNAs restores muscle integrity and improves muscle function in ΔEx44 mice

To evaluate whether systemic delivery of scAAV-packaged sgRNAs was able to rescue pathological hallmarks seen in dystrophic mice, we performed hematoxylin and eosin (H&E) staining of skeletal muscles and heart isolated from ΔEx44 mice four weeks after CRISPR/Cas9-mediated genome editing. The percentage of regenerating myofibers with central nuclei declined as the dose of scAAV-packaged sgRNA increased (Figures 3.7 and 3.8). Less than 5% of myofibers showed central nuclei in TA and triceps in mice receiving 1.6×10^{13} vg/kg of scAAV-packaged sgRNA (Figure 3.7, panel A and Figure 3.8, panel A).

In contrast, mice receiving the same dose of ssAAV-packaged sgRNA still showed over 70% of regenerating myofibers with central nuclei, together with signs of muscle necrosis and inflammatory infiltration (Figure 3.7, panel B and Figure 3.8, panel B). In addition, skeletal muscles isolated from mice receiving the highest dose of scAAV-packaged sgRNA (8×10^{13} vg/kg) were virtually indistinguishable from those of wild-type (WT) littermates, whereas the

ssAAV-treated cohort still showed 30% central nuclei in the TA and triceps (Figures 3.7 and 3.8).

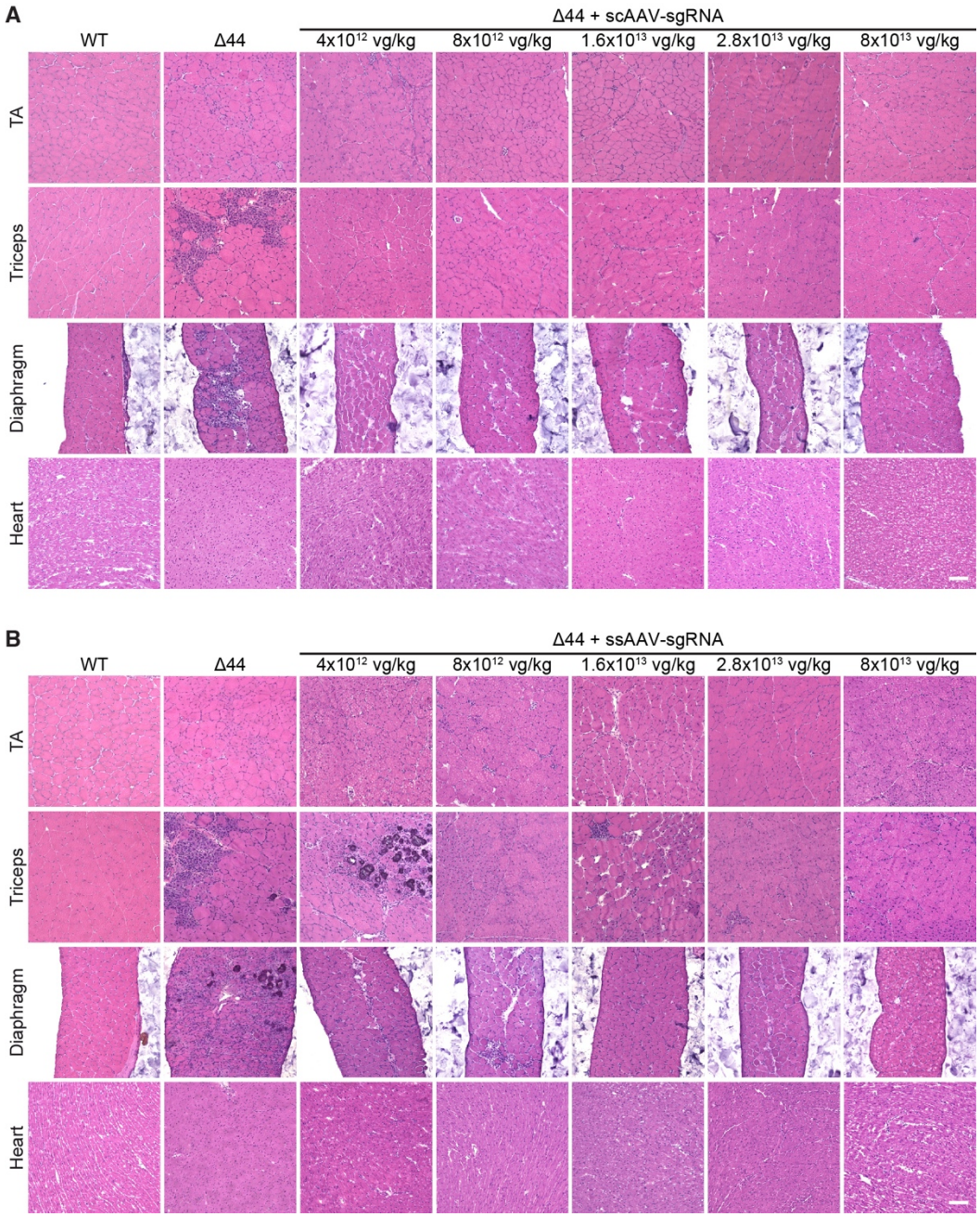

Figure 3.7 Muscle histology of ΔEx44 mice after systemic delivery of AAV expressing CRISPR/Cas9 genome editing components. (**A** and **B**) H&E staining of TA, triceps, diaphragm, and heart of ΔEx44 mice 4

weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV-packaged sgRNA (**A**) or ssAAV-packaged sgRNA (**B**). *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. n = 5 for each muscle group. Scale bar, 100 μm.

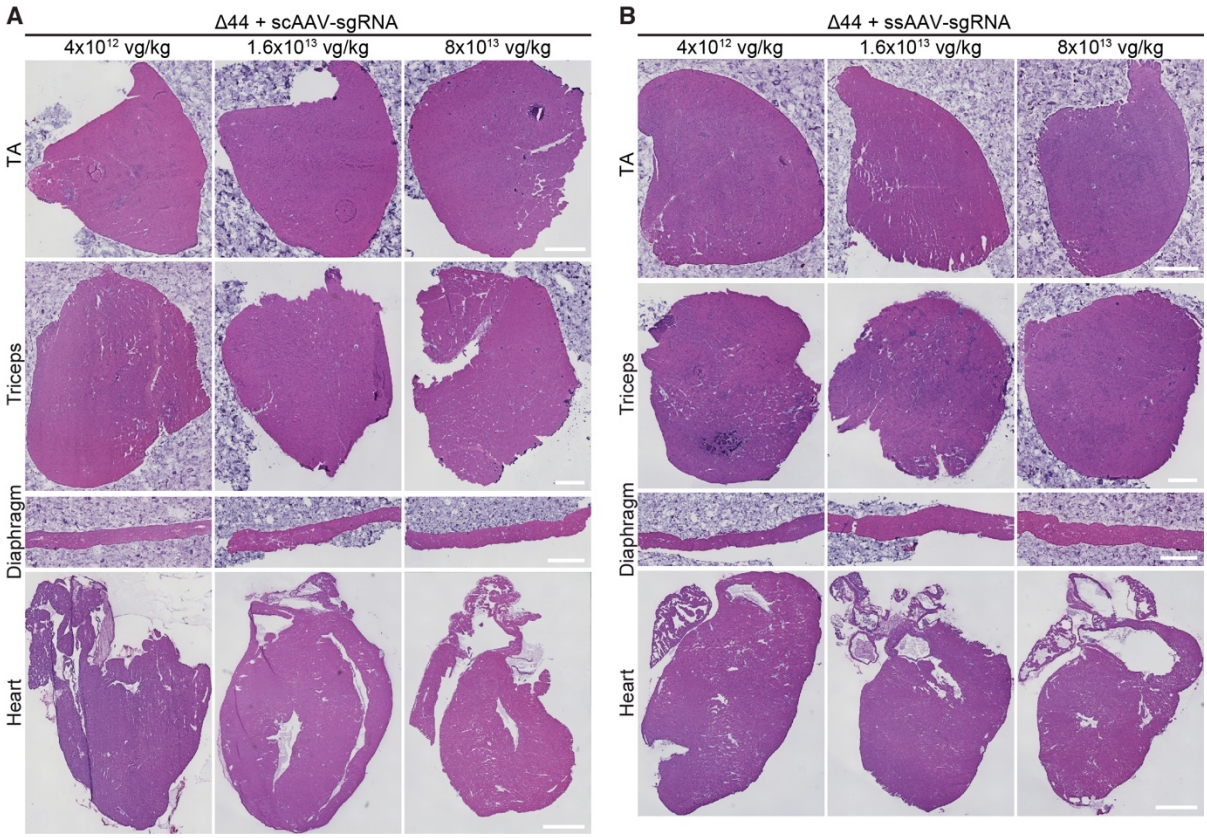

Figure 3.8 Whole muscle scanning of H&E staining of TA, triceps, diaphragm, and heart of CRISPR/Cas9-corrected ΔEx44 mice. (**A** and **B**) Whole muscle scanning of H&E staining of TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV-packaged sgRNA (**A**) or ssAAV-packaged sgRNA (**B**). *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. n = 5 for each muscle group. Scale bar in TA, triceps, diaphragm is 500 μm, in heart is 1.5mm.

To examine the effect of dystrophin restoration on muscle function after systemic delivery of scAAV or ssAAV-packaged sgRNA, we performed electrophysiological analyses on extensor digitorum longus (EDL) and soleus muscles isolated from ΔEx44 mice four weeks after receiving the middle dose of AAV-sgRNA (1.6×10^{13} vg/kg) or the high dose of AAV-sgRNA (8×10^{13} vg/kg). Without CRISPR/Cas9 genome editing, muscle-specific force, which was calibrated by the muscle cross-sectional area, was reduced by 46% in fast-twitch EDL muscle and by 42% in slow-twitch soleus muscle (Figure 3.9, panel A and B). After systemic delivery of scAAV-packaged sgRNA, muscle-specific force of the EDL was increased from 54% to 83% and to 82% for the middle and high doses; in contrast, for the ssAAV-treated cohort, muscle-specific force of the EDL was only increased from 54% to 62% and to 66% for the middle and high doses (Figure 3.9, panel A).

For the slow-twitch soleus muscle, muscle-specific force was increased from 58% to 93% and to 96% after receiving the middle and high doses of scAAV-packaged sgRNA; in contrast, for the ssAAV-treated cohort, only high dose treatment was able to improve muscle-specific force of the soleus to 85%, while no improvement was observed with the middle dose (Figure 3.9, panel B). The maximal tetanic force of the EDL and soleus followed a similar pattern to the muscle-specific force. Specifically, ΔEx44 mice receiving the middle or high doses of scAAV-packaged sgRNA showed improved maximal tetanic force of the EDL muscle to over 80% of WT, whereas the ssAAV-treated cohort was only able to improve to 60% of WT (Figure 3.9, panel C).

The maximal tetanic force of the soleus was improved to over 90% of WT after receiving the middle or high doses of scAAV-packaged sgRNA; high dose of ssAAV-

packaged sgRNA improved maximal tetanic force of the soleus to 85% of WT, while the middle dose did not provide any improvement (Figure 3.9, panel D).

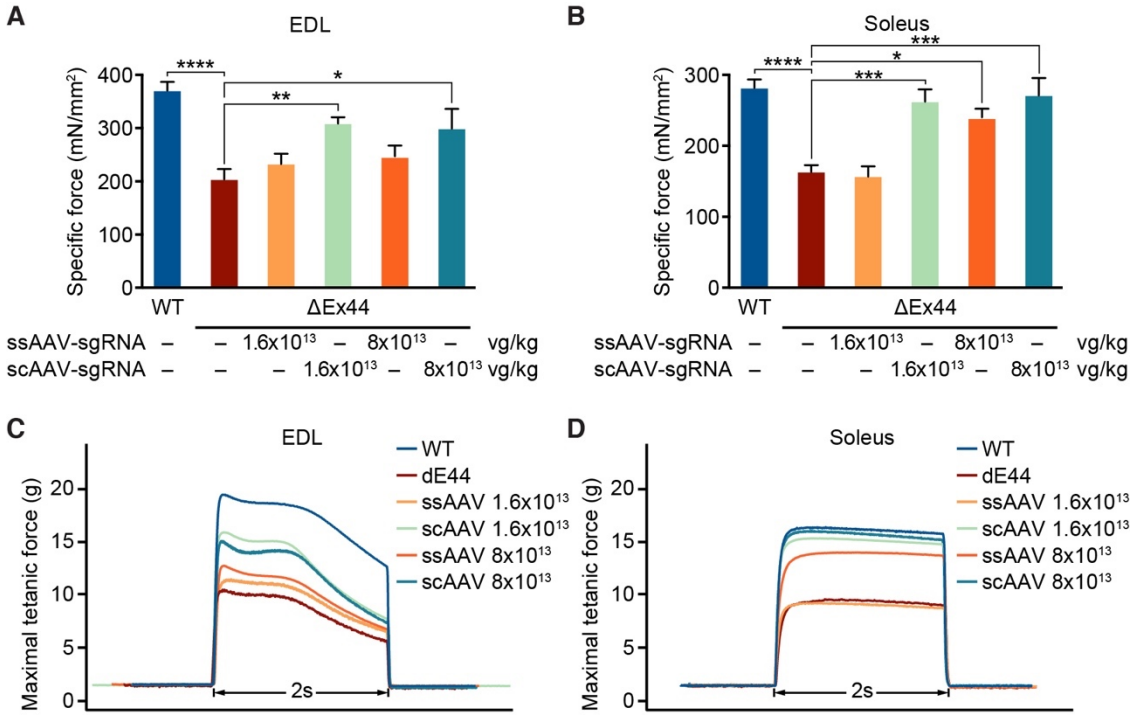

Figure 3.9 Rescue of skeletal muscle function after systemic AAV delivery of CRISPR/Cas9 genome editing components. (**A** and **B**) Specific force (mN/mm^2) of the extensor digitorum longus (EDL) (**A**) and soleus (**B**) muscles in WT, ΔEx44 mice untreated, and ΔEx44 mice treated with ssAAV-packaged SpCas9 and scAAV or ssAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, **P<0.005, ***P<0.001, ****P<0.0001 (n=6). (**C** and **D**) Maximal tetanic force of the EDL (**C**) and soleus (**D**) muscles in WT, ΔEx44 untreated mice, and ΔEx44 mice treated with ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. SpCas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. (n=6).

After receiving the middle and high doses of scAAV-packaged sgRNA, serum creatine kinase (CK) levels in the ΔEx44 mice were reduced by 87% and 95%, respectively, compared with ΔEx44 mice without treatment (Figure 3.10). In contrast, serum CK levels in the ΔEx44 mice receiving the same doses of ssAAV-packaged sgRNA were still 18.6- and 8.5-fold higher, respectively, than the WT littermates (Figure 3.10). These findings indicate that the double-stranded scAAV vector is highly efficient in in vivo gene therapy and can significantly improve muscle integrity and function.

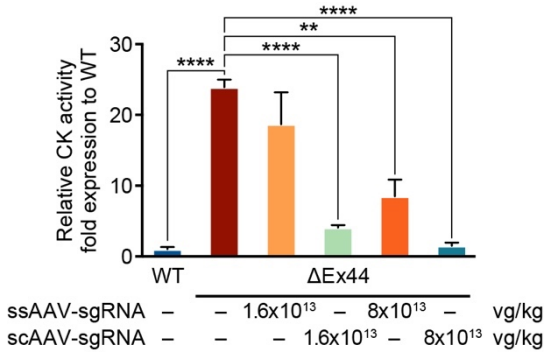

Figure 3.10 Serum creatine kinase (CK) analysis of CRISPR/Cas9-corrected ΔEx44 mice. Serum CK was measured in WT, ΔEx44 mice untreated, and ΔEx44 mice 4 weeks after treatment with ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Serum CK was normalized to WT mice and shown as fold expression. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. **P<0.005, ****P<0.0001 (n=5).

The scAAV system induces significant INDELs within Dmd exon 45 and maintains higher copies of the viral genome in vivo

To determine the mechanism whereby the scAAV system significantly improves in vivo genome editing in ΔEx44 mice, we performed deep sequencing analysis to determine the INDEL frequency at the genomic level (Figure 3.11 and Table 3.1). The percentage of total genomic INDELs and +1 nt insertions at exon 45 correlated with ascending doses of AAV-sgRNA. ΔEx44 mice receiving the high dose (8×10^{13} vg/kg) of scAAV-packaged sgRNA showed more than 28% and 30% of total NHEJ events in TA and triceps (Figure 3.11). Of note, over 60% of total NHEJ events were +1 nt insertions, which restores the *Dmd* exon 45 ORF. In contrast, TA and triceps from ΔEx44 mice receiving the same dose of ssAAV-packaged sgRNA had only 10% and 11% of total NHEJ events (Figure 3.11). We did not observe a significant difference between scAAV and ssAAV in inducing total NHEJ and +1 nt insertion events in the diaphragm or heart (Figure 3.11). We observed a low percentage of AAV integration events at the sgRNA targeting site in ssAAV- and scAAV-treated mice (Table 3.1).

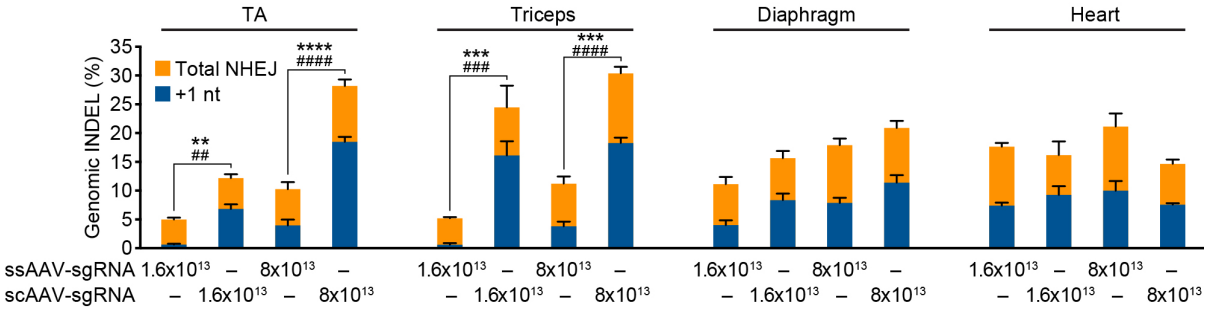

Figure 3.11 scAAV vector induces significant INDELs at genomic level. Genomic INDEL analysis by deep sequencing of TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. SpCas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Data are represented as mean ± SEM. Two-way

ANOVA was performed with post-hoc Tukey's multiple comparisons test. **P<0.005, ***P<0.001, ****P<0.0001 for total NHEJ event (n=3). ##P<0.005, ###P<0.001, ####P<0.0001 for +1 nt insertion event (n=3).

Table 3.1 Genomic INDEL analysis by deep sequencing of TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after AAV-CRISPR/Cas9 genome editing.

Dose (vg/kg)	Treatment	Analyzed Muscle	Samples	Barcode	Total Reads	Non-edited %	Total NHEJ %	NHEJ (+1 nt) %	ITR Integration %
1.60E+13	ssAAV-sgRNA	TA	BC2-ss58	TGGTCA	265,313	94.59	5.41	0.85	<0.2
			BC3-ss59	CACTGT	430,500	94.77	5.23	0.69	<0.2
			BC4-ss60	ATTGGC	343,805	95.69	4.31	0.34	<0.2
		Triceps	BC10-ss58	AGCTAG	336,778	94.97	5.03	0.25	<0.2
			BC11-ss59	GTCGTC	339,549	94.39	5.61	1.18	<0.2
			BC12-ss60	CGATTA	322,347	95.02	4.98	0.32	<0.2
		Diaphragm	BC18-ss58	ATCAGT	423,848	86.51	13.49	5.64	<0.2
			BC19-ss59	TATACT	508,066	89.32	10.68	3.65	<0.2
			BC20-ss60	CAACAA	297,096	90.79	9.21	2.77	<0.2
		Heart	BC26-ss58	AATGGT	414,968	81.29	18.71	7.84	<0.2
			BC27-ss59	GGCGGT	314,206	83.56	16.44	6.36	<0.2
			BC28-ss60	ATCACG	444,263	82.25	17.75	7.99	<0.2
	scAAV-sgRNA	TA	BC2-sc65	TGGTCA	351,570	87.11	12.89	8.29	<0.2
			BC3-sc66	CACTGT	397,695	89.08	10.92	5.47	<0.2
			BC4-sc67	ATTGGC	347,505	87.22	12.78	6.71	0.28
		Triceps	BC10-sc65	AGCTAG	445,116	73.92	26.08	18.96	0.47
			BC11-sc66	GTCGTC	391,352	82.71	17.29	11.15	0.39
			BC12-sc67	CGATTA	413,467	69.96	30.04	18.21	<0.2
		Diaphragm	BC18-sc65	ATCAGT	445,539	86.61	13.39	6.32	<0.2
			BC19-sc66	TATACT	404,070	82.33	17.67	10.3	<0.2
			BC20-sc67	CAACAA	331,964	84.09	15.91	8.41	<0.2
		Heart	BC26-sc65	AATGGT	471,201	79.04	20.96	12.3	0.35
			BC27-sc66	GGCGGT	328,191	86.48	13.52	7.75	<0.2
			BC28-sc67	ATCACG	435,717	85.92	14.08	7.77	0.48
8.00E+13	ssAAV-sgRNA	TA	BC5-ss83	GATCTG	418,020	91.34	8.66	2.67	0.24
			BC6-ss84	TACAAG	337,358	87.34	12.66	5.99	0.23
			BC7-ss85	CGTGAT	291,510	90.57	9.43	3.16	<0.2
		Triceps	BC13-ss83	GAATGA	348,423	90.21	9.79	3.53	<0.2
			BC14-ss84	CTTCGA	394,987	86.25	13.75	5.32	<0.2
			BC15-ss85	CTCTAC	245,385	89.91	10.09	2.55	<0.2
		Diaphragm	BC21-ss83	GTTGTT	339,005	84.47	15.53	6.11	<0.2
			BC22-ss84	TCGGTT	335,216	80.78	19.22	8.98	<0.2
			BC23-ss85	AGTATT	405,971	81.09	18.91	8.51	<0.2
		Heart	BC29-ss83	CGATGT	349,844	81.33	18.67	8.33	<0.2
			BC30-ss84	TTAGGC	329,313	80.97	19.03	8.32	0.23
			BC31-ss85	TGACCA	472,649	74.32	25.68	13.36	<0.2
	scAAV-sgRNA	TA	BC5-sc14	GATCTG	294,064	70.91	29.09	19.76	<0.2
			BC6-sc15	TACAAG	356,949	74.01	25.99	16.8	0.24
			BC7-sc16	CGTGAT	368,727	70.49	29.51	18.87	<0.2
		Triceps	BC13-sc14	GAATGA	380,792	71.12	28.88	16.95	0.24
			BC14-sc15	CTTCGA	339,794	70.4	29.6	17.8	0.7
			BC15-sc16	CTCTAC	365,034	67.38	32.62	20.08	<0.2
		Diaphragm	BC21-sc14	GTTGTT	407,519	77.15	22.85	13.67	<0.2
			BC22-sc15	TCGGTT	393,010	78.78	21.22	11.34	<0.2
			BC23-sc16	AGTATT	443,769	81.34	18.66	9.12	<0.2
		Heart	BC29-sc14	CGATGT	420,871	85.18	14.82	7.7	<0.2
			BC30-sc15	TTAGGC	415,579	84.14	15.86	7.04	0.64
			BC31-sc16	TGACCA	411,603	86.73	13.27	7.93	0.97

We also performed TIDE analysis on dystrophin cDNA transcripts isolated from skeletal muscles and heart. Total cDNA INDEL rate and +1 nt insertion events at exon 45 followed similar ascending patterns seen in the genomic TIDE analysis while the absolute percentage increased significantly (Figure 3.12), indicating enrichment of the reframed cDNA transcript after nonsense-mediated decay of unedited transcript with a premature stop codon in exon 45. These findings indicate that the scAAV system is highly efficient in inducing INDELs at the targeted genomic locus, and the majority of the INDEL events contain +1 nt insertions, which is able to repair the out-of-frame mutation in *Dmd* exon 45.

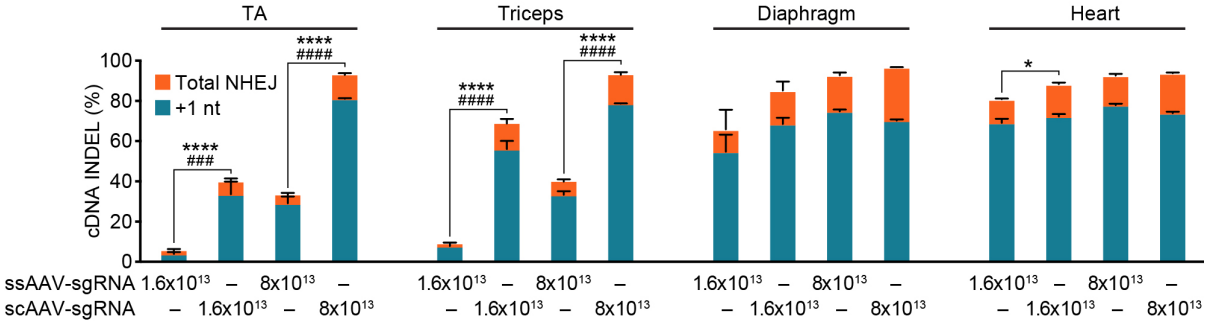

Figure 3.12 scAAV vector induces significant INDELs at cDNA level. Dystrophin cDNA INDEL analysis by TIDE analysis of TA, triceps, diaphragm, and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. SpCas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Data are represented as mean ± SEM. Two-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, ****P<0.0001 for total NHEJ event (n=3). ###P<0.001, ####P<0.0001 for +1 nt insertion event (n=3).

Next, we performed quantitative PCR analysis to detect viral genome copies in skeletal muscles and heart of ΔEx44 mice four weeks after systemic delivery of AAV-CRISPR/Cas9 genome editing components. Mice receiving scAAV treatment showed significantly higher copy numbers of sgRNA viral genomes than those receiving the same dose of ssAAV-packaged sgRNA (Figure 3.13, panel A). Moreover, the sgRNA transcripts transcribed from the scAAV vector were also significantly higher than those transcribed from the ssAAV vector (Figure 3.13, panel B). These findings indicate that there is a significant depletion of ssAAV-packaged sgRNA vector in skeletal muscles in vivo.

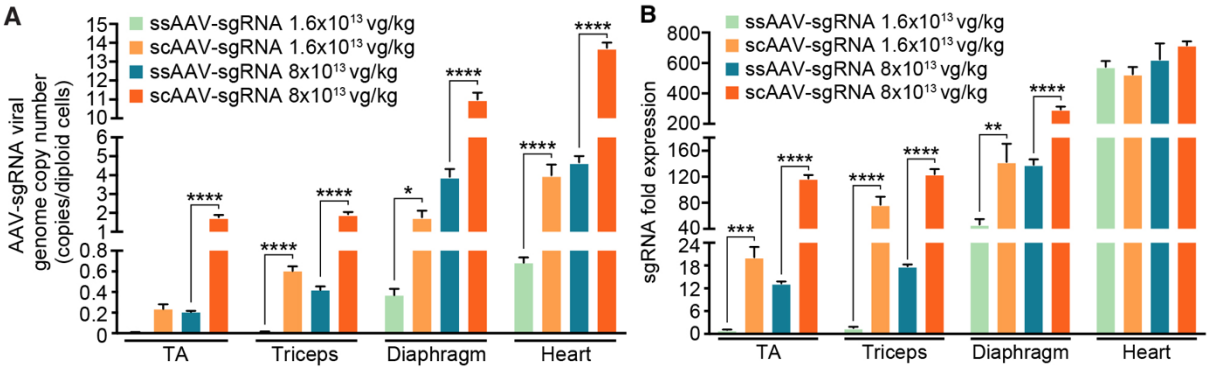

Figure 3.13. ΔEx44 mice sustain higher copies of scAAV-sgRNA viral genome and express more sgRNA transcripts after systemic delivery of scAAV-packaged sgRNA. (**A** and **B**) sgRNA viral genome copy number (**A**) and sgRNA transcript fold expression (**B**) from skeletal muscles and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, **P<0.005, ***P<0.001, ****P<0.0001 (n=3).

Interestingly, although the dose of Cas9 vector was kept constant at 8×10^{13} vg/kg during initial systemic injection, the viral genomes of Cas9 vector in TA and triceps persisted with higher copies from mice receiving scAAV-packaged sgRNA vector than those receiving same dose of ssAAV-packaged sgRNA vector (Figure 3.14, panel A). However, AAV-Cas9 viral genomes showed relatively high copies in diaphragm and heart independent of the identity of AAV-sgRNA vector (Figure 3.14, panel A). These findings are consistent with Cas9 cDNA transcript analysis (Figure 3.14, panel B). Together, these data suggest that the high efficiency of scAAV-mediated in vivo genome editing is attributed to higher viral genome persistence of the sgRNA vector and Cas9 vector.

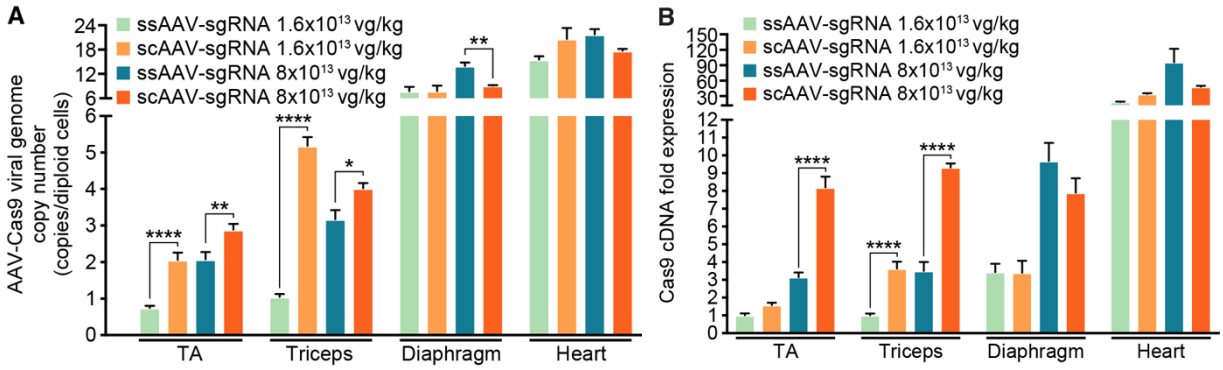

Figure 3.14. ΔEx44 mice sustain higher copies of ssAAV-*Sp*Cas9 viral genome and express more *Sp*Cas9 transcripts after systemic delivery of scAAV-packaged sgRNA. (**A** and **B**) *Sp*Cas9 viral genome copy number (**A**) and *Sp*Cas9 transcript fold expression (**B**) from skeletal muscles and heart of ΔEx44 mice 4 weeks after systemic delivery of ssAAV-packaged *Sp*Cas9 and scAAV or ssAAV-packaged sgRNA. *Sp*Cas9 vector was kept at constant dose of 8×10^{13} vg/kg. The dose of sgRNA vector was shown in the figure. Data are represented as mean ± SEM. One-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. *P<0.05, **P<0.005, ****P<0.0001 (n=3).

Discussion

Owing to non-pathogenic and low-immunogenic characteristics, recombinant AAV has been chosen as a delivery vector for multiple gene therapy clinical trials and three have been approved for treating lipoprotein lipase deficiency, inherited retinal dystrophy, and spinal muscular atrophy (Dunbar et al., 2018). In this study, we developed a new genome editing strategy in which the Cas9 nuclease is encoded by conventional ssAAV while sgRNAs are expressed by double-stranded scAAV. After a single high dose systemic injection of this dual AAV system into ΔEx44 mice (8×10^{13} vg/kg of Cas9 vector, 8×10^{13} vg/kg of sgRNA vector), dystrophin protein expression in multiple muscle groups was restored by ~80% and skeletal muscle function was improved by ~82% in fast-twitch EDL muscle and by ~96% in slow-twitch soleus muscle. Importantly, a low dose of scAAV-expressed sgRNAs (4×10^{12} vg/kg) is sufficient to restore 18%, 14% and 50% of dystrophin protein in TA, triceps and diaphragm, respectively, representing a 20-fold improvement in efficiency compared with the ssAAV-packaged sgRNA vector. Several potential explanations may account for these observations. First, it has been widely accepted that most recombinant AAV genomes persist as double-stranded episomes in vivo, either in the form of circular or linear concatemers (Duan et al., 1998; Miao et al., 1998; Nakai et al., 2001). During the concatemerization process, the double-stranded DNA intermediate is an indispensable prerequisite. Thus, the scAAV undergoes concatemerization more rapidly than ssAAV because scAAV-based concatemerization bypasses second-strand synthesis, which is a rate-limiting step for ssAAV (Ferrari et al., 1996; Fisher et al., 1996). Second, it has been reported that monomeric viral genome degradation is significantly slower in scAAV transduced skeletal muscle compared with ssAAV (Ren et al.,

2005). Therefore, scAAV is more stable than ssAAV during initial viral transduction, leading to higher episomal persistence in the long-term. Third, DNA DSBs in post-mitotic cells are repaired by the classical NHEJ pathway, which requires the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) (Ciccia and Elledge, 2010). It is known that DNA-PKcs is required for AAV viral genome concatemerization in AAV transduced skeletal muscle (Duan et al., 2003). In this study, we found that scAAV-packaged sgRNA leads to a higher incidence of DNA DSB at the target site. This may induce higher DNA-PKcs expression, which in turn facilitates AAV concatemerization and long-term gene expression. Indeed, we found higher viral genome persistence of both sgRNA vector and Cas9 vector in mice treated with scAAV-packaged sgRNA. In summary, the scAAV-sgRNA delivery system has many appealing features, including stable persistence of AAV viral genomes, higher INDEL frequency at the targeted genomic locus, and highly efficient genome editing in vitro and in vivo at low viral dose.

Initial studies of Cas9-induced DNA DSBs suggested that the breakage point was blunt-ended (Gasiunas et al., 2012; Jinek et al., 2012). However, molecular dynamics simulations of the *Sp*Cas9-sgRNA-dsDNA system suggest that *Sp*Cas9-induced cleavage generates a staggered cut, producing a single nucleotide 5' overhang at the breakage point, which is prone to be filled with one additional nucleotide by the DNA polymerase, leading to a high frequency of +1 nt insertion after NHEJ-mediated repair (Lemos et al., 2018; Zuo and Liu, 2016). Based on this mechanism, our lab developed CRISPR/Cas9-mediated "single-cut" technology and successfully restored the ORF of exon 51 in mice and dogs with exon 50 deletion, and exon 45 ORF in mice lacking exon 44 (Amoasii et al., 2018; Amoasii et al., 2017;

Min et al., 2019). This technology, in theory, can be applied to correct diverse mutations in any exon containing a single nucleotide out-of-frame deletion (Long et al., 2018). However, there remains another group of mutations in which the frameshift mutation needs to be reframed by 3n+2 insertions or 3n-1 deletion. In such cases, the CRISPR/Cas9-mediated "single-cut" repair strategy may be less efficient. Several studies have shown that it is possible to restore the *Dmd* ORF by removing one or multiple exons by using two sgRNAs (Bengtsson et al., 2017; Hakim et al., 2018; Long et al., 2016; Nelson et al., 2016; Nelson et al., 2019; Tabebordbar et al., 2016). However, this "double-cut" strategy is only effective when two cooperative DNA DSBs occur simultaneously. If the first DNA DSB is rapidly rejoined by NHEJ-mediated repair, the second DSB alone is not sufficient to excise the entire exon. Moreover, a high frequency of AAV ITR integration events is observed at the Cas9 target site when two sgRNAs are used to excise large genomic intervening regions (Maeder et al., 2019; Nelson et al., 2019). Therefore, the CRISPR/Cas9-mediated "single-cut" repair strategy has unique advantages, including predictable DNA repair outcome, minimum genomic modification at a precise location, and low frequency of off-target effects.

Nevertheless, many questions remain unaddressed for AAV-delivered CRISPR/Cas9-mediated therapeutic genome editing. The first concern is the potential for adaptive immune responses elicited by AAV capsid protein and Cas9 nuclease. It has been reported that AAV neutralizing antibodies, anti-*Sp*Cas9 antibodies, and *Sp*Cas9-specific T cells are found in the human population (Boutin et al., 2010; Calcedo et al., 2009; Charlesworth et al., 2019; Chew et al., 2016). While no apparent immune response to AAV or Cas9 was observed in neonatal mice, it remains unclear whether this observation also will also apply to humans. Potential

solutions to address these concerns include (i) large-scale functional variant profiling of AAV and *Sp*Cas9 for epitope mutagenesis to block antibody binding; (ii) plasmapheresis to reduce neutralizing antibody titer; and (iii) transient immunosuppression (Zhang et al., 2018). The second concern is durability of CRISPR/Cas9-mediated therapeutic genome editing. Our TIDE analysis in multiple skeletal muscle samples showed an average of 21% editing events at the genomic level, among which, 75% of editing events are effective and able to restore the *Dmd* ORF. Moreover, recent studies in the *mdx* mouse model of DMD showed that sustained genome editing and dystrophin expression can be observed for 12-18 months after single intravenous injection of AAV9 encoded *Sa*Cas9 (Cas9 derived from *Staphylococcus aureus*) (Hakim et al., 2018; Nelson et al., 2019). However, skeletal muscle contains stem cells capable of undergoing *de novo* myogenesis and contributing to pre-existing myofibers (Yin et al., 2013). Whether these events will gradually dilute out genome-edited, dystrophin-positive myofibers in the long-term remains to be determined. In this study, we were able to restore more than 70% of dystrophin expression in the heart. In addition, adult human cardiomyocytes have a very low turnover rate (1% at the age of 25 to 0.45% at the age of 75) (Bergmann et al., 2009; Bergmann et al., 2015). Thus, in theory, long-term clinical benefit by CRISPR/Cas9-mediated genome editing should be sustained in the human heart.

In summary, a low dose of scAAV-delivered CRISPR-Cas genome editing components is sufficient to restore dystrophin protein expression, reduce DMD pathological phenotypes, and improve muscle function in a DMD mouse model. Therefore, this robust scAAV delivery system combined with the efficient CRISPR-Cas9 genome editing technology represents a

promising therapy for permanent correction of diverse genetic mutations in neuromuscular diseases.

Materials and Methods

Study design

This study was designed with the primary aim to investigate the feasibility of using the scAAV system to deliver the CRISPR/Cas9 genome editing components for the correction of DMD mutations. The secondary objective was to compare the efficiency between the conventional ssAAV and scAAV in delivering CRISPR sgRNA for in vivo therapeutic genome editing. We did not use exclusion, randomization, or blind approaches to assign the animals for the experiments. For each experiment, sample size reflects the number of independent biological replicates and was provided in the figure legends.

AAV vector cloning and viral production

The sgRNA targeting mouse *Dmd* exon 45, listed in Table 3.2, was first cloned into TRISPR-sgRNA-CK8e-GFP plasmid, a modified gift from D. Grimm, using Golden Gate Assembly (New England Biolabs). A detailed cloning protocol was previously described (Amoasii et al., 2017). The sgRNA expression cassette containing three copies of same sgRNA driven by the U6, H1, and 7SK promoter was PCR amplified and subcloned into the pSJG self-complementary AAV plasmid (scAAV plasmid), a gift from S. Gray, or into the pSSV9 single-stranded AAV plasmid (ssAAV plasmid), using In-Fusion Cloning Kit (Takara Bio). A 2.3 kb

stuffer sequence was cloned into the ssAAV plasmid for optimal viral packaging. Both the scAAV and the ssAAV genome contain the same sgRNA expression cassette, consisting of three copies of sgRNA sequence driven by three RNA polymerase III promoters. Cloning primers are listed in Table 3.2. All AAV viral plasmids were column purified and digested with Sma I and Ahd I to check ITR integrity. AAVs were packaged by Boston Children's Hospital Viral Core and serotype 9 was chosen for capsid assembly. AAV titers were determined by Droplet Digital PCR (ddCPR) (Bio-Rad Laboratories) according to the manufacturer's protocol. Primers and probes used for titration were listed in Table 3.2.

Alkaline agarose gel electrophoresis

AAV virus (2 x 10^{11} vg) was equalized with water to 13 μl and digested with 10 μl DNase solution (10mM Tris-HCl, pH 7.5, 10mM $CaCl_2$, 10mM $MgCl_2$, 0.1mg/ml DNase) at 37 °C for 1 hour, followed by chelating Mg^{2+} and Ca^{2+} with 5 μl 0.5M EDTA. Then the capsid was denatured by adding 2 μl 10% SDS. The reaction mixture was mixed with 6 ul 6X alkaline agarose gel loading dye (Alfa Aesar) and loaded into 1% alkaline agarose gel. Denaturing gel electrophoresis was performed in a cold room at 50 V for 15 hours. The gel was neutralized with neutralization buffer (0.5 M Tris-HCl pH 7.5, 1M NaCl) and stained with SYBR Gold (Thermo Fisher Scientific) for visualization.

in vitro AAV viral transduction in C2C12 myotubes

Cas9-expressing C2C12 myoblasts were cultured in 96-well dishes with growth medium (DMEM with 10% FBS) until reaching 90% confluency. Then the myoblasts were allowed to

differentiate in myotube in differentiation medium (DMEM with 2% horse serum) for 5 days. Two-hours before viral transduction, myotubes were treated with *Vibrio cholerae* Neuraminidase (50 mu/ml) (Sigma-Aldrich), followed by washing with differentiation medium twice (Shen et al., 2011). Myotubes were incubated with varying doses of scAAV or ssAAV and centrifuged at 1,000 xg at 4 ^{o}C for 1.5 hour. After spin transduction, the virus was aspirated and the myotubes were washed with differentiation medium three times. The myotubes were cultured in differentiation medium for an additional week prior genomic DNA isolation for INDEL analysis.

in vivo AAV delivery into ΔEx44 mice

Postnatal day 4 (P4) ΔEx44 mice were injected intraperitoneally with 80 μl of AAV9 viral mixture containing 8×10^{13} vg/kg AAV9-*Sp*Cas9 and varying doses of scAAV or ssAAV-packaged sgRNA using an ultrafine BD insulin syringe (Becton Dickinson). The doses of scAAV or ssAAV-packaged sgRNA are indicated in the figure legends. Four weeks after systemic delivery, ΔEx44 mice and WT littermates were dissected for physiological, biochemical and histological analysis. Animal work described in this manuscript has been approved and conducted under the oversight of the UT Southwestern Institutional Animal Care and Use Committee.

Genomic DNA and RNA isolation, cDNA synthesis, and PCR amplification

Genomic DNA of mouse C2C12 myotubes, skeletal muscles and heart was isolated using DirectPCR (cell) lysis reagent (Viagen Biotech) according to the manufacturer's protocol.

Total RNA of skeletal muscles and heart was isolated using miRNeasy (QIAGEN) according to the manufacturer's protocol. cDNA was reverse-transcribed from total RNA using SuperScript III First-Strand Synthesis SuperMix (Thermo Fisher Scientific) according to the manufacturer's protocol. PCR amplification was performed as previously described (Min et al., 2019). Primer sequences were listed in Table 3.2.

INDELs analysis of genomic DNA and cDNA

INDELs in genomic DNA and cDNA were analyzed using Tracking of INDELs by Decomposition (TIDE) software package (https://tide.deskgen.com). Briefly, the sgRNA sequence targeting mouse *Dmd* exon 45 was first uploaded to the software to define *Sp*Cas9-mediated DSB site. Then, the CRISPR/Cas9-edited sequence and non-edited control sequence were uploaded and aligned using Smith-Waterman local alignment algorithm. The percentage of INDELs was calculated based on the relative abundance of aberrant nucleotides over the length of the whole sequence trace.

Amplicon deep sequencing analysis

PCR of genomic DNA was performed using primers designed against the *Dmd* exon 45 region. A second round of PCR was performed to add Illumina flow cell binding sequence and barcodes. All primer sequences are listed in Table 3.2. Deep sequencing and data analysis were performed as previously described (Min et al., 2019).

AAV viral genome copy number quantification

The AAV viral genome copy number was determined by quantitative PCR using custom-designed primers (Table 3.2). The primer sets used in AAV-sgRNA and AAV-Cas9 viral genome quantification anneal to the 7SK promoter and Cas9 gene, respectively. The threshold cycle value of each reaction was converted to the viral genome copy number by measuring against the copy number standard curve of the AAV plasmids used for AAV packaging in this study. Mouse 18S ribosomal RNA gene was used as the reference gene to calibrate genomic DNA quantity.

Dystrophin and SpCas9 Western blot analysis

Heart and skeletal muscles were crushed and lysed with lysis buffer [10% SDS, 62.5 mM tris (pH 6.8), 1 mM EDTA, and protease inhibitor]. Total 50 μg of protein was loaded onto 4-20% Criterion™ TGX™ Precast Midi Protein Gel (Bio-Rad Laboratories). Details of Western blot running, transferring, and developing were previously described (Min et al., 2019). Primary antibodies used in Western blot were mouse anti-dystrophin antibody (MANDYS8, Sigma-Aldrich, D8168), mouse anti-Cas9 antibody (Clone 7A9, Millipore, MAC133), mouse anti-vinculin antibody (Sigma-Aldrich, V9131). Secondary antibodies used in Western blot were goat anti-mouse horseradish peroxidase (HRP) antibody or goat anti-rabbit HRP antibody (Bio-Rad Laboratories).

Histological analysis of skeletal muscle and heart

Skeletal muscles and heart were cryosectioned into eight-micron transverse sections. Immunohistochemistry was performed as previously described (Min et al., 2019). Antibodies used in immunohistochemistry were mouse anti-dystrophin antibody (MANDYS8, Sigma-Aldrich, D8168) and M.O.M. biotinylated anti-mouse IgG (BMK-2202, Vector Laboratories).

Electrophysiological analysis of isolated EDL and soleus muscles

Four weeks after systemic AAV-CRISPR/Cas9 genome editing, EDL and soleus muscles from ΔEx44 mice and WT littermates were isolated for electrophysiological analysis, as previously described (Min et al., 2019). Specific force (mN/mm^2) was calculated by normalizing contraction force to muscle cross-sectional area.

Statistics

All data are shown as means ± SEM. One-way ANOVA or two-way ANOVA was performed with post-hoc Tukey's multiple comparisons test. A $P < 0.05$ value was considered statistically significant.

Table 3.2 Primers used in this study.

Primer function	Primer name	Primer sequence
ssAAV-sgRNA cloning primers	IF-SpeI-PacI-U6-F	ATGTATACCTTCGTCACTAGTTTAATTAAATCAGCGTTTGAGTAAGAGCCCG
	IF-XbaI-KpnI-GFP-R	GCTGGCGCGCCTTTTTCTAGAGGTACCCATGGACGAGCTGTACAAGTAAAGGC
scAAV-sgRNA cloning primers	IF-SpeI-KpnI-U6-F	ATGTATACCTTCGTCACTAGTGGTACCATCAGCGTTTGAGTAAGAGCCCG
	IF-XbaI-MluI-GFP-R	GCTGGCGCGCCTTTTTCTAGAACGCGTCATGGACGAGCTGTACAAGTAAAGG
AAV titer ddPCR primers	AAV-Cas9-ddPCR-F	GGACTCCCGGATGAACACTAAG
	AAV-Cas9-ddPCR-R	GTTGTTGATCTCGCGCACTTT
	AAV-Cas9-ddPCR-probe	TGGTGTCCGATTTCCGGA
	AAV-sgRNA-ddPCR-F	AGGGCCTATTTCCCATGATT
	AAV-sgRNA-ddPCR-R	AAACTGCAAACTACCCAAGAAA
	AAV-sgRNA-ddPCR-probe	TGCATATACGATACAAGGCTGTTAGAGAGA
Genomic PCR for T7E1 analysis	T7E1-mE45-g4-F	CCCTGAGCTGAAGTGAGAGG
	T7E1-mE45-g4-R	ACCTCTTTCTCCTTTCTGCCAG
AAV vector genome copy number quantification	AAV-Cas9-copy-F	TGAAAGAGGACTACTTCAAGAAAATC
	AAV-Cas9-copy-R	TTGTCCTTGATAATTTTCAGCAGATC
	AAV-sgRNA-copy-F	CCGGAAATCAAGTCCGTTTATC
	AAV-sgRNA-copy-R	GAGCTTCAGCAGGAAATTTAACTA
cDNA RT-PCR quantification	mE43-qPCR-F	AGGTGAAAGTACAGGAAGCCGT
	mE46-qPCR-R	CTGCTGCTCATCTCCAAGTGGA
	Cas9-qPCR-F	TGAAAGAGGACTACTTCAAGAAAATC
	Cas9-qPCR-R	TTGTCCTTGATAATTTTCAGCAGATC
	sgRNA-qPCR-F	GGCTTACAGGAACTCCAGGA
	sgRNA-qPCR-R	CGACTCGGTGCCACTTTTTC
	18s-qPCR-F	GTAACCCGTTGAACCCCATT
	18s-qPCR-R	CCATCCAATCGGTAGTAGCG
On-target deep sequencing primers	On-target-F-Nextera	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTATTGGATGTGTACATGTCAGGTTCA
	On-target-R-Nextera	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTCTGCTAAAAAGTCTCTGTCACC
	miSeq-Univ-F	AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTC
	miSeq-Barcode-R	CAAGCAGAAGACGGCATACGAGATXXXXXXGTCTCGTGGGCTCGG

Green: Nextera adaptor sequence
Red: Transposon end sequence
Blue: On-target primers to *Dmd* exon 45
XXXXXX: barcode sequence see table S1

CHAPTER FOUR

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Acknowledgement

Parts of this chapter, including figures, have been reproduced, with or without modifications, from my previously published work (Zhang et al., 2018)

Challenges of Therapeutic Genome Editing

Immunogenicity

One of the greatest challenges of using rAAV as a delivery system for CRISPR/Cas-mediated therapeutic genome editing is the immune response to the vector. Potential immunogenicity elicited by rAAV-based delivery of the CRISPR/Cas system can be evoked by: i) the restored protein product; ii) the CRISPR/Cas system; and iii) capsid proteins on the surface of rAAV virus.

Mutated genes in monogenic disorders encode abnormal proteins or cause a complete loss of protein. After CRISPR/Cas-mediated correction, epitopes derived from the newly restored protein may elicit immunogenicity. However, in the case of DMD, due to somatic mutation or alternative splicing, more than 50% of DMD patients display low level of dystrophin-positive revertant fibers (0.2 to 4%), which may mitigate a potential immune response (Burrow et al., 1991; Nicholson et al., 1989; Nicholson et al., 1993). Indeed, in a gene transfer study, expression of murine full-length or mini-dystrophin in *mdx* mice did not evoke

humoral or cytotoxic immune responses (Ferrer et al., 2000). Therefore, immunogenicity elicited by the rescued protein may not be a significant concern, at least in the case of DMD.

In regard to immunogenicity elicited by the CRISPR/Cas system, it was demonstrated that Cas9 endonucleases delivered by rAAV did evoke a humoral immune response in mice (Chew et al., 2016). However, they did not observe significant muscle cell damage or a repair response at 2 weeks after rAAV administration. Another concern is if the rAAV needs to be re-administered since once the *Sp*Cas9-mediated humoral immunity is established in the host, further application of *Sp*Cas9 for therapeutic genome editing may no longer be effective. Several potential strategies could be applied to address this issue, including i) large-scale functional variant profiling of *Sp*Cas9 for epitope mutation, ii) replacing *Sp*Cas9 with other Cas endonucleases such as *Sa*Cas9, *Lb*Cpf1 and *As*Cpf1 after initial *Sp*Cas9 administration, and iii) performing plasmapheresis or transient immunosuppression to reduce the circulating antibody titer.

In addition to the immunogenicity response elicited by the CRISPR/Cas system, the humoral immune response evoked by rAAV is another challenge for *in vivo* therapeutic genome editing. Several studies in non-human primates have shown that high incidence of neutralizing antibodies (NAbs) after initial exposure to AAV can block AAV transduction, rendering gene delivery ineffective (Jiang et al., 2006; Wang et al., 2010). Moreover, the prevalence of AAV NAbs in human populations is also relatively high, ranging from 30-60% to AAV2, to 15-30% to AAV7, AAV8, and AAV9 serotypes (Boutin et al., 2010; Calcedo et al., 2009). Several strategies have been developed to overcome AAV-induced humoral immune responses, including i) AAV capsid mutagenesis to alter NAb binding epitopes (Nance

and Duan, 2015), ii) plasmapheresis to reduce AAV NAb titer (Hurlbut et al., 2010; Monteilhet et al., 2011; Wang et al., 2011), and iii) transient immunosuppression (Mingozzi et al., 2013).

Off-target effects

Another concern around using programmable nucleases for therapeutic genome editing is the potential for mutagenesis caused by off-target effects. Initial programmable nuclease-like TALENs generally show minimal off-target effects because TALEN-mediated DNA DSBs require dimerization of two TALEN monomers and the dimerized TALEN pairs can recognize 30-40 bp of DNA sequence (Kim and Kim, 2014). In contrast, DNA DSBs induced by *Sp*Cas9, currently the most prevalently used Cas endonuclease, only requires a 20-nt sgRNA forming DNA-RNA duplex with the target DNA strand, which may increase mismatching frequency. Indeed, several early studies about *Sp*Cas9 specificity have shown that high-frequency mutagenesis caused by off-target effects is possible at mismatched sites (Fu et al., 2013; Hsu et al., 2013; Lin et al., 2014b). For example, INDELs caused by *Sp*Cas9 off-target cleavage can be detected at certain sites with up to five mismatches relative to the on-target site (Fu et al., 2013; Hsu et al., 2013; Pattanayak et al., 2013). Several approaches can be used to evaluate potential CRISPR/Cas-induced off-target effects. Computational prediction of off-target sites followed by DNA mismatch cleavage assay serves as a rapid and convenient method to evaluate sgRNA specificity. More recently, computational prediction of off-target sites followed by deep sequencing or unbiased whole-genome sequencing represents a more reliable method for systematic evaluation of CRISPR/Cas specificity (Tsai and Joung, 2016).

Many strategies have been developed to reduce off-target effects caused by CRISPR/Cas9 non-specific cleavage. Using truncated sgRNA with 2-3 nt deletion at the 5'-end can reduce the off-target-induced mutagenesis by 5,000-fold or more, and reduce the genome-wide off-target sites by 2-5-fold (Fu et al., 2014; Tsai et al., 2015). Another strategy is to use a pair of Cas9 nickases to generate two DNA nicks in close proximity and this strategy has been shown to reduce off-target mutations by 50-1,500-fold (Mali et al., 2013a; Ran et al., 2013; Shen et al., 2014).

Resolving the crystal structure of *Sp*Cas9 in complex with sgRNA and target DNA provided many insights into the potential improvement of Cas9 specificity through logistic engineering (Nishimasu et al., 2014). The high-fidelity *Sp*Cas9 (*Sp*Cas9-HF1) was generated by alanine substitution to disrupt non-specific interaction with the target DNA strand (Kleinstiver et al., 2016). Recent structural analysis demonstrated that binding *Sp*Cas9-HF1 to a substrate with even a single bp mismatch at the PAM distal end completely abolishes stable docking of the HNH nuclease domain (Chen et al., 2017). In addition, enhanced *Sp*Cas9 (e*Sp*Cas9 1.1) was developed by neutralizing positively-charged residues to weaken interactions with the non-target DNA strand (Slaymaker et al., 2016). Additionally, a hyper-accurate Cas9 variant (HypaCas9) was developed in which the HNH nuclease activation is allosterically regulated by REC3, leading to high specificity of on-target cleavage (Chen et al., 2017). All of these engineered *Sp*Cas9 variants have been shown to significantly reduce off-target mutagenesis and have the potential to push therapeutic genome editing toward high specificity.

Long term effects and benefits of post-natal genome editing

Post-natal genome editing can correct the genetic mutation in muscular dystrophies, and in the short term, leads to improved muscle function. It has been reported that AAV CRISPR/Cas9 genome edited *mdx* mice can sustain dystrophin protein expression for more than 1 year (Hakim et al., 2018; Nelson et al., 2019). However, it is unknown whether this long term efficacy seen in DMD mouse models can also be observed in humans. This is because DMD patients can live more than 20 years, it is uncertain whether genome-edited, dystrophin-positive myofibers will be diluted out by dystrophin-negative myofibers generated from satellite cells carrying the *DMD* mutation. In contrast, cardiomyocytes in the adult human heart have a very low turnover rate (1% at the age of 25 to 0.45% at the age of 75) (Bergmann et al., 2009; Bergmann et al., 2015). Therefore, therapeutic genome editing in the human heart should provide long-term clinical benefit. Indeed, several studies have demonstrated that CRISPR/Cas-mediated genome editing can restore cardiac function in dystrophic mice (El Refaey et al., 2017; Long et al., 2016).

Conclusions and Future Perspectives

The ultimate goal for therapeutic CRISPR/Cas-mediated genome editing is to permanently correct mutations contributing to human genetic diseases. Skeletal muscle is an ideal tissue for CRISPR/Cas9-mediated therapeutic genome editing because of the post-mitotic and multinucleation features. Studies have shown that genomic mutation correction in a subset population of muscle nuclei is sufficient to result in continued improvement of muscle function (Amoasii et al., 2018; Amoasii et al., 2019; Amoasii et al., 2017; Bengtsson et al., 2017; El

Refaey et al., 2017; Long et al., 2016; Long et al., 2014; Min et al., 2019; Nelson et al., 2016; Tabebordbar et al., 2016; Zhang et al., 2017b). In addition to genome editing in post-mitotic myofibers, mutation correction in skeletal muscle satellite cells should provide significant regeneration benefit in the long term, although it is still debatable whether rAAV can efficiently transduce satellite cells *in vivo* (Arnett et al., 2014; Tabebordbar et al., 2016). Recently, our lab discovered a new myogenic lineage, which is molecularly and anatomically distinct from satellite cells. These progenitor cells, expressing Twist2, are capable of contributing to type IIb/x myofibers by fusing to myofibers during adulthood and muscle regeneration (Liu et al., 2017). In the future, it would be interesting to test whether rAAV can efficiently transduce Twist2^{+} progenitor cells, leading to CRISPR/Cas-mediated genome editing in these progenitor cells.

In addition to mutations in the nuclear genome, mutations in the mitochondrial genome also cause primary mitochondrial DNA (mtDNA)-related diseases. Mutations in mitochondrial tRNAs or protein-coding genes can lead to MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes) or MERRF (myoclonus epilepsy and ragged red fibers) (DiMauro, 2004; DiMauro and Hirano, 1993a, b; Schon et al., 2012). These genetic disorders not only affect muscle function but also result in pathogenesis in other systems, including brain, blood vessels and the endocrine system. Currently, mitochondrial replacement therapy (MRT) is used to treat genetic disorders caused by mtDNA mutations, in which the meiotic spindle apparatus with chromosomes from an unfertilized maternal oocyte is transferred into a donor oocyte cytoplasm containing healthy mtDNA (Wolf et al., 2017). However, healthy mtDNA replacement is not absolute and less than 1% carryover of mutant

mtDNA can be present after MRT. This carryover may lead to a gradual loss of healthy donor mtDNA and reversal to the maternal haplotype by genetic drift (Kang et al., 2016a; Yamada et al., 2016). This brings up an interesting question of whether programmable nucleases can be used to eliminate mutant mtDNA. Indeed, ZFNs and TALENs have been applied to selectively degrade pathogenic mitochondrial genomes (Bacman et al., 2014; Gammage et al., 2014; Reddy et al., 2015). Currently, efficient mitochondrial genome editing by the CRISPR/Cas system remains controversial (Gammage et al., 2017; Jo et al., 2015). Perhaps, efficient delivery of sgRNA into the mitochondrial matrix is an impediment for this application. We anticipate that in the near future, CRISPR/Cas-mediated genome editing can be further expanded to the mitochondrial genome.

To date, there are 840 neuromuscular diseases known to be caused by mutations in 465 different genes, with 72 mapped loci awaiting gene identification (Bonne et al., 2017). These debilitating diseases cause early death or significantly impair the quality of life. Currently, there is no effective treatment for these diseases since most therapies developed to date focus on alleviating the symptoms or targeting the secondary effects, while the source of mutations is still present in the human genome. The discovery and application of programmable nucleases for site-specific DNA DSBs provides a powerful tool for precise genome engineering. In particular, the CRISPR/Cas system has revolutionized the genome editing field and provides a new path for disease treatment that is beyond the reach of current therapies.

BIBLIOGRAPHY

Aartsma-Rus, A. (2012). Overview on DMD exon skipping. Methods Mol Biol *867*, 97-116.

Aartsma-Rus, A., Fokkema, I., Verschuuren, J., Ginjaar, I., van Deutekom, J., van Ommen, G.J., and den Dunnen, J.T. (2009). Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. Hum Mutat *30*, 293-299.

Aartsma-Rus, A., Ginjaar, I.B., and Bushby, K. (2016). The importance of genetic diagnosis for Duchenne muscular dystrophy. J Med Genet *53*, 145-151.

Aartsma-Rus, A., Van Deutekom, J.C., Fokkema, I.F., Van Ommen, G.J., and Den Dunnen, J.T. (2006). Entries in the Leiden Duchenne muscular dystrophy mutation database: an overview of mutation types and paradoxical cases that confirm the reading-frame rule. Muscle Nerve *34*, 135-144.

Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L.*, et al.* (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science *353*, aaf5573.

Ahn, A.H., and Kunkel, L.M. (1993). The structural and functional diversity of dystrophin. Nat Genet *3*, 283-291.

Amoasii, L., Hildyard, J.C.W., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., Harron, R., Stathopoulou, T.R., Massey, C., Shelton, J.M.*, et al.* (2018). Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. Science *362*, 86-91.

Amoasii, L., Li, H., Zhang, Y., Min, Y.L., Sanchez-Ortiz, E., Shelton, J.M., Long, C., Mireault, A.A., Bhattacharyya, S., McAnally, J.R.*, et al.* (2019). In vivo non-invasive monitoring of dystrophin correction in a new Duchenne muscular dystrophy reporter mouse. Nat Commun *10*, 4537.

Amoasii, L., Long, C., Li, H., Mireault, A.A., Shelton, J.M., Sanchez-Ortiz, E., McAnally, J.R., Bhattacharyya, S., Schmidt, F., Grimm, D.*, et al.* (2017). Single-cut genome editing restores dystrophin expression in a new mouse model of muscular dystrophy. Sci Transl Med *9*.

Araki, E., Nakamura, K., Nakao, K., Kameya, S., Kobayashi, O., Nonaka, I., Kobayashi, T., and Katsuki, M. (1997). Targeted disruption of exon 52 in the mouse dystrophin gene induced muscle degeneration similar to that observed in Duchenne muscular dystrophy. Biochem Biophys Res Commun *238*, 492-497.

Arnett, A.L., Konieczny, P., Ramos, J.N., Hall, J., Odom, G., Yablonka-Reuveni, Z., Chamberlain, J.R., and Chamberlain, J.S. (2014). Adeno-associated viral (AAV) vectors do not efficiently target muscle satellite cells. Mol Ther Methods Clin Dev *1*.

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

Atencia-Fernandez, S., Shiel, R.E., Mooney, C.T., and Nolan, C.M. (2015). Muscular dystrophy in the Japanese Spitz: an inversion disrupts the DMD and RPGR genes. Anim Genet *46*, 175-184.

Audebert, M., Salles, B., and Calsou, P. (2004). Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. J Biol Chem *279*, 55117-55126.

Auer, T.O., Duroure, K., De Cian, A., Concordet, J.P., and Del Bene, F. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome Res *24*, 142-153.

Bacman, S.R., Williams, S.L., Pinto, M., and Moraes, C.T. (2014). The use of mitochondria-targeted endonucleases to manipulate mtDNA. Methods Enzymol *547*, 373-397.

Baghdadi, M.B., and Tajbakhsh, S. (2017). Regulation and phylogeny of skeletal muscle regeneration. Dev Biol.

Balsam, L.B., Wagers, A.J., Christensen, J.L., Kofidis, T., Weissman, I.L., and Robbins, R.C. (2004). Haematopoietic stem cells adopt mature haematopoietic fates in ischaemic myocardium. Nature *428*, 668-673.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709-1712.

Baskin, K.K., Grueter, C.E., Kusminski, C.M., Holland, W.L., Bookout, A.L., Satapati, S., Kong, Y.M., Burgess, S.C., Malloy, C.R., Scherer, P.E.*, et al.* (2014). MED13-dependent signaling from the heart confers leanness by enhancing metabolism in adipose tissue and liver. EMBO Mol Med *6*, 1610-1621.

Bengtsson, N.E., Hall, J.K., Odom, G.L., Phelps, M.P., Andrus, C.R., Hawkins, R.D., Hauschka, S.D., Chamberlain, J.R., and Chamberlain, J.S. (2017). Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. Nat Commun *8*, 14454.

Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabe-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H.*, et al.* (2009). Evidence for cardiomyocyte renewal in humans. Science *324*, 98-102.

Bergmann, O., Zdunek, S., Felker, A., Salehpour, M., Alkass, K., Bernard, S., Sjostrom, S.L., Szewczykowska, M., Jackowska, T., Dos Remedios, C.*, et al.* (2015). Dynamics of Cell Generation and Turnover in the Human Heart. Cell *161*, 1566-1575.

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

Bhakta, M.S., and Segal, D.J. (2010). The generation of zinc finger proteins by modular assembly. Methods Mol Biol *649*, 3-30.

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

Bladen, C.L., Salgado, D., Monges, S., Foncuberta, M.E., Kekou, K., Kosma, K., Dawkins, H., Lamont, L., Roy, A.J., Chamova, T.*, et al.* (2015). The TREAT-NMD DMD Global Database: analysis of more than 7,000 Duchenne muscular dystrophy mutations. Hum Mutat *36*, 395-402.

Blankinship, M.J., Gregorevic, P., Allen, J.M., Harper, S.Q., Harper, H., Halbert, C.L., Miller, A.D., and Chamberlain, J.S. (2004). Efficient transduction of skeletal muscle using vectors based on adeno-associated virus serotype 6. Mol Ther *10*, 671-678.

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

Bolderson, E., Tomimatsu, N., Richard, D.J., Boucher, D., Kumar, R., Pandita, T.K., Burma, S., and Khanna, K.K. (2010). Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. Nucleic Acids Res *38*, 1821-1831.

Bolotin, A., Quinquis, B., Sorokin, A., and Ehrlich, S.D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology *151*, 2551-2561.

Bonamassa, B., Hai, L., and Liu, D. (2011). Hydrodynamic gene delivery and its applications in pharmaceutical research. Pharm Res *28*, 694-701.

Bonne, G., Rivier, F., and Hamroun, D. (2017). The 2018 version of the gene table of monogenic neuromuscular disorders (nuclear genome). Neuromuscul Disord *27*, 1152-1183.

Bostick, B., Shin, J.H., Yue, Y., and Duan, D. (2011). AAV-microdystrophin therapy improves cardiac performance in aged female mdx mice. Mol Ther *19*, 1826-1832.

Boutin, S., Monteilhet, V., Veron, P., Leborgne, C., Benveniste, O., Montus, M.F., and Masurier, C. (2010). Prevalence of serum IgG and neutralizing factors against adeno-associated virus (AAV) types 1, 2, 5, 6, 8, and 9 in the healthy population: implications for gene therapy using AAV vectors. Hum Gene Ther *21*, 704-712.

Brack, A.S., and Rando, T.A. (2012). Tissue-specific stem cells: lessons from the skeletal muscle satellite cell. Cell Stem Cell *10*, 504-514.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Briggs, A.W., Rios, X., Chari, R., Yang, L., Zhang, F., Mali, P., and Church, G.M. (2012). Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. Nucleic Acids Res *40*, e117.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science *321*, 960-964.

Bulfield, G., Siller, W.G., Wight, P.A., and Moore, K.J. (1984). X chromosome-linked muscular dystrophy (mdx) in the mouse. Proc Natl Acad Sci U S A *81*, 1189-1192.

Burridge, P.W., Holmstrom, A., and Wu, J.C. (2015). Chemically Defined Culture and Cardiomyocyte Differentiation of Human Pluripotent Stem Cells. Curr Protoc Hum Genet *87*, 21 23 21-15.

Burrow, K.L., Coovert, D.D., Klein, C.J., Bulman, D.E., Kissel, J.T., Rammohan, K.W., Burghes, A.H., and Mendell, J.R. (1991). Dystrophin expression and somatic reversion in prednisone-treated and untreated Duchenne dystrophy. CIDD Study Group. Neurology *41*, 661-666.

Bushby, K., Muntoni, F., Urtizberea, A., Hughes, R., and Griggs, R. (2004). Report on the 124th ENMC International Workshop. Treatment of Duchenne muscular dystrophy; defining the gold standards of management in the use of corticosteroids. 2-4 April 2004, Naarden, The Netherlands. Neuromuscul Disord *14*, 526-534.

Calcedo, R., Vandenberghe, L.H., Gao, G., Lin, J., and Wilson, J.M. (2009). Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. J Infect Dis *199*, 381-390.

Campbell, K.P., and Kahl, S.D. (1989). Association of dystrophin and an integral membrane glycoprotein. Nature *338*, 259-262.

Capecchi, M.R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet *6*, 507-512.

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

Chamberlain, J.S., Metzger, J., Reyes, M., Townsend, D., and Faulkner, J.A. (2007). Dystrophin-deficient mdx mice display a reduced life span and are susceptible to spontaneous rhabdomyosarcoma. FASEB J *21*, 2195-2204.

Chan, S.H., Yu, A.M., and McVey, M. (2010). Dual roles for DNA polymerase theta in alternative end-joining repair of double-strand breaks in Drosophila. PLoS Genet *6*, e1001005.

Chandrasegaran, S., and Carroll, D. (2016). Origins of Programmable Nucleases for Genome Engineering. J Mol Biol *428*, 963-989.

Chang, N.C., and Rudnicki, M.A. (2014). Satellite cells: the architects of skeletal muscle. Curr Top Dev Biol *107*, 161-181.

Chao, H., Liu, Y., Rabinowitz, J., Li, C., Samulski, R.J., and Walsh, C.E. (2000). Several log increase in therapeutic transgene delivery by distinct adeno-associated viral serotype vectors. Mol Ther *2*, 619-623.

Chapman, V.M., Miller, D.R., Armstrong, D., and Caskey, C.T. (1989). Recovery of induced mutations for X chromosome-linked muscular dystrophy in mice. Proc Natl Acad Sci U S A *86*, 1292-1296.

Charleston, J.S., Schnell, F.J., Dworzak, J., Donoghue, C., Lewis, S., Chen, L., Young, G.D., Milici, A.J., Voss, J., DeAlwis, U.*, et al.* (2018). Eteplirsen treatment for Duchenne muscular dystrophy: Exon skipping and dystrophin production. Neurology *90*, e2146-e2154.

Charlesworth, C.T., Deshpande, P.S., Dever, D.P., Camarena, J., Lemgart, V.T., Cromer, M.K., Vakulskas, C.A., Collingwood, M.A., Zhang, L., Bode, N.M.*, et al.* (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. Nat Med *25*, 249-254.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. Nature.

Chen, S., Lee, B., Lee, A.Y., Modzelewski, A.J., and He, L. (2016). Highly Efficient Mouse Genome Editing by CRISPR Ribonucleoprotein Electroporation of Zygotes. J Biol Chem *291*, 14457-14467.

Chen, Y., Zheng, Y., Kang, Y., Yang, W., Niu, Y., Guo, X., Tu, Z., Si, C., Wang, H., Xing, R.*, et al.* (2015). Functional disruption of the dystrophin gene in rhesus monkey using CRISPR/Cas9. Hum Mol Genet *24*, 3764-3774.

Chew, W.L., Tabebordbar, M., Cheng, J.K., Mali, P., Wu, E.Y., Ng, A.H., Zhu, K., Wagers, A.J., and Church, G.M. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. Nat Methods *13*, 868-874.

Chiurazzi, M., Ray, A., Viret, J.F., Perera, R., Wang, X.H., Lloyd, A.M., and Signer, E.R. (1996). Enhancement of somatic intrachromosomal homologous recombination in Arabidopsis by the HO endonuclease. Plant Cell *8*, 2057-2066.

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

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

Ciccia, A., and Elledge, S.J. (2010). The DNA damage response: making it safe to play with knives. Mol Cell *40*, 179-204.

Cirak, S., Arechavala-Gomeza, V., Guglieri, M., Feng, L., Torelli, S., Anthony, K., Abbs, S., Garralda, M.E., Bourke, J., Wells, D.J.*, et al.* (2011). 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.

Conboy, M.J., Karasov, A.O., and Rando, T.A. (2007). High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. PLoS Biol *5*, e102.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Cox, D.B., Platt, R.J., and Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. Nat Med *21*, 121-131.

Cozzi, F., Cerletti, M., Luvoni, G.C., Lombardo, R., Brambilla, P.G., Faverzani, S., Blasevich, F., Cornelio, F., Pozza, O., and Mora, M. (2001). Development of muscle pathology in canine X-linked muscular dystrophy. II. Quantitative characterization of histopathological progression during postnatal skeletal muscle development. Acta Neuropathol *101*, 469-478.

Davis, A.J., and Chen, D.J. (2013). DNA double strand break repair via non-homologous end-joining. Transl Cancer Res *2*, 130-143.

Deconinck, A.E., Rafael, J.A., Skinner, J.A., Brown, S.C., Potter, A.C., Metzinger, L., Watt, D.J., Dickson, J.G., Tinsley, J.M., and Davies, K.E. (1997). Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. Cell *90*, 717-727.

Deng, C., and Capecchi, M.R. (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol Cell Biol *12*, 3365-3371.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.K., Shi, Y., and Yan, N. (2012). Structural basis for sequence-specific recognition of DNA by TAL effectors. Science *335*, 720-723.

DeWitt, M.A., Magis, W., Bray, N.L., Wang, T., Berman, J.R., Urbinati, F., Heo, S.J., Mitros, T., Munoz, D.P., Boffelli, D.*, et al.* (2016). Selection-free genome editing of the sickle mutation in human adult hematopoietic stem/progenitor cells. Sci Transl Med *8*, 360ra134.

DiMauro, S. (2004). Mitochondrial diseases. Biochim Biophys Acta *1658*, 80-88.

DiMauro, S., and Hirano, M. (1993a). Melas. In GeneReviews(R), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, H.C. Mefford, K. Stephens, A. Amemiya, and N. Ledbetter, eds. (Seattle (WA)).

DiMauro, S., and Hirano, M. (1993b). Merrf. In GeneReviews(R), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, H.C. Mefford, K. Stephens, A. Amemiya, and N. Ledbetter, eds. (Seattle (WA)).

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

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

Dreier, B., Beerli, R.R., Segal, D.J., Flippin, J.D., and Barbas, C.F., 3rd (2001). Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. J Biol Chem *276*, 29466-29478.

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

Duan, D. (2018). Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy. Mol Ther *26*, 2337-2356.

Duan, D., Sharma, P., Yang, J., Yue, Y., Dudus, L., Zhang, Y., Fisher, K.J., and Engelhardt, J.F. (1998). Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. J Virol *72*, 8568-8577.

Duan, D., Yan, Z., Yue, Y., Ding, W., and Engelhardt, J.F. (2001). Enhancement of muscle gene delivery with pseudotyped adeno-associated virus type 5 correlates with myoblast differentiation. J Virol *75*, 7662-7671.

Duan, D., Yue, Y., and Engelhardt, J.F. (2003). Consequences of DNA-dependent protein kinase catalytic subunit deficiency on recombinant adeno-associated virus genome circularization and heterodimerization in muscle tissue. J Virol *77*, 4751-4759.

Dunbar, C.E., High, K.A., Joung, J.K., Kohn, D.B., Ozawa, K., and Sadelain, M. (2018). Gene therapy comes of age. Science *359*.

East-Seletsky, A., O'Connell, M.R., Burstein, D., Knott, G.J., and Doudna, J.A. (2017). RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. Mol Cell *66*, 373-383 e373.

El Refaey, M., Xu, L., Gao, Y., Canan, B.D., Adesanya, T.A., Warner, S.C., Akagi, K., Symer, D.E., Mohler, P.J., Ma, J.*, et al.* (2017). In Vivo Genome Editing Restores Dystrophin Expression and Cardiac Function in Dystrophic Mice. Circ Res *121*, 923-929.

Ellison, G.M., Vicinanza, C., Smith, A.J., Aquila, I., Leone, A., Waring, C.D., Henning, B.J., Stirparo, G.G., Papait, R., Scarfo, M.*, et al.* (2013). Adult c-kit(pos) cardiac stem cells are necessary and sufficient for functional cardiac regeneration and repair. Cell *154*, 827-842.

Endo, A., Masafumi, M., Kaya, H., and Toki, S. (2016). Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from Francisella novicida. Sci Rep *6*, 38169.

Ferrari, F.K., Samulski, T., Shenk, T., and Samulski, R.J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. J Virol *70*, 3227-3234.

Ferrer, A., Wells, K.E., and Wells, D.J. (2000). Immune responses to dystropin: implications for gene therapy of Duchenne muscular dystrophy. Gene Ther *7*, 1439-1446.

Fisher, K.J., Gao, G.P., Weitzman, M.D., DeMatteo, R., Burda, J.F., and Wilson, J.M. (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. J Virol *70*, 520-532.

Flotte, T.R., and Berns, K.I. (2005). Adeno-associated virus: a ubiquitous commensal of mammals. Hum Gene Ther *16*, 401-407.

Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A., and Charpentier, E. (2016). The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. Nature *532*, 517-521.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol *31*, 822-826.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol *32*, 279-284.

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

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

Gammage, P.A., Moraes, C.T., and Minczuk, M. (2017). Mitochondrial Genome Engineering: The Revolution May Not Be CRISPR-Ized. Trends Genet.

Gammage, P.A., Rorbach, J., Vincent, A.I., Rebar, E.J., and Minczuk, M. (2014). Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. EMBO Mol Med *6*, 458-466.

Gao, G.P., Alvira, M.R., Wang, L., Calcedo, R., Johnston, J., and Wilson, J.M. (2002). Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. Proc Natl Acad Sci U S A *99*, 11854-11859.

Gao, Q.Q., and McNally, E.M. (2015). The Dystrophin Complex: Structure, Function, and Implications for Therapy. Compr Physiol *5*, 1223-1239.

Gao, Y., Guo, X., Santostefano, K., Wang, Y., Reid, T., Zeng, D., Terada, N., Ashizawa, T., and Xia, G. (2016). Genome Therapy of Myotonic Dystrophy Type 1 iPS Cells for Development of Autologous Stem Cell Therapy. Mol Ther *24*, 1378-1387.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A *109*, E2579-2586.

Genovese, P., Schiroli, G., Escobar, G., Tomaso, T.D., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C.*, et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. Nature *510*, 235-240.

Gonzalez, B., Schwimmer, L.J., Fuller, R.P., Ye, Y., Asawapornmongkol, L., and Barbas, C.F., 3rd (2010). Modular system for the construction of zinc-finger libraries and proteins. Nat Protoc *5*, 791-810.

Grady, R.M., Teng, H., Nichol, M.C., Cunningham, J.C., Wilkinson, R.S., and Sanes, J.R. (1997). Skeletal and cardiac myopathies in mice lacking utrophin and dystrophin: a model for Duchenne muscular dystrophy. Cell *90*, 729-738.

Gregorevic, P., Blankinship, M.J., Allen, J.M., Crawford, R.W., Meuse, L., Miller, D.G., Russell, D.W., and Chamberlain, J.S. (2004). Systemic delivery of genes to striated muscles using adeno-associated viral vectors. Nat Med *10*, 828-834.

Guiraud, S., Aartsma-Rus, A., Vieira, N.M., Davies, K.E., van Ommen, G.J., and Kunkel, L.M. (2015). The Pathogenesis and Therapy of Muscular Dystrophies. Annu Rev Genomics Hum Genet *16*, 281-308.

Gundry, M.C., Brunetti, L., Lin, A., Mayle, A.E., Kitano, A., Wagner, D., Hsu, J.I., Hoegenauer, K.A., Rooney, C.M., Goodell, M.A.*, et al.* (2016). Highly Efficient Genome Editing of Murine and Human Hematopoietic Progenitor Cells by CRISPR/Cas9. Cell Rep *17*, 1453-1461.

Haft, D.H., Selengut, J., Mongodin, E.F., and Nelson, K.E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. PLoS Comput Biol *1*, e60.

Hakim, C.H., Wasala, N.B., Nelson, C.E., Wasala, L.P., Yue, Y., Louderman, J.A., Lessa, T.B., Dai, A., Zhang, K., Jenkins, G.J.*, et al.* (2018). AAV CRISPR editing rescues cardiac and muscle function for 18 months in dystrophic mice. JCI Insight *3*.

Himeda, C.L., Chen, X., and Hauschka, S.D. (2011). Design and testing of regulatory cassettes for optimal activity in skeletal and cardiac muscles. Methods Mol Biol *709*, 3-19.

Hinderer, C., Katz, N., Buza, E.L., Dyer, C., Goode, T., Bell, P., Richman, L.K., and Wilson, J.M. (2018). Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN. Hum Gene Ther *29*, 285-298.

Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., Yamamoto, T., and Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. Sci Rep *5*, 8841.

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

Hollinger, K., and Chamberlain, J.S. (2015). Viral vector-mediated gene therapies. Curr Opin Neurol *28*, 522-527.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc Natl Acad Sci U S A *110*, 15644-15649.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O.*, et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol *31*, 827-832.

Hur, J.K., Kim, K., Been, K.W., Baek, G., Ye, S., Hur, J.W., Ryu, S.M., Lee, Y.S., and Kim, J.S. (2016). Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. Nat Biotechnol *34*, 807-808.

Hurlbut, G.D., Ziegler, R.J., Nietupski, J.B., Foley, J.W., Woodworth, L.A., Meyers, E., Bercury, S.D., Pande, N.N., Souza, D.W., Bree, M.P.*, et al.* (2010). Preexisting immunity and low expression in primates highlight translational challenges for liver-directed AAV8-mediated gene therapy. Mol Ther *18*, 1983-1994.

Inagaki, K., Fuess, S., Storm, T.A., Gibson, G.A., McTiernan, C.F., Kay, M.A., and Nakai, H. (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. Mol Ther *14*, 45-53.

Ira, G., Pellicioli, A., Balijja, A., Wang, X., Fiorani, S., Carotenuto, W., Liberi, G., Bressan, D., Wan, L., Hollingsworth, N.M.*, et al.* (2004). DNA end resection,

homologous recombination and DNA damage checkpoint activation require CDK1. Nature *431*, 1011-1017.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol *169*, 5429-5433.

Iyombe-Engembe, J.P., Ouellet, D.L., Barbeau, X., Rousseau, J., Chapdelaine, P., Lague, P., and Tremblay, J.P. (2016). Efficient Restoration of the Dystrophin Gene Reading Frame and Protein Structure in DMD Myoblasts Using the CinDel Method. Mol Ther Nucleic Acids *5*, e283.

Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Mol Microbiol *43*, 1565-1575.

Jiang, H., Couto, L.B., Patarroyo-White, S., Liu, T., Nagy, D., Vargas, J.A., Zhou, S., Scallan, C.D., Sommer, J., Vijay, S.*, et al.* (2006). Effects of transient immunosuppression on adenoassociated, virus-mediated, liver-directed gene transfer in rhesus macaques and implications for human gene therapy. Blood *108*, 3321-3328.

Jiang, Y., Qian, F., Yang, J., Liu, Y., Dong, F., Xu, C., Sun, B., Chen, B., Xu, X., Li, Y.*, et al.* (2017). CRISPR-Cpf1 assisted genome editing of Corynebacterium glutamicum. Nat Commun *8*, 15179.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Jo, A., Ham, S., Lee, G.H., Lee, Y.I., Kim, S., Lee, Y.S., Shin, J.H., and Lee, Y. (2015). Efficient Mitochondrial Genome Editing by CRISPR/Cas9. Biomed Res Int *2015*, 305716.

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

Kabadi, A.M., Ousterout, D.G., Hilton, I.B., and Gersbach, C.A. (2014). Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. Nucleic Acids Res *42*, e147.

Kanagawa, M., and Toda, T. (2006). The genetic and molecular basis of muscular dystrophy: roles of cell-matrix linkage in the pathogenesis. J Hum Genet *51*, 915-926.

Kang, E., Wu, J., Gutierrez, N.M., Koski, A., Tippner-Hedges, R., Agaronyan, K., Platero-Luengo, A., Martinez-Redondo, P., Ma, H., Lee, Y.*, et al.* (2016a). Mitochondrial

replacement in human oocytes carrying pathogenic mitochondrial DNA mutations. Nature *540*, 270-275.

Kang, X., He, W., Huang, Y., Yu, Q., Chen, Y., Gao, X., Sun, X., and Fan, Y. (2016b). Introducing precise genetic modifications into human 3PN embryos by CRISPR/Cas-mediated genome editing. J Assist Reprod Genet *33*, 581-588.

Kemaladewi, D.U., Maino, E., Hyatt, E., Hou, H., Ding, M., Place, K.M., Zhu, X., Bassi, P., Baghestani, Z., Deshwar, A.G.*, et al.* (2017). Correction of a splicing defect in a mouse model of congenital muscular dystrophy type 1A using a homology-directed-repair-independent mechanism. Nat Med *23*, 984-989.

Kim, H., and Kim, J.S. (2014). A guide to genome engineering with programmable nucleases. Nat Rev Genet *15*, 321-334.

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

Kim, S.K., Kim, H., Ahn, W.C., Park, K.H., Woo, E.J., Lee, D.H., and Lee, S.G. (2017). Efficient Transcriptional Gene Repression by Type V-A CRISPR-Cpf1 from Eubacterium eligens. ACS Synth Biol *6*, 1273-1282.

Kim, Y., Cheong, S.A., Lee, J.G., Lee, S.W., Lee, M.S., Baek, I.J., and Sung, Y.H. (2016). Generation of knockout mice by Cpf1-mediated gene targeting. Nat Biotechnol *34*, 808-810.

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

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

Kim, Y.G., Li, L., and Chandrasegaran, S. (1994). Insertion and deletion mutants of FokI restriction endonuclease. J Biol Chem *269*, 31978-31982.

Kim, Y.G., Smith, J., Durgesha, M., and Chandrasegaran, S. (1998). Chimeric restriction enzyme: Gal4 fusion to FokI cleavage domain. Biol Chem *379*, 489-495.

Kimura, Y., Hisano, Y., Kawahara, A., and Higashijima, S. (2014). Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. Sci Rep *4*, 6545.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature *529*, 490-495.

Klymiuk, N., Blutke, A., Graf, A., Krause, S., Burkhardt, K., Wuensch, A., Krebs, S., Kessler, B., Zakhartchenko, V., Kurome, M.*, et al.* (2013). Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. Hum Mol Genet *22*, 4368-4382.

Koenig, M., Hoffman, E.P., Bertelson, C.J., Monaco, A.P., Feener, C., and Kunkel, L.M. (1987). Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. Cell *50*, 509-517.

Kornegay, J.N., Bogan, J.R., Bogan, D.J., Childers, M.K., Li, J., Nghiem, P., Detwiler, D.A., Larsen, C.A., Grange, R.W., Bhavaraju-Sanka, R.K.*, et al.* (2012). Canine models of Duchenne muscular dystrophy and their use in therapeutic strategies. Mamm Genome *23*, 85-108.

Kornegay, J.N., Li, J., Bogan, J.R., Bogan, D.J., Chen, C., Zheng, H., Wang, B., Qiao, C., Howard, J.F., Jr., and Xiao, X. (2010). Widespread muscle expression of an AAV9 human mini-dystrophin vector after intravenous injection in neonatal dystrophin-deficient dogs. Mol Ther *18*, 1501-1508.

Kotterman, M.A., and Schaffer, D.V. (2014). Engineering adeno-associated viruses for clinical gene therapy. Nat Rev Genet *15*, 445-451.

Kuang, S., Kuroda, K., Le Grand, F., and Rudnicki, M.A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. Cell *129*, 999-1010.

Kyrychenko, V., Kyrychenko, S., Tiburcy, M., Shelton, J.M., Long, C., Schneider, J.W., Zimmermann, W.H., Bassel-Duby, R., and Olson, E.N. (2017). Functional correction of dystrophin actin binding domain mutations by genome editing. JCI Insight *2*.

Larcher, T., Lafoux, A., Tesson, L., Remy, S., Thepenier, V., Francois, V., Le Guiner, C., Goubin, H., Dutilleul, M., Guigand, L.*, et al.* (2014). Characterization of dystrophin deficient rats: a new model for Duchenne muscular dystrophy. PLoS One *9*, e110371.

Lattanzi, A., Duguez, S., Moiani, A., Izmiryan, A., Barbon, E., Martin, S., Mamchaoui, K., Mouly, V., Bernardi, F., Mavilio, F.*, et al.* (2017). Correction of the Exon 2 Duplication in DMD Myoblasts by a Single CRISPR/Cas9 System. Mol Ther Nucleic Acids *7*, 11-19.

Lemos, B.R., Kaplan, A.C., Bae, J.E., Ferrazzoli, A.E., Kuo, J., Anand, R.P., Waterman, D.P., and Haber, J.E. (2018). CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. Proc Natl Acad Sci U S A *115*, E2040-E2047.

Lepper, C., Partridge, T.A., and Fan, C.M. (2011). An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. Development *138*, 3639-3646.

Li, B., Zhao, W., Luo, X., Zhang, X., Li, C., Zeng, C., and Dong, Y. (2017). Engineering CRISPR-Cpf1 crRNAs and mRNAs to maximize genome editing efficiency. Nat Biomed Eng *1*.

Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H.*, et al.* (2015). Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. Stem Cell Reports *4*, 143-154.

Li, L., and Chandrasegaran, S. (1993). Alteration of the cleavage distance of Fok I restriction endonuclease by insertion mutagenesis. Proc Natl Acad Sci U S A *90*, 2764-2768.

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

Li, L., Wu, L.P., Clarke, R., and Chandrasegaran, S. (1993). C-terminal deletion mutants of the FokI restriction endonuclease. Gene *133*, 79-84.

Liang, P., Xu, Y., Zhang, X., Ding, C., Huang, R., Zhang, Z., Lv, J., Xie, X., Chen, Y., Li, Y.*, et al.* (2015). CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. Protein Cell *6*, 363-372.

Lin, S.R., Yang, H.C., Kuo, Y.T., Liu, C.J., Yang, T.Y., Sung, K.C., Lin, Y.Y., Wang, H.Y., Wang, C.C., Shen, Y.C.*, et al.* (2014a). The CRISPR/Cas9 System Facilitates Clearance of the Intrahepatic HBV Templates In Vivo. Mol Ther Nucleic Acids *3*, e186.

Lin, Y., Cradick, T.J., Brown, M.T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B.M., Vertino, P.M., Stewart, F.J., and Bao, G. (2014b). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. Nucleic Acids Res *42*, 7473-7485.

Liu, N., Garry, G.A., Li, S., Bezprozvannaya, S., Sanchez-Ortiz, E., Chen, B., Shelton, J.M., Jaichander, P., Bassel-Duby, R., and Olson, E.N. (2017). A Twist2-dependent progenitor cell contributes to adult skeletal muscle. Nat Cell Biol *19*, 202-213.

Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400-403.

Long, C., Li, H., Tiburcy, M., Rodriguez-Caycedo, C., Kyrychenko, V., Zhou, H., Zhang, Y., Min, Y.L., Shelton, J.M., Mammen, P.P.A.*, et al.* (2018). Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. Sci Adv *4*, eaap9004.

Long, C., McAnally, J.R., Shelton, J.M., Mireault, A.A., Bassel-Duby, R., and Olson, E.N. (2014). Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. Science *345*, 1184-1188.

Ma, H., Marti-Gutierrez, N., Park, S.W., Wu, J., Lee, Y., Suzuki, K., Koski, A., Ji, D., Hayama, T., Ahmed, R.*, et al.* (2017a). Correction of a pathogenic gene mutation in human embryos. Nature *548*, 413-419.

Ma, S., Liu, Y., Liu, Y., Chang, J., Zhang, T., Wang, X., Shi, R., Lu, W., Xia, X., Zhao, P.*, et al.* (2017b). An integrated CRISPR Bombyx mori genome editing system with improved efficiency and expanded target sites. Insect Biochem Mol Biol *83*, 13-20.

Maeder, M.L., Stefanidakis, M., Wilson, C.J., Baral, R., Barrera, L.A., Bounoutas, G.S., Bumcrot, D., Chao, H., Ciulla, D.M., DaSilva, J.A.*, et al.* (2019). Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. Nat Med *25*, 229-233.

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

Magadan, A.H., Dupuis, M.E., Villion, M., and Moineau, S. (2012). Cleavage of phage DNA by the Streptococcus thermophilus CRISPR3-Cas system. PLoS One *7*, e40913.

Maggio, I., Liu, J., Janssen, J.M., Chen, X., and Goncalves, M.A. (2016a). Adenoviral vectors encoding CRISPR/Cas9 multiplexes rescue dystrophin synthesis in unselected populations of DMD muscle cells. Sci Rep *6*, 37051.

Maggio, I., Stefanucci, L., Janssen, J.M., Liu, J., Chen, X., Mouly, V., and Goncalves, M.A. (2016b). Selection-free gene repair after adenoviral vector transduction of designer nucleases: rescue of dystrophin synthesis in DMD muscle cell populations. Nucleic Acids Res *44*, 1449-1470.

Mak, A.N., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. Science *335*, 716-719.

Makarova, K.S., Aravind, L., Wolf, Y.I., and Koonin, E.V. (2011a). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biol Direct *6*, 38.

Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and Koonin, E.V. (2006). A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol Direct *1*, 7.

Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F.*, et al.* (2011b). Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol *9*, 467-477.

Makarova, K.S., Wolf, Y.I., Alkhnbashi, O.S., Costa, F., Shah, S.A., Saunders, S.J., Barrangou, R., Brouns, S.J., Charpentier, E., Haft, D.H.*, et al.* (2015). An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol *13*, 722-736.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013a). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat Biotechnol *31*, 833-838.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013b). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Mansour, S.L., Thomas, K.R., and Capecchi, M.R. (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature *336*, 348-352.

Maresca, M., Lin, V.G., Guo, N., and Yang, Y. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. Genome Res *23*, 539-546.

Marraffini, L.A. (2015). CRISPR-Cas immunity in prokaryotes. Nature *526*, 55-61.

Marraffini, L.A., and Sontheimer, E.J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science *322*, 1843-1845.

Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol *33*, 538-542.

McCarthy, J.J., Mula, J., Miyazaki, M., Erfani, R., Garrison, K., Farooqui, A.B., Srikuea, R., Lawson, B.A., Grimes, B., Keller, C.*, et al.* (2011). Effective fiber hypertrophy in satellite cell-depleted skeletal muscle. Development *138*, 3657-3666.

McCarty, D.M. (2008). Self-complementary AAV vectors; advances and applications. Mol Ther *16*, 1648-1656.

McCarty, D.M., Fu, H., Monahan, P.E., Toulson, C.E., Naik, P., and Samulski, R.J. (2003). Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. Gene Ther *10*, 2112-2118.

McGreevy, J.W., Hakim, C.H., McIntosh, M.A., and Duan, D. (2015). Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy. Dis Model Mech *8*, 195-213.

Meca-Cortes, O., Guerra-Rebollo, M., Garrido, C., Borros, S., Rubio, N., and Blanco, J. (2017). CRISPR/Cas9-Mediated Knockin Application in Cell Therapy: A Non-viral Procedure for Bystander Treatment of Glioma in Mice. Mol Ther Nucleic Acids *8*, 395-403.

Mendell, J.R., Al-Zaidy, S., Shell, R., Arnold, W.D., Rodino-Klapac, L.R., Prior, T.W., Lowes, L., Alfano, L., Berry, K., Church, K.*, et al.* (2017). Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. N Engl J Med *377*, 1713-1722.

Mercuri, E., and Muntoni, F. (2013). Muscular dystrophies. Lancet *381*, 845-860.

Miao, C.H., Snyder, R.O., Schowalter, D.B., Patijn, G.A., Donahue, B., Winther, B., and Kay, M.A. (1998). The kinetics of rAAV integration in the liver. Nat Genet *19*, 13-15.

Millay, D.P., Sargent, M.A., Osinska, H., Baines, C.P., Barton, E.R., Vuagniaux, G., Sweeney, H.L., Robbins, J., and Molkentin, J.D. (2008). Genetic and pharmacologic inhibition of mitochondrial-dependent necrosis attenuates muscular dystrophy. Nat Med *14*, 442-447.

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

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

Min, Y.L., Li, H., Rodriguez-Caycedo, C., Mireault, A.A., Huang, J., Shelton, J.M., McAnally, J.R., Amoasii, L., Mammen, P.P.A., Bassel-Duby, R.*, et al.* (2019). CRISPR-Cas9 corrects Duchenne muscular dystrophy exon 44 deletion mutations in mice and human cells. Sci Adv *5*, eaav4324.

Mingozzi, F., Chen, Y., Edmonson, S.C., Zhou, S., Thurlings, R.M., Tak, P.P., High, K.A., and Vervoordeldonk, M.J. (2013). Prevalence and pharmacological modulation of humoral immunity to AAV vectors in gene transfer to synovial tissue. Gene Ther *20*, 417-424.

Mojica, F.J., Diez-Villasenor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. J Mol Evol *60*, 174-182.

Mojica, F.J., Diez-Villasenor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol *36*, 244-246.

Monaco, A.P., Bertelson, C.J., Liechti-Gallati, S., Moser, H., and Kunkel, L.M. (1988). An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. Genomics *2*, 90-95.

Monteilhet, V., Saheb, S., Boutin, S., Leborgne, C., Veron, P., Montus, M.F., Moullier, P., Benveniste, O., and Masurier, C. (2011). A 10 patient case report on the impact of plasmapheresis upon neutralizing factors against adeno-associated virus (AAV) types 1, 2, 6, and 8. Mol Ther *19*, 2084-2091.

Monteys, A.M., Ebanks, S.A., Keiser, M.S., and Davidson, B.L. (2017). CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. Mol Ther *25*, 12-23.

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

Mou, H., Smith, J.L., Peng, L., Yin, H., Moore, J., Zhang, X.O., Song, C.Q., Sheel, A., Wu, Q., Ozata, D.M.*, et al.* (2017). CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion. Genome Biol *18*, 108.

Mourkioti, F., Kustan, J., Kraft, P., Day, J.W., Zhao, M.M., Kost-Alimova, M., Protopopov, A., DePinho, R.A., Bernstein, D., Meeker, A.K.*, et al.* (2013). Role of telomere dysfunction in cardiac failure in Duchenne muscular dystrophy. Nat Cell Biol *15*, 895-904.

Muller, M., Lee, C.M., Gasiunas, G., Davis, T.H., Cradick, T.J., Siksnys, V., Bao, G., Cathomen, T., and Mussolino, C. (2016). Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. Mol Ther *24*, 636-644.

Muntoni, F., Torelli, S., and Ferlini, A. (2003). Dystrophin and mutations: one gene, several proteins, multiple phenotypes. Lancet Neurol *2*, 731-740.

Murovec, J., Pirc, Z., and Yang, B. (2017). New variants of CRISPR RNA-guided genome editing enzymes. Plant Biotechnol J *15*, 917-926.

Murphy, M.M., Lawson, J.A., Mathew, S.J., Hutcheson, D.A., and Kardon, G. (2011). Satellite cells, connective tissue fibroblasts and their interactions are crucial for muscle regeneration. Development *138*, 3625-3637.

Murry, C.E., Soonpaa, M.H., Reinecke, H., Nakajima, H., Nakajima, H.O., Rubart, M., Pasumarthi, K.B., Virag, J.I., Bartelmez, S.H., Poppa, V.*, et al.* (2004). Haematopoietic stem cells do not transdifferentiate into cardiac myocytes in myocardial infarcts. Nature *428*, 664-668.

Nakade, S., Tsubota, T., Sakane, Y., Kume, S., Sakamoto, N., Obara, M., Daimon, T., Sezutsu, H., Yamamoto, T., Sakuma, T.*, et al.* (2014). Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat Commun *5*, 5560.

Nakai, H., Yant, S.R., Storm, T.A., Fuess, S., Meuse, L., and Kay, M.A. (2001). Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo. J Virol *75*, 6969-6976.

Nakamura, K., Fujii, W., Tsuboi, M., Tanihata, J., Teramoto, N., Takeuchi, S., Naito, K., Yamanouchi, K., and Nishihara, M. (2014). Generation of muscular dystrophy model rats with a CRISPR/Cas system. Sci Rep *4*, 5635.

Nance, M.E., and Duan, D. (2015). Perspective on Adeno-Associated Virus Capsid Modification for Duchenne Muscular Dystrophy Gene Therapy. Hum Gene Ther *26*, 786-800.

Nayerossadat, N., Maedeh, T., and Ali, P.A. (2012). Viral and nonviral delivery systems for gene delivery. Adv Biomed Res *1*, 27.

Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X.*, et al.* (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403-407.

Nelson, C.E., Wu, Y., Gemberling, M.P., Oliver, M.L., Waller, M.A., Bohning, J.D., Robinson-Hamm, J.N., Bulaklak, K., Castellanos Rivera, R.M., Collier, J.H.*, et al.* (2019). Long-term evaluation of AAV-CRISPR genome editing for Duchenne muscular dystrophy. Nat Med *25*, 427-432.

Nicholson, L.V., Davison, K., Johnson, M.A., Slater, C.R., Young, C., Bhattacharya, S., Gardner-Medwin, D., and Harris, J.B. (1989). Dystrophin in skeletal muscle. II. Immunoreactivity in patients with Xp21 muscular dystrophy. J Neurol Sci *94*, 137-146.

Nicholson, L.V., Johnson, M.A., Bushby, K.M., Gardner-Medwin, D., Curtis, A., Ginjaar, I.B., den Dunnen, J.T., Welch, J.L., Butler, T.J., Bakker, E.*, et al.* (1993). Integrated study of 100 patients with Xp21 linked muscular dystrophy using clinical, genetic, immunochemical, and histopathological data. Part 1. Trends across the clinical groups. J Med Genet *30*, 728-736.

Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935-949.

Niu, Y., Shen, B., Cui, Y., Chen, Y., Wang, J., Wang, L., Kang, Y., Zhao, X., Si, W., Li, W.*, et al.* (2014). Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. Cell *156*, 836-843.

Nonnenmacher, M., and Weber, T. (2012). Intracellular transport of recombinant adeno-associated virus vectors. Gene Ther *19*, 649-658.

O'Brien, K.F., and Kunkel, L.M. (2001). Dystrophin and muscular dystrophy: past, present, and future. Mol Genet Metab *74*, 75-88.

Orlic, D., Kajstura, J., Chimenti, S., Jakoniuk, I., Anderson, S.M., Li, B., Pickel, J., McKay, R., Nadal-Ginard, B., Bodine, D.M.*, et al.* (2001). Bone marrow cells regenerate infarcted myocardium. Nature *410*, 701-705.

Ousterout, D.G., Kabadi, A.M., Thakore, P.I., Majoros, W.H., Reddy, T.E., and Gersbach, C.A. (2015a). Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. Nat Commun *6*, 6244.

Ousterout, D.G., Kabadi, A.M., Thakore, P.I., Perez-Pinera, P., Brown, M.T., Majoros, W.H., Reddy, T.E., and Gersbach, C.A. (2015b). Correction of dystrophin expression in cells from Duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. Mol Ther *23*, 523-532.

Ousterout, D.G., Perez-Pinera, P., Thakore, P.I., Kabadi, A.M., Brown, M.T., Qin, X., Fedrigo, O., Mouly, V., Tremblay, J.P., and Gersbach, C.A. (2013). Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. Mol Ther *21*, 1718-1726.

Padgett, R.A. (2012). New connections between splicing and human disease. Trends Genet *28*, 147-154.

Pan, X., Yue, Y., Zhang, K., Hakim, C.H., Kodippili, K., McDonald, T., and Duan, D. (2015). AAV-8 is more efficient than AAV-9 in transducing neonatal dog heart. Hum Gene Ther Methods *26*, 54-61.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat Biotechnol *31*, 839-843.

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

Pavletich, N.P., and Pabo, C.O. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809-817.

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

Pingoud, A., Wilson, G.G., and Wende, W. (2014). Type II restriction endonucleases--a historical perspective and more. Nucleic Acids Res *42*, 7489-7527.

Popplewell, L., Koo, T., Leclerc, X., Duclert, A., Mamchaoui, K., Gouble, A., Mouly, V., Voit, T., Paques, F., Cedrone, F.*, et al.* (2013). Gene correction of a duchenne muscular dystrophy mutation by meganuclease-enhanced exon knock-in. Hum Gene Ther *24*, 692-701.

Port, F., and Bullock, S.L. (2016). Augmenting CRISPR applications in Drosophila with tRNA-flanked sgRNAs. Nat Methods *13*, 852-854.

Porteus, M.H. (2015). Towards a new era in medicine: therapeutic genome editing. Genome Biol *16*, 286.

Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology *151*, 653-663.

Pozsgai, E.R., Griffin, D.A., Heller, K.N., Mendell, J.R., and Rodino-Klapac, L.R. (2017). Systemic AAV-Mediated beta-Sarcoglycan Delivery Targeting Cardiac and Skeletal Muscle Ameliorates Histological and Functional Deficits in LGMD2E Mice. Mol Ther *25*, 855-869.

Prakash, V., Moore, M., and Yanez-Munoz, R.J. (2016). Current Progress in Therapeutic Gene Editing for Monogenic Diseases. Mol Ther *24*, 465-474.

Puchta, H., Dujon, B., and Hohn, B. (1996). Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc Natl Acad Sci U S A *93*, 5055-5060.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S.*, et al.* (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186-191.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y.*, et al.* (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380-1389.

Rass, E., Grabarz, A., Plo, I., Gautier, J., Bertrand, P., and Lopez, B.S. (2009). Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. Nat Struct Mol Biol *16*, 819-824.

Reddy, P., Ocampo, A., Suzuki, K., Luo, J., Bacman, S.R., Williams, S.L., Sugawara, A., Okamura, D., Tsunekawa, Y., Wu, J.*, et al.* (2015). Selective elimination of mitochondrial mutations in the germline by genome editing. Cell *161*, 459-469.

Ren, C., Kumar, S., Shaw, D.R., and Ponnazhagan, S. (2005). Genomic stability of self-complementary adeno-associated virus 2 during early stages of transduction in mouse muscle in vivo. Hum Gene Ther *16*, 1047-1057.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol *30*, 460-465.

Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L., and Corn, J.E. (2016). Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol *34*, 339-344.

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

Sacco, A., Mourkioti, F., Tran, R., Choi, J., Llewellyn, M., Kraft, P., Shkreli, M., Delp, S., Pomerantz, J.H., Artandi, S.E.*, et al.* (2010). Short telomeres and stem cell exhaustion model Duchenne muscular dystrophy in mdx/mTR mice. Cell *143*, 1059-1071.

Saera-Vila, A., Kasprick, D.S., Junttila, T.L., Grzegorski, S.J., Louie, K.W., Chiari, E.F., Kish, P.E., and Kahana, A. (2015). Myocyte Dedifferentiation Drives Extraocular Muscle Regeneration in Adult Zebrafish. Invest Ophthalmol Vis Sci *56*, 4977-4993.

Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.T., and Yamamoto, T. (2016). MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems. Nat Protoc *11*, 118-133.

Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., and Yamamoto, T. (2014). Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep *4*, 5400.

Sambasivan, R., Yao, R., Kissenpfennig, A., Van Wittenberghe, L., Paldi, A., Gayraud-Morel, B., Guenou, H., Malissen, B., Tajbakhsh, S., and Galy, A. (2011). Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. Development *138*, 3647-3656.

Samulski, R.J., and Muzyczka, N. (2014). AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annu Rev Virol *1*, 427-451.

Sander, J.D., Ramirez, C.L., Linder, S.J., Pattanayak, V., Shoresh, N., Ku, M., Foden, J.A., Reyon, D., Bernstein, B.E., Liu, D.R.*, et al.* (2013). In silico abstraction of zinc finger nuclease cleavage profiles reveals an expanded landscape of off-target sites. Nucleic Acids Res *41*, e181.

Sandoval-Guzman, T., Wang, H., Khattak, S., Schuez, M., Roensch, K., Nacu, E., Tazaki, A., Joven, A., Tanaka, E.M., and Simon, A. (2014). Fundamental differences in dedifferentiation and stem cell recruitment during skeletal muscle regeneration in two salamander species. Cell Stem Cell *14*, 174-187.

Sather, B.D., Romano Ibarra, G.S., Sommer, K., Curinga, G., Hale, M., Khan, I.F., Singh, S., Song, Y., Gwiazda, K., Sahni, J.*, et al.* (2015). Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. Sci Transl Med *7*, 307ra156.

Schatzberg, S.J., Olby, N.J., Breen, M., Anderson, L.V., Langford, C.F., Dickens, H.F., Wilton, S.D., Zeiss, C.J., Binns, M.M., Kornegay, J.N.*, et al.* (1999). Molecular analysis of a spontaneous dystrophin 'knockout' dog. Neuromuscul Disord *9*, 289-295.

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

Schnepp, B.C., Jensen, R.L., Chen, C.L., Johnson, P.R., and Clark, K.R. (2005). Characterization of adeno-associated virus genomes isolated from human tissues. J Virol *79*, 14793-14803.

Schon, E.A., DiMauro, S., and Hirano, M. (2012). Human mitochondrial DNA: roles of inherited and somatic mutations. Nat Rev Genet *13*, 878-890.

Seale, P., Sabourin, L.A., Girgis-Gabardo, A., Mansouri, A., Gruss, P., and Rudnicki, M.A. (2000). Pax7 is required for the specification of myogenic satellite cells. Cell *102*, 777-786.

Segal, D.J., Dreier, B., Beerli, R.R., and Barbas, C.F., 3rd (1999). Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. Proc Natl Acad Sci U S A *96*, 2758-2763.

Segal, D.J., and Meckler, J.F. (2013). Genome engineering at the dawn of the golden age. Annu Rev Genomics Hum Genet *14*, 135-158.

Selsby, J.T., Ross, J.W., Nonneman, D., and Hollinger, K. (2015). Porcine models of muscular dystrophy. ILAR J *56*, 116-126.

Sfeir, A., and Symington, L.S. (2015). Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? Trends Biochem Sci *40*, 701-714.

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X.*, et al.* (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat Methods *11*, 399-402.

Shen, S., Bryant, K.D., Brown, S.M., Randell, S.H., and Asokan, A. (2011). Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. J Biol Chem *286*, 13532-13540.

Shimizu-Motohashi, Y., Miyatake, S., Komaki, H., Takeda, S., and Aoki, Y. (2016). Recent advances in innovative therapeutic approaches for Duchenne muscular dystrophy: from discovery to clinical trials. Am J Transl Res *8*, 2471-2489.

Shinin, V., Gayraud-Morel, B., Gomes, D., and Tajbakhsh, S. (2006). Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. Nat Cell Biol *8*, 677-687.

Shmakov, S., Abudayyeh, O.O., Makarova, K.S., Wolf, Y.I., Gootenberg, J.S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K.*, et al.* (2015). Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. Mol Cell *60*, 385-397.

Sicinski, P., Geng, Y., Ryder-Cook, A.S., Barnard, E.A., Darlison, M.G., and Barnard, P.J. (1989). The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. Science *244*, 1578-1580.

Silva, G., Poirot, L., Galetto, R., Smith, J., Montoya, G., Duchateau, P., and Paques, F. (2011). Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. Curr Gene Ther *11*, 11-27.

Simsek, D., Brunet, E., Wong, S.Y., Katyal, S., Gao, Y., McKinnon, P.J., Lou, J., Zhang, L., Li, J., Rebar, E.J.*, et al.* (2011). DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. PLoS Genet *7*, e1002080.

Sipo, I., Fechner, H., Pinkert, S., Suckau, L., Wang, X., Weger, S., and Poller, W. (2007). Differential internalization and nuclear uncoating of self-complementary adeno-associated virus pseudotype vectors as determinants of cardiac cell transduction. Gene Ther *14*, 1319-1329.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84-88.

Smith, B.F., Yue, Y., Woods, P.R., Kornegay, J.N., Shin, J.H., Williams, R.R., and Duan, D. (2011). An intronic LINE-1 element insertion in the dystrophin gene aborts dystrophin expression and results in Duchenne-like muscular dystrophy in the corgi breed. Lab Invest *91*, 216-231.

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

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

Srivastava, A., Lusby, E.W., and Berns, K.I. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. J Virol *45*, 555-564.

Stoddard, B.L. (2005). Homing endonuclease structure and function. Q Rev Biophys *38*, 49-95.

Stoddard, B.L. (2011). Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. Structure *19*, 7-15.

Suda, T., and Liu, D. (2007). Hydrodynamic gene delivery: its principles and applications. Mol Ther *15*, 2063-2069.

Sultana, N., Zhang, L., Yan, J., Chen, J., Cai, W., Razzaque, S., Jeong, D., Sheng, W., Bu, L., Xu, M.*, et al.* (2015). Resident c-kit(+) cells in the heart are not cardiac stem cells. Nat Commun *6*, 8701.

Sun, J.Y., Anand-Jawa, V., Chatterjee, S., and Wong, K.K. (2003). Immune responses to adeno-associated virus and its recombinant vectors. Gene Ther *10*, 964-976.

Sun, N., and Zhao, H. (2013). Transcription activator-like effector nucleases (TALENs): a highly efficient and versatile tool for genome editing. Biotechnol Bioeng *110*, 1811-1821.

Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z.*, et al.* (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature *540*, 144-149.

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

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A.*, et al.* (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407-411.

Tang, L., Zeng, Y., Du, H., Gong, M., Peng, J., Zhang, B., Lei, M., Zhao, F., Wang, W., Li, X.*, et al.* (2017). CRISPR/Cas9-mediated gene editing in human zygotes using Cas9 protein. Mol Genet Genomics *292*, 525-533.

Thomas, K.R., and Capecchi, M.R. (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell *51*, 503-512.

Thomas, K.R., Folger, K.R., and Capecchi, M.R. (1986). High frequency targeting of genes to specific sites in the mammalian genome. Cell *44*, 419-428.

Toth, E., Weinhardt, N., Bencsura, P., Huszar, K., Kulcsar, P.I., Talas, A., Fodor, E., and Welker, E. (2016). Cpf1 nucleases demonstrate robust activity to induce DNA modification by exploiting homology directed repair pathways in mammalian cells. Biol Direct *11*, 46.

Truong, L.N., Li, Y., Shi, L.Z., Hwang, P.Y., He, J., Wang, H., Razavian, N., Berns, M.W., and Wu, X. (2013). Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. Proc Natl Acad Sci U S A *110*, 7720-7725.

Tsai, S.Q., and Joung, J.K. (2016). Defining and improving the genome-wide specificities of CRISPR-Cas9 nucleases. Nat Rev Genet *17*, 300-312.

Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat Biotechnol *32*, 569-576.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P.*, et al.* (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol *33*, 187-197.

Turan, S., Farruggio, A.P., Srifa, W., Day, J.W., and Calos, M.P. (2016). Precise Correction of Disease Mutations in Induced Pluripotent Stem Cells Derived From Patients With Limb Girdle Muscular Dystrophy. Mol Ther *24*, 685-696.

Ungerer, J., and Pakrasi, H.B. (2016). Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. Sci Rep *6*, 39681.

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

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

Valentine, B.A., Cooper, B.J., Cummings, J.F., and deLahunta, A. (1986). Progressive muscular dystrophy in a golden retriever dog: light microscope and ultrastructural features at 4 and 8 months. Acta Neuropathol *71*, 301-310.

Valentine, B.A., Cooper, B.J., de Lahunta, A., O'Quinn, R., and Blue, J.T. (1988). Canine X-linked muscular dystrophy. An animal model of Duchenne muscular dystrophy: clinical studies. J Neurol Sci *88*, 69-81.

van Berlo, J.H., Kanisicak, O., Maillet, M., Vagnozzi, R.J., Karch, J., Lin, S.C., Middleton, R.C., Marban, E., and Molkentin, J.D. (2014). c-kit+ cells minimally contribute cardiomyocytes to the heart. Nature *509*, 337-341.

Vilenchik, M.M., and Knudson, A.G. (2003). Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. Proc Natl Acad Sci U S A *100*, 12871-12876.

Vulin, A., Wein, N., Simmons, T.R., Rutherford, A.M., Findlay, A.R., Yurkoski, J.A., Kaminoh, Y., and Flanigan, K.M. (2015). The first exon duplication mouse model of Duchenne muscular dystrophy: A tool for therapeutic development. Neuromuscul Disord *25*, 827-834.

Walmsley, G.L., Arechavala-Gomeza, V., Fernandez-Fuente, M., Burke, M.M., Nagel, N., Holder, A., Stanley, R., Chandler, K., Marks, S.L., Muntoni, F.*, et al.* (2010). A duchenne muscular dystrophy gene hot spot mutation in dystrophin-deficient cavalier king charles spaniels is amenable to exon 51 skipping. PLoS One *5*, e8647.

Wang, H., Rosidi, B., Perrault, R., Wang, M., Zhang, L., Windhofer, F., and Iliakis, G. (2005a). DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. Cancer Res *65*, 4020-4030.

Wang, H.X., Li, M., Lee, C.M., Chakraborty, S., Kim, H.W., Bao, G., and Leong, K.W. (2017). CRISPR/Cas9-Based Genome Editing for Disease Modeling and Therapy: Challenges and Opportunities for Nonviral Delivery. Chem Rev *117*, 9874-9906.

Wang, J., Exline, C.M., DeClercq, J.J., Llewellyn, G.N., Hayward, S.B., Li, P.W., Shivak, D.A., Surosky, R.T., Gregory, P.D., Holmes, M.C.*, et al.* (2015). Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. Nat Biotechnol *33*, 1256-1263.

Wang, L., Calcedo, R., Bell, P., Lin, J., Grant, R.L., Siegel, D.L., and Wilson, J.M. (2011). Impact of pre-existing immunity on gene transfer to nonhuman primate liver with adeno-associated virus 8 vectors. Hum Gene Ther *22*, 1389-1401.

Wang, L., Calcedo, R., Wang, H., Bell, P., Grant, R., Vandenberghe, L.H., Sanmiguel, J., Morizono, H., Batshaw, M.L., and Wilson, J.M. (2010). The pleiotropic effects of natural AAV infections on liver-directed gene transfer in macaques. Mol Ther *18*, 126-134.

Wang, L., Li, F., Dang, L., Liang, C., Wang, C., He, B., Liu, J., Li, D., Wu, X., Xu, X.*, et al.* (2016). In Vivo Delivery Systems for Therapeutic Genome Editing. Int J Mol Sci *17*.

Wang, Z., Ma, H.I., Li, J., Sun, L., Zhang, J., and Xiao, X. (2003). Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo. Gene Ther *10*, 2105-2111.

Wang, Z., Zhu, T., Qiao, C., Zhou, L., Wang, B., Zhang, J., Chen, C., Li, J., and Xiao, X. (2005b). Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart. Nat Biotechnol *23*, 321-328.

West, S.C. (2003). Molecular views of recombination proteins and their control. Nat Rev Mol Cell Biol *4*, 435-445.

Williams, R.S., Williams, J.S., and Tainer, J.A. (2007). Mre11-Rad50-Nbs1 is a keystone complex connecting DNA repair machinery, double-strand break signaling, and the chromatin template. Biochem Cell Biol *85*, 509-520.

Winand, N., Pradham, D., and Cooper, B. (1994). Molecular characterization of severe Duchenne-type muscular dystrophy in a family of Rottweiler dogs. Molecular mechanism of neuromuscular disease.

Wojtal, D., Kemaladewi, D.U., Malam, Z., Abdullah, S., Wong, T.W., Hyatt, E., Baghestani, Z., Pereira, S., Stavropoulos, J., Mouly, V.*, et al.* (2016). Spell Checking

Nature: Versatility of CRISPR/Cas9 for Developing Treatments for Inherited Disorders. Am J Hum Genet *98*, 90-101.

Wolf, D.P., Hayama, T., and Mitalipov, S. (2017). Mitochondrial genome inheritance and replacement in the human germline. EMBO J *36*, 2177-2181.

Wolff, J.A., Ludtke, J.J., Acsadi, G., Williams, P., and Jani, A. (1992). Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. Hum Mol Genet *1*, 363-369.

Wright, A.V., Nunez, J.K., and Doudna, J.A. (2016). Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell *164*, 29-44.

Wu, Y., Liang, D., Wang, Y., Bai, M., Tang, W., Bao, S., Yan, Z., Li, D., and Li, J. (2013). Correction of a genetic disease in mouse via use of CRISPR-Cas9. Cell Stem Cell *13*, 659-662.

Wyatt, H.D., and West, S.C. (2014). Holliday junction resolvases. Cold Spring Harb Perspect Biol *6*, a023192.

Xia, G., Gao, Y., Jin, S., Subramony, S.H., Terada, N., Ranum, L.P., Swanson, M.S., and Ashizawa, T. (2015). Genome modification leads to phenotype reversal in human myotonic dystrophy type 1 induced pluripotent stem cell-derived neural stem cells. Stem Cells *33*, 1829-1838.

Xie, A., Kwok, A., and Scully, R. (2009). Role of mammalian Mre11 in classical and alternative nonhomologous end joining. Nat Struct Mol Biol *16*, 814-818.

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci U S A *112*, 3570-3575.

Xu, L., Park, K.H., Zhao, L., Xu, J., El Refaey, M., Gao, Y., Zhu, H., Ma, J., and Han, R. (2016). CRISPR-mediated Genome Editing Restores Dystrophin Expression and Function in mdx Mice. Mol Ther *24*, 564-569.

Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G.*, et al.* (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature *514*, 380-384.

Yamada, M., Emmanuele, V., Sanchez-Quintero, M.J., Sun, B., Lallos, G., Paull, D., Zimmer, M., Pagett, S., Prosser, R.W., Sauer, M.V.*, et al.* (2016). Genetic Drift Can Compromise Mitochondrial Replacement by Nuclear Transfer in Human Oocytes. Cell Stem Cell *18*, 749-754.

Yamano, T., Zetsche, B., Ishitani, R., Zhang, F., Nishimasu, H., and Nureki, O. (2017). Structural Basis for the Canonical and Non-canonical PAM Recognition by CRISPR-Cpf1. Mol Cell *67*, 633-645 e633.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013a). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell *154*, 1370-1379.

Yang, J., Li, S.Y., Li, Y.Q., Cao, J.Q., Feng, S.W., Wang, Y.Y., Zhan, Y.X., Yu, C.S., Chen, F., Li, J.*, et al.* (2013b). MLPA-based genotype-phenotype analysis in 1053 Chinese patients with DMD/BMD. BMC Med Genet *14*, 29.

Yang, S., Chang, R., Yang, H., Zhao, T., Hong, Y., Kong, H.E., Sun, X., Qin, Z., Jin, P., Li, S.*, et al.* (2017). CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. J Clin Invest *127*, 2719-2724.

Yin, H., Kanasty, R.L., Eltoukhy, A.A., Vegas, A.J., Dorkin, J.R., and Anderson, D.G. (2014a). Non-viral vectors for gene-based therapy. Nat Rev Genet *15*, 541-555.

Yin, H., Price, F., and Rudnicki, M.A. (2013). Satellite cells and the muscle stem cell niche. Physiol Rev *93*, 23-67.

Yin, H., Xue, W., Chen, S., Bogorad, R.L., Benedetti, E., Grompe, M., Koteliansky, V., Sharp, P.A., Jacks, T., and Anderson, D.G. (2014b). Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. Nat Biotechnol *32*, 551-553.

You, Z., and Bailis, J.M. (2010). DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. Trends Cell Biol *20*, 402-409.

Young, C.S., Hicks, M.R., Ermolova, N.V., Nakano, H., Jan, M., Younesi, S., Karumbayaram, S., Kumagai-Cresse, C., Wang, D., Zack, J.A.*, et al.* (2016). A Single CRISPR-Cas9 Deletion Strategy that Targets the Majority of DMD Patients Restores Dystrophin Function in hiPSC-Derived Muscle Cells. Cell Stem Cell *18*, 533-540.

Young, C.S., Mokhonova, E., Quinonez, M., Pyle, A.D., and Spencer, M.J. (2017). Creation of a Novel Humanized Dystrophic Mouse Model of Duchenne Muscular Dystrophy and Application of a CRISPR/Cas9 Gene Editing Therapy. J Neuromuscul Dis *4*, 139-145.

Yu, A.M., and McVey, M. (2010). Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. Nucleic Acids Res *38*, 5706-5717.

Yu, H.H., Zhao, H., Qing, Y.B., Pan, W.R., Jia, B.Y., Zhao, H.Y., Huang, X.X., and Wei, H.J. (2016). Porcine Zygote Injection with Cas9/sgRNA Results in DMD-Modified Pig with Muscle Dystrophy. Int J Mol Sci *17*.

Yue, Y., Shin, J.H., and Duan, D. (2011). Whole body skeletal muscle transduction in neonatal dogs with AAV-9. Methods Mol Biol *709*, 313-329.

Zaidi, S.S., Mahfouz, M.M., and Mansoor, S. (2017). CRISPR-Cpf1: A New Tool for Plant Genome Editing. Trends Plant Sci *22*, 550-553.

Zechner, C., Lai, L., Zechner, J.F., Geng, T., Yan, Z., Rumsey, J.W., Collia, D., Chen, Z., Wozniak, D.F., Leone, T.C.*, et al.* (2010). Total skeletal muscle PGC-1 deficiency uncouples mitochondrial derangements from fiber type determination and insulin sensitivity. Cell Metab *12*, 633-642.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A.*, et al.* (2015). Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell *163*, 759-771.

Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S.*, et al.* (2017). Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. Nat Biotechnol *35*, 31-34.

Zhang, J.P., Li, X.L., Li, G.H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen, W., Fu, Y.W.*, et al.* (2017a). Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. Genome Biol *18*, 35.

Zhang, Y., Long, C., Bassel-Duby, R., and Olson, E.N. (2018). Myoediting: Toward Prevention of Muscular Dystrophy by Therapeutic Genome Editing. Physiol Rev *98*, 1205-1240.

Zhang, Y., Long, C., Li, H., McAnally, J.R., Baskin, K.K., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2017b). CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice. Sci Adv *3*, e1602814.

Zhang, Y., Li, H., Min, Y.L., Sanchez-Ortiz, E., Huang, J., Mireault, A.A., Shelton, J.M., Kim, J., Mammen, P.P.A., Bassel-Duby, R., and Olson, E.N. (2020). Enhanced CRISPR/Cas9 Correction of Duchenne Muscular Dystrophy In Mice By A Self-Complementary AAV Delivery System. Sci Adv 2.

Zhen, S., Hua, L., Liu, Y.H., Gao, L.C., Fu, J., Wan, D.Y., Dong, L.H., Song, H.F., and Gao, X. (2015). Harnessing the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated Cas9 system to disrupt the hepatitis B virus. Gene Ther *22*, 404-412.

Zhu, P., Wu, F., Mosenson, J., Zhang, H., He, T.C., and Wu, W.S. (2017). CRISPR/Cas9-Mediated Genome Editing Corrects Dystrophin Mutation in Skeletal Muscle Stem Cells in a Mouse Model of Muscle Dystrophy. Mol Ther Nucleic Acids *7*, 31-41.

Zincarelli, C., Soltys, S., Rengo, G., Koch, W.J., and Rabinowitz, J.E. (2010). Comparative cardiac gene delivery of adeno-associated virus serotypes 1-9 reveals that AAV6 mediates the most efficient transduction in mouse heart. Clin Transl Sci *3*, 81-89.

Zuo, Z., and Liu, J. (2016). Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics Simulations. Sci Rep *5*, 37584.