

Development of a
CRISPR-based
gene therapy approach
to correct duplications causing
Duchenne Muscular Dystrophy

Veronica Pini

UCL Great Ormond Street Institute of Child Health

University College London

Thesis submitted to University College London
in fulfilment of the requirements for the title of
Doctor of Philosophy

I, Veronica Pini, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

ACKNOWLEDGMENTS

I wish to truly thank Dr. Francesco Conti and Dr. Helen O'Neill, my supervisors that stood by my side during these three years, and in particular Prof. Francesco Muntoni for the guidance, support and encouragement shown during the realization of this thesis project.

My deepest gratitude goes to Prof. Jennifer Morgan, Dr. Federica Montanaro and Dr. Silvia Torelli for the feedbacks, brainstorming and continuous aids and suggestions throughout this experience. They taught me how to think and act professionally as a researcher.

I would also like to thank Dr. John Counsell for supervision during the production of lentiviral particles and the scientific discussions held together with Dr. Marc Moore, Dr. Jinhong Meng for helping me with planning the iPSCs experimental work and, finally, Dr. Julie Dumonceaux and Dr. Virginie Mariot (collaborators working at the Translational Myology section of our Institute who provided the NEON device), for their undeniable help in pursuing cell electroporation.

Thanks to the whole Dubowitz Neuromuscular Team for becoming my second family and in particular to Sara, Francesco, Mario, Pier and Elena, friends who bore with me and with whom I shared (too many) coffees, achievements and frustrations. Thanks also to my friends Mariagrazia and Nadia, who I have been lucky enough to cross path with: I know you always believed in me. Finally, thanks to Dominic for the daily encouragement and the “stai calma” and to Dimitri for having survived not only my most awful dinners, but also for having listened to my outbursts during lunches and tea-breaks full of gossips.

Thanks to my friends around the world: thanks Emily for the dancing nights with flamingoes and for the intercontinental Skype calls on Saturdays, thanks Giuseppe for the daily updates from Denmark and for being my over-protecting “daddy cool”, and finally thanks Daniela for being as crazy as me and becoming a constant presence in my life.

Thanks to my dear Andrea, amazing singing partner and passionate friend, for sharing with me much more than the joys and doubts of every PhD student and for being such an amazing and caring person.

Thanks to Emma, Nicole and Stefania, the most special acquired sisters I could ever had. You are the friends of a lifetime always ready to be on my side and protect me. Despite the distance, we are part of each other's life. You are the most precious gift.

Thanks to Mayflower House where I had my “writing retreats” and to Matteo, my favourite partner-in-crime in the most absurd, funny and serious situations. You are the only person who can read my mind and understand almost every mood, emotion and doubt I experience at a glance. Without even knowing it, you taught me how to be the best version of myself and that soulmates truly exist.

Thanks to my parents and my sister, my life. They gave me wings to explore the world and design my path, despite the fears of the unknown and the difficulties I could encounter in life. Thanks to them, I know I am stronger than what I thought. They are the ones to which I dedicate this thesis: the “beast” did not win, not even this time.

Thanks to everyone who believed I could have never made it since here, discouraging me in virtue of my limitations and making me question who and what I want to be: I know the answer, more than ever.

Finally, but not least, thanks to those voices who left me too soon but keep speaking in a whisper to my heart. Your memory encourages me to be brave in hard times and never to stop to look for a sincere happiness. You will always be with me.

ABSTRACT

Duchenne Muscular Dystrophy is a severe neurodegenerative disorder caused by deletions, duplications or point mutations in the DMD gene, which encodes dystrophin. In absence of dystrophin, muscle fibres degenerate and patients become wheelchair dependent by their early teens. Cardiac and respiratory muscles are also affected, causing premature death by the third decade of life. Among the approaches currently being tested in clinical trials to treat this disease, none is suitable to permanently restore dystrophin by removing either small or large multi-exon dystrophin duplications, which account for 10-15% of DMD cases.

In this thesis, I designed a genome editing approach to correct duplications in the *DMD* gene by using a single CRISPR/Cas9 target site.

First, I identified a CRISPR/Cas9 nuclease able to efficiently target *DMD* intron 9, which would be suitable for gene editing in patients harbouring *DMD* duplications in the mutational hotspot 2-20¹.

Then, I tested both integrating lentiviral particles and nuclear electroporation as tools to deliver and express CRISPR/Cas9 in patient-derived cells carrying different dystrophin duplications. Patient-derived myoblasts allowed me to assess dystrophin restoration at the genomic, transcriptional and protein level by means of the T7 assay, quantitative-PCR and western blot, respectively.

I confirmed dystrophin correction in transduced as well as electroporated cells expressing CRISPR/Cas9, and I demonstrated that both a constitutive and a transient nuclease expression led to a similar extent of protein restoration of around 50%.

These outcomes allowed me to conclude that CRISPR/Cas9 editing tool is a suitable approach to remove large genomic duplications *in vitro*. Furthermore, the data presented in this thesis provides the basis for the design of new therapeutic approaches to be tested *in vivo* in Duchenne Muscular Dystrophy animal models. These include both *in vivo* CRISPR/Cas9-mediated gene therapy and cell-therapy based on transplantation of *ex vivo* corrected myoblasts expressing corrected wild-type dystrophin.

IMPACT STATEMENT

The work presented in this thesis provides a proof-of-concept of the applicability of a single CRISPR/Cas9 nuclease as a therapeutic strategy for repairing genomic duplications. Specifically, this tool was designed to target duplications arising in dystrophin gene and causative of Duchenne Muscular Dystrophy, one of the most debilitating and severe neuromuscular conditions that leads to a fatal outcome early in life^{1,2}.

I demonstrated that the CRISPR/Cas9 system can remove dystrophin duplication and restore a wild-type dystrophin protein *in vitro* in patient-derived cells with an efficiency of about 50%. These results represented a foundation for the design of further *in vivo* experiments in diseased DMD animal models, either as gene therapy by *in vivo* adeno-associated virus-mediated administration of CRISPR/Cas9 or cell-therapy based on *ex vivo* expansion of edited cells to be transplanted in immunodeficient dystrophic animals.

If successful *in vivo*, the approach I designed will be advantageous compared to the available therapeutic strategies³, as it will be the only existing approach allowing the permanent restoration, in targeted nuclei, of the genomic mutation that is responsible for the production of a mutant non-functional protein.

This strategy can potentially be beneficial for the Duchenne Muscular Dystrophy population carrying dystrophin duplications, which represents 10-15% of patients¹. However, the single nuclease approach can be easily customized to target and remove duplicated regions occurring elsewhere in the genome; genomic duplications are in fact at the basis of a large number of human diseases and syndromes like the Rett syndrome, cystic fibrosis and β -thalassemia, among others^{4,5}. As such, the single CRISPR/Cas9 nuclease approach developed in the present work is relevant beyond the field of neuromuscular disorders and therefore represent a useful tool with a wider applicability.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	I
ABSTRACT	III
IMPACT STATEMENT	IV
TABLE OF CONTENTS	V
FIGURES	X
TABLES	XII
ABBREVIATIONS LIST	XIII
CHAPTER 1. GENERAL INTRODUCTION	1
1.1 MUSCULAR DYSTROPHIES	1
1.2 DUCHENNE MUSCULAR DYSTROPHY	1
1.3 THE <i>DMD</i> GENE	4
1.3.1. DMD GENE CHARACTERIZATION	4
1.3.2. THE MUTATIONAL SPECTRUM OF DMD GENE AND DMD DUPLICATIONS	5
1.4 DYSTROPHIN VERSATILITY: STRUCTURAL AND SIGNALLING ROLES	9
1.5 DMD PATHOGENICITY: INSIGHTS INTO THE MOLECULAR CHANGES FOLLOWING DYSTROPHIN ABSENCE	13
1.5.1. FAULTY REGENERATION AND DEVELOPMENT OF FIBROSIS	13
1.5.2. DYSREGULATED INTRACELLULAR HOMEOSTASIS	14
1.6 MOLECULAR ADVANCES IN DMD TREATMENT	15
1.6.1. DYSTROPHIN-INDEPENDENT STRATEGIES	15
a) Corticosteroids	15
b) Muscle growth strategies	16
c) Utrophin modulation	17
d) Modulation of the nNOS signalling pathway	17
e) Upregulation of components of the ECM	18
1.6.2. DYSTROPHIN-BASED STRATEGIES	18
a) Aminoglycosides	18
b) Exon-skipping	19
c) Cell-based therapies	20
d) Gene-therapies	22
1.7 THE EVOLUTION OF ENGINEERED NUCLEASES	24
1.7.1 MEGANUCLEASES	28
1.7.2. ZINC-FINGER NUCLEASES	29
1.7.3. TRANSCRIPTION ACTIVATOR-LIKE EFFECTOR NUCLEASES (TALENs)	30
1.7.4. CLUSTERED REGULARLY INTERSPACED SHORT PALINDROMIC REPEATS (CRISPR)/Cas9	30
a) The discovery of CRISPR system	30
b) Heterogeneity of CRISPR loci	32

c) CRISPR in action	33
1.8 CRISPR/CAS9 AS GENOME EDITING TOOL	38
1.9 AIM OF THE PROPOSED WORK	41
CHAPTER 2. MATERIALS AND METHODS	43
2.1 MOLECULAR CLONING	43
2.1.1. BACTERIAL MANIPULATION	43
a) Preparation of LB agar plates	43
b) Preparation of LB broth	43
c) Bacterial plating and growth	43
d) Heat-shock transformation of chemically competent bacteria	43
e) Plasmid amplification	44
f) Extraction of plasmid DNA from small volumes of bacterial cultures	44
g) Extraction of plasmid DNA from large volumes of bacterial cultures	45
2.1.2. LIST OF VECTORS	45
a) LentiCRISPRv1 vector	45
b) pL-CRISPR.EFS.GFP vector	45
c) LentiCRISPR-EGFP sgRNA5 vector	46
d) pLJM1-EGFP vector	46
e) psPAX2 vector	46
f) pCMV-VSV-G vector	46
g) pLV-U6g-EPCG vector	47
h) U6gRNA-CMVCas9-GFP vector	47
i) pEGFP-C2 vector	47
j) pCMV-GFP vector	47
2.1.3 PLASMID VERIFICATION	58
2.1.4 BACTERIA STORAGE	59
2.1.5 GENERATION OF FUNCTIONAL CRISPR/Cas9 VECTORS	59
a) Identification of CRISPR/Cas9 genomic target sequences and sgRNAs.	59
b) Digestion of lentiviral plasmids expressing CRISPR/Cas9	59
c) sgRNAs oligos design	60
d) sgRNAs cloning into the LentiCRISPRv1 vector	61
2.1.6 PRODUCTION OF INTEGRATING LENTIVIRAL PARTICLES	61
2.1.7 TITRATION OF GENERATED LENTIVIRAL PARTICLES	64
2.2 CELL CULTURE	65
2.2.1. ETHICS STATEMENT	65
2.2.2. LIST OF HUMAN CELL LINES	66
2.2.3. MATRIGEL PREPARATION AND PLATE COATING	66
2.2.4. CULTURE OF HUMAN EMBRYONIC KIDNEY (HEK 293T) CELLS	66
2.2.5. CULTURE OF HUMAN FIBROBLASTS	67
2.2.6. CULTURE AND DIFFERENTIATION OF HUMAN MYOBLASTS	67
2.2.7. CULTURE AND DIFFERENTIATION OF HUMAN INDUCED PLURIPOTENT STEM CELLS (hiPSCs)	67
2.2.8. CELL PASSAGING	68
2.2.9. COUNT OF VIABLE CELLS	68
2.2.10. CRYOPRESERVATION OF CELLS	68
2.2.11. CELL TRANSFECTION BY USING COMMERCIALY AVAILABLE REAGENTS	69
a) Lipofectamine transfection of HEK 293T cells	69
b) Lipofectamine transfection of primary myoblasts	69

c) TurboFect transfection of primary myoblasts	69
d) GeneJuice transfection of primary myoblasts	70
2.2.12. CELL TRANSFECTION BY USING NUCLEAR ELECTROPORATION	70
a) Nuclear electroporation of human fibroblasts and myoblasts - Amaxa®	70
b) Nuclear electroporation of human myoblasts - NEON®	70
c) Nuclear electroporation of human induced pluripotent stem cells (hiPSCs) - NEON®	71
2.2.13. CELL TRANSDUCTION	71
a) Lentiviral transduction	71
b) Adenoviral transduction	72
2.2.14. FLUORESCENCE ACTIVATED CELL SORTING (FACS)	72
a) GFP sorting	72
b) Sorting of iPSC-derived myoblast-like cells expressing ErbB3	72
2.3. GENERAL MOLECULAR BIOLOGY PROCEDURES	73
2.3.1. EXTRACTION OF GENOMIC DNA FROM HUMAN CELL LINES	73
2.3.2. DNA AND RNA QUANTIFICATION	73
2.3.3. PRIMERS DESIGN	73
2.3.4. TOUCHDOWN PCR	73
2.3.5. PREPARATION OF AGAROSE GELS AND GEL ELECTROPHORESIS	75
2.3.6. DNA PURIFICATION	75
a) DNA purification from PCR	75
b) DNA purification from agarose gel	76
2.3.7. T7 NUCLEASE ASSAY	76
2.3.8. TBE GEL ELECTROPHORESIS	77
2.3.9. DETERMINATION OF CRISPR/Cas9 CLEAVAGE EFFICIENCY	77
2.3.10. EXTRACTION OF RNA FROM HUMAN CELL LINES	79
2.3.11. RETRO-TRANSCRIPTION	79
a) Retro-transcription (SuperScript™ VILO™ cDNA Synthesis kit)	79
b) Retro-transcription (High-Capacity RNA-to-cDNA™ Kit)	79
2.3.12. REVERSE-TRANSCRIPTASE (RT)-PCR	80
2.3.13. QUANTITATIVE PCR	80
a) One-step qPCR	80
b) Two-step qPCR	81
2.3.14. $\Delta\Delta C_T$ COMPARATIVE QPCR ANALYSIS	83
2.3.15. CELL FIXATION AND PERMEABILIZATION	83
2.3.16. MYOGENICITY ASSAY	83
2.3.17. CELL IMMUNOSTAINING	84
a) Desmin immunostaining	84
b) MyoD immunostaining	84
2.3.18. PHALLOIDIN STAINING	84
2.3.19. PROTEIN EXTRACTION	84
2.3.20. PROTEIN QUANTIFICATION	85
2.3.21. PROTEIN ELECTROPHORESIS	86
2.3.22. WESTERN TRANSFER AND IMMUNOBLOT	86
a) Turbo Blot	86
b) Wet Blot	86
2.3.23. MICROSCOPY AND IMAGE CAPTURE	87
2.3.24. STATISTICAL ANALYSIS	87

CHAPTER 3. DESIGN AND VALIDATION OF CRISPR/CAS NUCLEASES TARGETING DUPLICATIONS IN DMD GENE

3.1 INTRODUCTION	89
3.1.1 THE ORIGIN OF GENOMIC DUPLICATIONS	89
3.1.2 PATHOGENIC GENE DUPLICATIONS AND DUCHENNE MUSCULAR DYSTROPHY	90
3.1.3 THERAPEUTIC APPROACHES TO DYSTROPHIN DUPLICATIONS	90
3.2 AIMS	92
3.3 RESULTS	94
3.3.1. SELECTION OF GENOMIC TARGETS WITHIN DYSTROPHIN	94
3.3.2. DESIGN OF CRISPR/Cas9 NUCLEASES TARGETED TO DYSTROPHIN INTRON 9	96
a) sgRNA design	96
b) Generation of functional CRISPR/Cas9 nucleases	98
3.3.3 LENTICRISPRs TRANSFECTION INTO HEK 293T CELLS	102
3.3.4. DETERMINING TARGETING SPECIFICITY AND EFFICIENCY OF CRISPR/Cas9 INTO HEK 293T CELLS	104
3.4 DISCUSSION	109

CHAPTER 4. DELIVERY OF CRISPR/CAS9 NUCLEASES TO PATIENT-DERIVED FIBROBLASTS CARRYING DYSTROPHIN DUPLICATIONS **114**

4.1 INTRODUCTION	114
4.1.1. FIBROBLASTS AS A CELLULAR MODEL TO PROVE CRISPR/Cas9 EFFICIENCY	114
4.1.2. MYOD: FORCED FIBROBLAST TRANSITION TO MYOGENIC SPECIFICATION	114
4.2 AIMS	116
4.3 RESULTS	118
4.3.1. GENERATION OF LENTIVIRAL PARTICLES EXPRESSING THE BEST LENTICRISPRs	118
4.3.2. TRANSDUCTION OF HEK 293T CELLS	119
4.3.3. TRANSDUCTION OF PRIMARY FIBROBLASTS	122
a) Transduction of primary fibroblasts with in-house produced lentiviral vectors	122
b) Transduction of primary fibroblasts with commercial lentiviral particles	126
4.3.4. NUCLEAR ELECTROPORATION OF PRIMARY FIBROBLASTS	130
a) Validation of CRISPR2 σ and CRISPR0 in HEK 293T cells via Lipofectamine transfection	131
b) Nucleoporation of CRISPR2 σ and CRISPR0 in primary fibroblasts	134
4.3.5. MYOD TRANSDUCTION	140
4.4 DISCUSSION	145

CHAPTER 5. DELIVERY OF CRISPR/CAS9 NUCLEASES TO PATIENT-DERIVED MYOBLASTS AND RESTORATION OF DYSTROPHIN DUPLICATIONS **149**

5.1 INTRODUCTION	149
5.1.1. MYOBLASTS: THE ANCESTORS OF THE MUSCLE FIBER	149
5.1.2. PATIENT-DERIVED MYOBLASTS: DOWNSIDES AND BENEFITS	149
5.2 AIMS	151
5.3 RESULTS	153
5.3.1. SELECTION OF MYOGENIC MYOBLAST CELL LINE	153
5.3.2. CRISPR/Cas9 DELIVERY IN PATIENT-DERIVED MYOBLASTS: VIRAL APPROACH	156
a) Verification of CRISPR/Cas9-mediated cleavage at genomic level	159
b) Verification of dystrophin transcript repair following CRISPR/Cas9 activity	159
c) Verification of dystrophin protein repair following CRISPR/Cas9 activity in DUP2 cell line	167
5.3.3. CRISPR/Cas9 DELIVERY IN PATIENT-DERIVED MYOBLASTS: NON-VIRAL APPROACHES	170
a) Transfection by using commercially available transfection reagents	170

b) Myoblast electroporation: Amaxa versus NEON	176
b.i) Verification of CRISPR/Cas9-mediated cleavage at genomic level	179
b.ii) Verification of dystrophin transcript repair following CRISPR/Cas9 activity in DUP2i cells	185
b.iii) Verification of dystrophin protein repair following CRISPR/Cas9 activity	185
c) Contribution of Cas9 expression levels to dystrophin editing	187
c.i) Assessing CRISPR/Cas9-mediated cleavage efficiency at genomic level	187
c.ii) Verification of dystrophin transcript repair following CRISPR/Cas9 activity	191
c.iii) Verification of dystrophin repair following CRISPR/Cas9 activity	191
5.4 DISCUSSION	194

CHAPTER 6. ADDRESSING SINGLE CRISPR/CAS9 NUCLEASES TO LARGE DYSTROPHIN DUPLICATIONS IN PATIENT-DERIVED INDUCED PLURIPOTENT STEM CELLS **198**

6.1 INTRODUCTION	198
6.1.1. THE ADVENT OF INDUCED PLURIPOTENT STEM CELLS	198
6.1.2. MIMICKING DEVELOPMENTAL MYOGENESIS TO INDUCE MYOGENIC SPECIFICATION IN IPSCs	199
6.2 AIMS	200
6.3 RESULTS	203
6.3.1 NEON ELECTROPORATION OF IPSCs WITH CRISPR2 ϵ	203
6.3.2. DIFFERENTIATION OF IPSCs INTO MYOBLASTS	205
6.3.3. ANALYSIS OF MYOGENIC MARKERS EXPRESSION FOLLOWING IPSCs MYOGENIC INDUCTION	209
6.3.4. NEON ELECTROPORATION OF MYOBLAST-LIKE CELLS DERIVED FROM IPSCs	212
6.3.5. T7 ASSAY ON ELECTROPORATED AND FACS SORTED IPSC-DERIVED MYOBLAST-LIKE CELLS	215
6.3.6. INDUCING TERMINAL DIFFERENTIATION OF IPSCs-DERIVED MYOBLAST-LIKE CELLS	215
6.4 DISCUSSION	218

CHAPTER 7. GENERAL DISCUSSION **221**

7.1 THE ROLE OF CRISPR/CAS9 AS A THERAPEUTIC STRATEGY FOR DMD	221
7.2 CRISPR/CAS9 AND DYSTROPHIN EDITING: THE IMPORTANCE OF THE STUDY DESIGN	223
7.3 IN VITRO CRISPR/CAS9 GOES IN VIVO: FUTURE PERSPECTIVES	226
7.3.1 GENOME-WIDE ANALYSIS OF OFF-TARGET EFFECTS	226
7.3.2. CONSIDERATIONS FOR FUTURE CRISPR/CAS9-BASED THERAPIES	227
7.3.3. <i>IN VIVO</i> CRISPR/CAS9 IMMUNITY: FURTHER CONSIDERATIONS	229
7.4 CONCLUDING REMARKS	230

BIBLIOGRAPHY **231**

APPENDIX I: FORMULATION OF CULTURE MEDIUM USED DURING IPSCS DIFFERENTIATION **283**

FIGURES

FIGURE 1.1. SCHEMATIC VIEW OF DMD GENE AND DYSTROPHIN ISOFORMS.	7
FIGURE 1.2. DYSTROPHIN STRUCTURE AND INTERACTIONS.	12
FIGURE 1.3. SCHEMATIC OF DNA-REPAIR PATHWAYS.	26
FIGURE 1.4. GENOME EDITING USING ENGINEERED NUCLEASES.	27
FIGURE 1.5. SCHEMATIC OF THE FOUR NUCLEASES CLASSES.	35
FIGURE 1.6. CRISPR LOCUS IN A BACTERIAL CHROMOSOME.	36
FIGURE 1.7. CRISPR/CAS TYPE II SYSTEM OF STREPTOCOCCUS PYOGENES.	37
FIGURE 1.8. CRISPR/CAS9-MEDIATED EDITING.	40
FIGURE 1.9. SCHEMATIC REPRESENTATION OF CRISPR/CAS9 MEDIATED DUPLICATION REPAIR IN DMD GENE.	42
FIGURE 2.1. SCHEMATIC OF LENTIVIRAL LENTICRISPRV1 PLASMID.	48
FIGURE 2.2. SCHEMATIC OF PL-CRISPR.EFS.GFP PLASMID.	49
FIGURE 2.3. SCHEMATIC OF LENTICRISPR-EGFP SGRNA5 PLASMID.	50
FIGURE 2.4. SCHEMATIC OF PLJM1-EGFP.	51
FIGURE 2.5. SCHEMATIC OF PSPAX2 PACKAGING PLASMID.	52
FIGURE 2.6. SCHEMATIC OF PCMV-VSV-G PLASMID.	53
FIGURE 2.7. SCHEMATIC OF THE PLV-U6G-EPCG LENTIVIRAL VECTOR.	54
FIGURE 2.8. SCHEMATIC OF THE PCMV-CAS9-GFP VECTOR.	55
FIGURE 2.9. SCHEMATIC OF THE PEGFP-C2 VECTOR.	56
FIGURE 2.10. SCHEMATIC OF THE PCMV-GFP VECTOR.	57
FIGURE 2.11. SCHEMATIC OF RESTRICTION DIGEST.	58
FIGURE 2.12. SCHEMATIC OF LENTIVIRAL PARTICLES PRODUCTION.	63
FIGURE 2.13. SCHEMATIC OF MISMATCH NUCLEASE ASSAY TO DETERMINE THE EFFICIENCY OF CRISPR CLEAVAGE.	78
FIGURE 2.14. QUANTIFICATION OF NEURONAL AND MYOGENIC MARKERS IN IPSCS INDUCED TOWARDS MYOGENIC DIFFERENTIATION: QPCR PLATE SET-UP.	82
FIGURE 3.1. SCHEMATIC OF THE EXPERIMENTAL STEPS REQUIRED FOR CRISPR/CAS9 DESIGN AND VALIDATION.	93
FIGURE 3.2. DESIGN OF SGRNAS TARGETED TO HUMAN DMD INTRON 9.	97
FIGURE 3.3. LENTICRISPRV1 PLASMID DIGESTION.	99
FIGURE 3.4. SCHEMATIC OF SGRNA PRIMERS DESIGN.	101
FIGURE 3.5. HEK 293T CELLS TRANSFECTION EFFICIENCY.	103
FIGURE 3.6. TOUCHDOWNPCR OF LENTICRISPR TARGET SITES.	105
FIGURE 3.7. ANALYSIS OF CRISPR/CAS9 EFFICIENCY IN HEK 293T CELLS.	108
FIGURE 3.8. POSSIBLE SCENARIOS DERIVED FROM NUCLEASES DIRECTED AT GENE DUPLICATIONS.	113
FIGURE 4.1. SCHEMATIC OF THE EXPERIMENTAL STEPS REQUIRED FOR RESTORING DYSTROPHIN IN PATIENT-DERIVED FIBROBLASTS.	117
FIGURE 4.2. T7 ASSAY ON TRANSDUCE HEK 293T CELLS.	121
FIGURE 4.3. PHALLOIDIN STAINING OF PRIMARY FIBROBLASTS.	123
FIGURE 4.4. T7 ASSAY ON TRANSDUCE HUMAN-DERIVED FIBROBLASTS.	125
FIGURE 4.5. TRANSDUCTION OF LENTIDMD2 IN PATIENT FIBROBLASTS.	127
FIGURE 4.6. COMPARISON OF TRANSDUCTION EFFICIENCY BETWEEN LENTIVIRAL PARTICLES EXPRESSING LENTIDMD2 AND PLMJ1-EGFP.	129
FIGURE 4.7. GFP FLUORESCENCE MONITORED ON TRANSFECTED HEK 293T CELLS.	132
FIGURE 4.8. T7 ASSAY PERFORMED ON HEK 293T CELLS TRANSFECTED WITH THE PCMV-CAS9-GFP VECTORS.	133
FIGURE 4.9. SCREENING OF ELECTROPORATION PROGRAMS IN PRIMARY CONTROL FIBROBLASTS.	135
FIGURE 4.10. T7 ASSAY PERFORMED ON THE TOTAL POPULATION OF PATIENT-DERIVED FIBROBLASTS ELECTROPORATED WITH THE PCMV-CAS9-GFP VECTORS.	136
FIGURE 4.11. T7 ASSAY PERFORMED ON THE GFP-POSITIVE PATIENT-DERIVED FIBROBLASTS ELECTROPORATED WITH THE PCMV-CAS9-GFP VECTORS.	138
FIGURE 4.12. T7 ASSAY PERFORMED ON THE TOP 20% GFP-POSITIVE PATIENT-DERIVED FIBROBLASTS ELECTROPORATED WITH THE PCMV-CAS9-GFP VECTORS.	139

FIGURE 4.13. FIBROBLAST TRANSDUCTION WITH AD(RGD)-GFP-H-MYOD1 VIRAL VECTOR.	141
FIGURE 4.14. THREE DIFFERENTIATION CONDITIONS TESTED IN TRANSDUCED FIBROBLASTS (MOI 100).	143
FIGURE 4.15. MYOD AND DESMIN IMMUNOSTAINING.	144
FIGURE 5.1. SCHEMATIC OF THE EXPERIMENTAL STEPS REQUIRED FOR RESTORING DYSTROPHIN IN PATIENT-DERIVED MYOBLASTS.	152
FIGURE 5.2. MYOTUBES ORIGINATED FROM TERMINALLY DIFFERENTIATED DMD1 MYOBLASTS.	154
FIGURE 5.3. MYOGENICITY OF DUP2 MYOBLASTS.	155
FIGURE 5.4. SELECTION OF THE BEST TRANSDUCING MOI.	157
FIGURE 5.5. FACS SORTING OF TRANSDUCED DUP2 MYOBLASTS.	158
FIGURE 5.6. T7 ASSAY ON TRANSDUCED DUP2 PATIENT-DERIVED MYOBLASTS.	160
FIGURE 5.7. TERMINAL DIFFERENTIATION OF DUP2 CELLS.	161
FIGURE 5.8. IDENTIFICATION OF CORRECTED DYSTROPHIN TRANSCRIPT VIA RT-PCR.	164
FIGURE 5.9. QUANTIFICATION OF DYSTROPHIN TRANSCRIPT CORRECTION IN TRANSDUCED DUP2 CELLS.	165
FIGURE 5.10. EFFECT OF CRISPR/CAS9 TRANSDUCTION ON MYOBLAST MYOGENICITY.	166
FIGURE 5.11. VALIDATION OF DYSTROPHIN REPAIR IN TRANSDUCED DUP2I MYOBLASTS.	169
FIGURE 5.12. LIPOFECTAMINE TRANSFECTION OF PRIMARY MYOBLASTS. IDENTIFICATION OF (LIPOFECTAMINE 2000).	171
FIGURE 5.13. FACS ANALYSIS OF PATIENT-DERIVED MYOBLASTS TRANSFECTED VIA LIPOFECTAMINE 2000 (LF2000).	172
FIGURE 5.14. IDENTIFICATION OF THE BEST TRANSFECTION CONDITION (TURBOFECT AND GENEJUICE).	174
FIGURE 5.15. FACS ANALYSIS OF PATIENT-DERIVED MYOBLASTS TRANSFECTED VIA TURBOFECT (TF) AND GENEJUICE (GJ).	175
FIGURE 5.16. ELECTROPORATION OF CRISPR2 Σ IN DUP2 PATIENT-DERIVED MYOBLASTS (AMAXA DEVICE).	177
FIGURE 5.17. ELECTROPORATION OF CRISPR2 Σ IN DUP2 PATIENT-DERIVED MYOBLASTS (NEON DEVICE).	178
FIGURE 5.18. T7 ASSAY ON ELECTROPORATED DUP2 PATIENT-DERIVED MYOBLASTS (NEON).	180
FIGURE 5.19. FACS ANALYSIS OF PATIENT-DERIVED DUP2 MYOBLASTS ELECTROPORATED BY NEON.	181
FIGURE 5.20. FACS ANALYSIS OF IMMORTALIZED DUP2I MYOBLASTS ELECTROPORATED BY NEON.	183
FIGURE 5.21. T7 ASSAY ON ELECTROPORATED DUP2I MYOBLASTS (NEON) EXPRESSING CRISPR/CAS9.	184
FIGURE 5.22. VALIDATION OF DYSTROPHIN REPAIR IN ELECTROPORATED DUP2I MYOBLASTS.	186
FIGURE 5.23. FACS ANALYSIS OF DUP2I MYOBLASTS ELECTROPORATED BY NEON.	188
FIGURE 5.24. T7 ASSAY ON ELECTROPORATED DUP2I MYOBLASTS (NEON) EXPRESSING LOW AND HIGH CRISPR/CAS9 LEVELS.	189
FIGURE 5.25. T7 ASSAY ON ELECTROPORATED DUP2I MYOBLASTS (NEON) EXPRESSING LOW AND HIGH CRISPR/CAS9 LEVELS.	190
FIGURE 5.26. QUANTIFICATION OF DYSTROPHIN TRANSCRIPT CORRECTION IN ELECTROPORATED DUP2I CELLS.	192
FIGURE 5.27. VALIDATION OF DYSTROPHIN REPAIR IN ELECTROPORATED DUP2I MYOBLASTS EXPRESSING EITHER HIGH OR LOW CRISPR/CAS9 LEVELS.	193
FIGURE 6.1. SCHEMATIC OF THE EXPERIMENTAL STEPS REQUIRED FOR RESTORING DYSTROPHIN IN IPSCS.	202
FIGURE 6.2. ELECTROPORATION OF DUP3 IPSCS WITH NEON ELECTROPORATION SYSTEM.	204
FIGURE 6.3. MYOGENIC DIFFERENTIATION OF DUP3 IPSCS CELLS SEEDED AT LOW DENSITY.	206
FIGURE 6.4. MYOGENIC DIFFERENTIATION OF DUP3 IPSCS CELLS SEEDED AT MEDIUM DENSITY.	207
FIGURE 6.5. MYOGENIC DIFFERENTIATION OF DUP3 IPSCS CELLS SEEDED AT HIGH DENSITY.	208
FIGURE 6.6. PAX7 AND PAX6 EXPRESSION UPON MYOGENIC DIFFERENTIATION OF DMD-DERIVED IPSCS.	210
FIGURE 6.7. MYOGENIN AND MYOSIN EXPRESSION UPON MYOGENIC DIFFERENTIATION OF DMD-DERIVED IPSCS CULTURED WITH 6 μ M CHIR99021.	211
FIGURE 6.8. NEON ELECTROPORATION OF DUP3 IPSCS-DERIVED MYOBLASTS.	213

FIGURE 6.9.FACS ANALYSIS OF ELECTROPORATED IPSCS-DERIVED PATIENT MYOBLASTS.	214
FIGURE 6.10.T7 ASSAY ON ELECTROPORATED MYOBLAST-LIKE CELLS DERIVED FROM DUP3 IPSCS.	216
FIGURE 6.11.SELECTION OF IPSCS-DERIVED MYOBLAST-LIKE CELLS EXPRESSING ERBB3.	217

TABLES

TABLE 1.1.DISTRIBUTION OF MUTATIONS WITHIN DMD GENE.	8
TABLE 1.2. THERAPEUTIC DESIGNS FOR DMD TREATMENT.	15
TABLE 1.3.SCHEMATIC OF FEATURES DISTINGUISHING THE MAIN CLASSES OF VIRAL VECTORS	22
TABLE 1.4.CLASSIFICATION OF CRISPR SYSTEM.	32
TABLE 2.1. HUMAN CELL LINES.	66
TABLE 2.2. MEDIA USED FOR INDUCING MYOGENIC DIFFERENTIATION IN HIPSCS.	67
TABLE 3.1.RANGE OF DUPLICATIONS IN DMD PATIENT CELLS LINES STORED IN THE MRC CENTRE FOR NEUROMUSCULAR DISEASES BIOBANK.	94
TABLE 3.2. LIST OF CONTROLS INCLUDED FOR CRISPR/CAS9 VALIDATION IN HEK 293T CELLS	102
TABLE 4.1. LIST OF CONTROLS INCLUDED FOR CRISPR/CAS9 TRANSDUCTION IN PRIMARY DMD FIBROBLASTS.	118
TABLE 4.2. LIST OF CONTROLS INCLUDED FOR CRISPR/CAS9 ELECTROPORATION IN PRIMARY DMD FIBROBLASTS.	130
TABLE 5.1.LIST OF CONTROLS INCLUDED IN THE RT-PCR TO ASSESS REPAIR OF DYSTROPHIN TRANSCRIPT.	159
TABLE 5.2.LIST OF CONTROLS INCLUDED IN THE WESTERN BLOT TO ASSESS REPAIR OF DYSTROPHIN PROTEIN.	167
TABLE 5.3. MYOBLASTS TRANSFECTION EFFICIENCY ASSESSED BY MICROSCOPIC ANALYSIS.	173
TABLE 6.1. LIST OF CONTROLS USED FOR THE ELECTROPORATION OF DUP3-DERIVED MYOBLAST-LIKE CELLS WITH CRISPR/CAS9 PLASMIDS.	212

ABBREVIATIONS LIST

AAV	adeno-associated virus
ABD	actin binding domain
ATP	adenosine triphosphate
bFGF	basal fibroblast growth factor
BMD	Becker muscular dystrophy
BMP	bone morphogenetic protein
bp	base pairs
BSA	bovine serum albumin
CAR	coxsackievirus and adenovirus receptor
CK	creatine kinase
CMV	cytomegalovirus
CRISPR	clustered regularly interspaced short palindromic repeats
Ct	cycle threshold
DAPC	dystrophin-associated protein complex
DHPR	dihydropyridine receptor
DMD	Duchenne muscular dystrophy
<i>DMD</i>	gene coding for dystrophin
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulphoxide
DNA-PKcs	DNA-dependent protein kinase
dsDNA	double-stranded DNA
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EF1 α	elongation factor 1 α
EFS	EF1 α short
EGFP	enhanced green fluorescent protein
ESCs	embryonic stem cells
FACS	fluorescence activated cell sorting
FBS	fetal bovine serum
FDA	Food and Drug Administration
FGF	fibroblast growth factor

GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein
Grb2	growth factor receptor-bound2
GSK-3 β	glycogen synthase kinase 3 β
hALB	human albumin
HDR	homology directed repair
HEK	human embryonic kidney
HGF	hepatocyte growth factor
hiPSCs	human induced pluripotent stem cells
HIV	human immunodeficiency virus
IGF	insulin-like growth factor
indels	insertions/deletions
iNOS	inducible nitric oxide synthase
iPSCs	induced pluripotent stem cells
IST	insulin/ selenite/ transferrin
LB	Luria-Bertani
LTR	long terminal repeats
MN	meganuclease
MOI	multiplicity of infection
NEAA	non-essential amino-acids
NGS	next-generation sequencing
NHEJ	non-homologous ends joining
nNOS	neuronal nitric oxide synthase
NO	nitric oxide
PAM	protospacer adjacent motif
PAR	pseudo-autosomal region
PBS	phosphate buffer saline
PCR	polymerase chain reaction
PVDF	polyvinylidene difluoride
qPCR	quantitative PCR
rAAV	recombinant adeno-associated virus
ROCK	Rho kinase

ROS	reactive oxygen species
RT-PCR	reverse-transcriptase PCR
RYR	ryanodine receptor
S.O.C.	super optimal broth with catabolite repression
SDHA	succinate dehydrogenase complex flavoprotein subunit A
sgRNA (or gRNA)	guide RNA
Shh	Sonic hedgehog
ssDNA	single stranded DNA
TALE	transcription activator-like effector
TALENs	transcription activator-like effector nucleases
TBE	tris borate EDTA
TBS	tris-buffered saline
TBS-T	tris-buffered saline supplemented with Tween
tdPCR	touchdown PCR
tEF1 α	truncated EF1 α
TGF	transforming growth factor
T _m	melting temperature
VSV-g	vesicular stomatis virus G protein
XRCC4	x-ray cross-complementation group 4
ZFN	zinc-finger nuclease

Chapter 1. General introduction

1.1 MUSCULAR DYSTROPHIES

The term muscular dystrophy defines a broad group of rare and monogenic inherited neuromuscular disorders characterized by the progressive loss of muscle function and strength⁶. These symptoms are caused by muscle degeneration and its replacement with adipose and fibrotic scar tissue⁷.

Mutations in genes involved in maintaining muscle function and homeostasis have been associated with different muscular dystrophy classes and subgroups, reviewed by Mercuri and Muntoni⁸. Each subtype is characterized by a specific age at disease onset, incidence in the population and inheritance pattern. The clinical severity of each muscular dystrophy associated with a recessive inheritance broadly correlates with the expression pattern of the mutated gene and the role played by the corresponding protein within muscle groups. The degree of weakness may range from partial mobility impairment to a severe diffuse weakness that also affects the cardiorespiratory systems, leading to premature death⁸.

This thesis will focus on one specific type of muscular dystrophy known as Duchenne Muscular Dystrophy (DMD), caused by mutations in the gene coding for dystrophin. DMD is extensively studied by many research groups as it is one of the most common dystrophies and it is associated with a short lifespan.

1.2 DUCHENNE MUSCULAR DYSTROPHY

DMD was first reported in the first decades of 1800s in the works of Bell⁹, Conte and Gioja¹⁰. However, the complete description of the disease has to be ascribed to Edward Meryon¹¹ and, soon after, to Guillaume Benjamin Amand Duchenne (from whom the disease gets the name)^{12,13}.

Despite at that time the molecular cause of the disease being unknown, these scientists identified one of the hallmarks of the disease, i.e. the progressive weakening of specific muscle groups followed by gradual atrophy, which impeded movements for which muscle force was needed.

To date, DMD affects roughly 1 in 5000 individuals worldwide¹. As it is transmitted through X-linked recessive inheritance, this disease affects almost exclusively males (although affected females have been rarely reported, usually as a result of skewed X-inactivation)^{12,14–16}.

DMD is not manifest at birth and, generally, affected infants remain asymptomatic until the 2nd-4th year of life. A recent study by Ricotti *et al.*¹⁷ indicated that, in the UK, the mean age of diagnosis is around 4.5 years, although symptoms are typically present for a few years before the diagnosis is made. The disease only manifests when a significant part of skeletal muscle, damaged by the pathological process, is lost due to the progressive degenerative process.

The levels of several serum enzymes indicative of ongoing muscle damage are significantly elevated from birth¹⁸. This indicates that muscle abnormalities that result in muscle degeneration and efflux of enzymes from muscle fibre are already present in the preclinical stage of the disease¹⁹. Particularly impressive is the change in the level of serum creatine kinase (CK), that can reach levels of > 50,000 IU/L versus a normal amount of < 200 IU/L²⁰. This enzyme is particularly abundant in skeletal and cardiac muscle²¹ tissue, where it mediates the reversible transfer of a phosphate group from adenosine triphosphate to creatine, exploited in muscle contraction. Such a big change that can exceed 100 folds of normal values is indicative of a severe muscle fibre injury and necrosis, and may also reflect an impaired sarcolemmal barrier²². For this reason, in some countries the screen for CK levels in neonates is offered as part of a newborn screening programme²³.

Muscle weakness in DMD is typically distributed symmetrically, progressively affecting from proximal to distal muscles in a highly selective fashion, as is particularly evident at the beginning of the disease. In the early clinical stages, the muscles of lower limbs are predominantly affected. Because of that, DMD patients show a peculiar waddling and unsteady ambulatory pattern (also observed in most muscular dystrophies) with an increased tendency to fall. Moreover, due to diffuse weakness of proximal muscles, simple movements like climbing stairs are difficult. Rising from the floor requires a specific tactic first reported by William R. Gowers (and therefore known as Gower's manoeuvre), which involves the use of upper limbs to climb up and reach an upward position from a squat²⁴. A further DMD hallmark observed in the early disease stage is the enlargement of calves, mostly derived from the replacement of functional muscle with scar tissue and fatty infiltration²⁵. This feature is recognized as pseudohypertrophy. However, the initial presence of a real hypertrophy that compensates for the weakness in other compromised muscles has been hypothesized²⁵. Occasionally, other muscles (predominantly limb muscles) might appear enlarged²⁶.

To counteract the weakness of hip extensors and gluteal muscles, DMD patients tend to develop lumbar lordosis that forces them to adopt an arcuate posture. Pronounced winging of scapulae also becomes apparent²⁷. Toe walking and equine deformity of the

feet are often reported in association with Achilles tendon contractures and the involvement of distal leg muscles, respectively, as the muscle wasting evolves. The stretch reflexes are typically suppressed²⁸.

After a plateau phase reached around the age of 5-7, in which the disease can remain relatively stable, the continuous decrease of muscle force finally leads DMD patients to become wheelchair dependent by 13 years of age²⁹. The confinement to a sitting position reduces the bone density and further aggravates the chest deformity and curvature of the spine (kyphoscoliosis) caused by the imbalance of affected trunk muscles³⁰.

The pulmonary function, not particularly compromised until ambulation is lost, is worsened due to the progressive weakening of respiratory muscles (such as diaphragm and intercostal muscles). Maximal inspiratory and expiratory pressures gradually reduce and the total lung capacity drops until night-time hypoventilation occurs³¹. Interestingly, Kim *et al.* recently showed that CK values might also play a role in predicting loss of lung functionality³². The progressive contractures that develop as the disease progresses further limit joint movement (elbows, knees, hips, hamstrings, shoulders and wrists)^{33,34}.

Progressive decline in the left ventricular function and dilated cardiomyopathy appear as a consequence of cardiac muscle involvement and predisposes these patients to cardiorespiratory failure³⁵. In fact, cardiac involvement is an important mortality factor in DMD and, similarly to skeletal muscle, the examination of post-mortem tissue revealed variability in the size of cardiac muscle fibres, the replacement of cardiac tissue with connective and fatty tissue and prominent fibrosis, especially in the outer ventricular walls³⁶. Therefore, once cardiac abnormality is detected, the use of angiotensin-converting enzyme (ACE) inhibitors (often in association with low doses of β -blockers), is recommended because of its cardio-protective effect^{37,38}.

50-70% of DMD cases also manifest a specific cognitive profile that involves reduced attention and impaired speech development, probably because of the lack of brain dystrophin isoforms^{39,40}. The brain dystrophin isoforms are expressed at an early developmental stage, and the intellectual impairment derived from their absence does not progress with time. Differences in cognitive profiles have been associated with the dystrophin isoforms affected by patient-specific mutations. Generally, mutations that affect all dystrophin isoforms inevitably result in a severe mental retardation⁴¹. Altered cognitive profiles have been reported both in absence of the brain-specific⁴² and general dystrophin isoforms⁴³ (section 1.3.2). Behavioural and emotional disturbances (such as attention deficit hyperactivity disorder) are also reported to be more common in boys

with mutations affecting both these isoforms⁴⁴, and often associate with depression and anxiety ⁴⁵.

The lack of dystrophin in smooth muscle is often the cause of constipation and gastric dilatation⁴⁶. The vascular system has shown to be affected and vasoconstriction is often causative of muscle ischemia⁴⁷. Occasionally, type 2 diabetes, bladder paralysis and renal failure have been also reported^{48–51}.

Despite DMD being an extremely debilitating pleiotropic disease, the life expectancy of affected patients has improved over time. Nowadays, the implementation of standards of care and interventions such as the use of corticosteroids in the ambulant phase, a stringent cardiac surveillance and non-invasive mechanical ventilation^{52,53} in the later stage of the diseases, have changed the natural history of the disease and extended the mean age of survival of from the middle teens to the late 30s^{54,55}.

In sharp contrast to DMD, life expectancy is only modestly affected in patients affected by the more benign version of the disease first observed by Becker (and therefore named Becker muscular dystrophy)^{56,57}. Becker muscular dystrophy (BMD) is nowadays known to be an allelic variant of DMD and is characterized by less severe symptoms that usually only appear at a later stage in life (often close to adolescence) and generally progress at a slower rate⁵⁸. However, the clinical detailed description of BMD lies outside the purpose of the present work.

1.3 THE *DMD* GENE

1.3.1. *DMD* gene characterization

Despite knowing that DMD followed an X-linked pattern of inheritance, the mutated gene and defective protein behind the disease remained unidentified until 1985, when Louis Kunkel identified the full transcript⁵⁹. By using a mapping approach similar to those designed by the concurrent works of Kay Davies and Ron Worton^{60,61}, Kunkel associated the *DMD* gene with a precise A/T-rich genomic locus on the short arm of the X chromosome (Xp21)(**Fig.1.1.a**).

The *DMD* locus is now known to span about 2.22 Mb and is therefore considered the largest in the human genome. The *DMD* gene includes 79 exons (mean size 0.2 kb) separated by considerably large introns (mean size 35 kb).

There are at least seven tissue-specific promoters located throughout the *DMD* gene (**Fig.1.1.a**). These are named according to the main tissue in which they are active (even though their expression might not be exclusive to only one tissue)^{62–65}. Dystrophin isoforms produced by each promoter are indicated by the acronym Dp followed by a

number indicative of their molecular size (**Fig.1.1.b**).

Three promoters, known as brain (B), muscle (M) and cerebellar Purkinje neurons (P), are located at the 5' end of the gene and produce the full-length dystrophin isoforms Dp427b (also known as Dp427c), Dp427m and Dp427p^{62,64,66}.

The remaining promoters are instead intragenic. These are known as the retinal (R), brain-specific (B3), Schwann cells (S) and general (G) promoters and are located in introns 30, 45, 56 and 63, respectively. R, B3, S and G promoters generate the shorter dystrophin isoforms called retinal (Dp260)⁶⁷, brain specific (Dp140)⁶⁸, Schwann cells (Dp116)⁶⁹ and general isoform G (Dp71)^{70,71}, respectively.

However, more dystrophin isoforms can be generated by the alternative splicing generated at the 3' end of the *DMD* transcript, as reported by Feener *et al.*⁷².

1.3.2. The mutational spectrum of DMD gene and DMD duplications

Three main kinds of mutations in the *DMD* gene cause DMD: intragenic deletions, duplications and point mutations⁷³ (**Table 1.1**). While these mutations can occur anywhere in the *DMD* gene, there are two main mutational hotspots: exons 2-20 at the 5' and exons 45-55 at the 3' of the gene.

According to a recent study published by Bladen *et al.*¹, mutations that were smaller than one exon (and therefore named small mutations) accounted for 20% of *DMD* mutations, while the majority of *DMD* mutations (80%) involved one or more exons (large mutations). Small mutations included small base pairs deletions (25%) or insertions (9%), splice sites mutations (14%) and point mutations (52%). These included nonsense mutations (50%) and missense mutations (2%). Atypical mutations occurring in introns or in the 5' UTR or 3' UTR accounted for a very limited number of *DMD* mutations (0.3%). Within large *DMD* mutations, 86% were deletions and 14% duplications (**Table 1.1**).

Hu *et al.* attributed the origin of dystrophin duplications to unequal exchange of either sister chromatids or homologous chromosomes in the germline of maternal grandparents⁷⁴. Studying the genomic breakpoints of DMD patients carrying tandem duplications, he hypothesized that the causative mechanisms could either be homologous or non-homologous recombination, depending on the degree of similarity of the recombining DNA sequences^{75,76}. More recent studies however, were also able to associate some duplications to a paternal transmission^{77,78}.

A study performed in late 1980s by Den Dunnen *et al.*⁷⁹ on a cohort of almost 200 patients showed that most of the observed dystrophin duplications localize close to the two mutational hotspots located close to the dystrophin 5' and 3' regions, with a

predominance of the 5' mutational hot spot. Interestingly, the frequency of duplications reported in Japanese patients seems to be slightly increased, with an higher tendency to cluster around the 3' mutational hot spot than those reported in European patients^{80,81}.

Usually, duplications are in the tandem head-to-tail orientation and span few exons. However, considerably larger duplications have been also reported in the literature^{82,83}: for example, a duplication spanning exons 4-42 was recently identified in a 9-month DMD patient⁸⁴. The analysis carried out by Bladen *et al.*¹ confirmed that the majority (11%) of the analysed duplications spanned multiple exons, as opposed to the duplications involving a single exon (3%). Generally, multi-exon duplications are more commonly observed within either the proximal (exon 2 and contiguous exons) and distal (exons 45-55) hotspots⁸⁵. However, they can also be found throughout the gene, randomly distributed. The most frequent single exon duplication is the duplication of exon 2, followed by the duplication of exon 17¹.

Finally, duplications can potentially co-occur with other mutations. Ishmukhametova *et al.*, for example, identified a tandem dystrophin duplication of almost 700 kb, displayed according to a duplication-inverted triplication-duplication pattern⁸⁶: the patient reported a duplication of exons 45 to 50, coupled with the inverted insertion of the triplicated exon 51 and finally followed by another duplication spanning exons from 52 to 60. A further novel mutation involving the noncontiguous duplication of exon 1, 42 and 43 was also lately described in a patient displaying a phenotype intermediate between DMD and BMD⁸⁷.

The functional consequences of each mutation do not depend on how extended a mutation is (with the exceptions of very large deletions removing most of the coding region), but on whether the dystrophin reading frame is maintained or not. Generally, DMD is associated with mutations that disrupt the reading frame (out-of-frame mutations) and give rise to premature stop codons, as they result in the production of a non-functional truncated protein. Conversely, mutations that maintain the reading frame (in-frame mutations) allow the synthesis of a mutated but still functional dystrophin and are therefore associated with the milder BMD phenotype⁸⁸ (section 1.2). This rule is valid for 90% of cases, even though exception to this rule (i.e. DMD patients with in-frame *DMD* mutations and BMD patients with out-of-frame *DMD* mutations) can be found¹⁴⁻¹⁶.

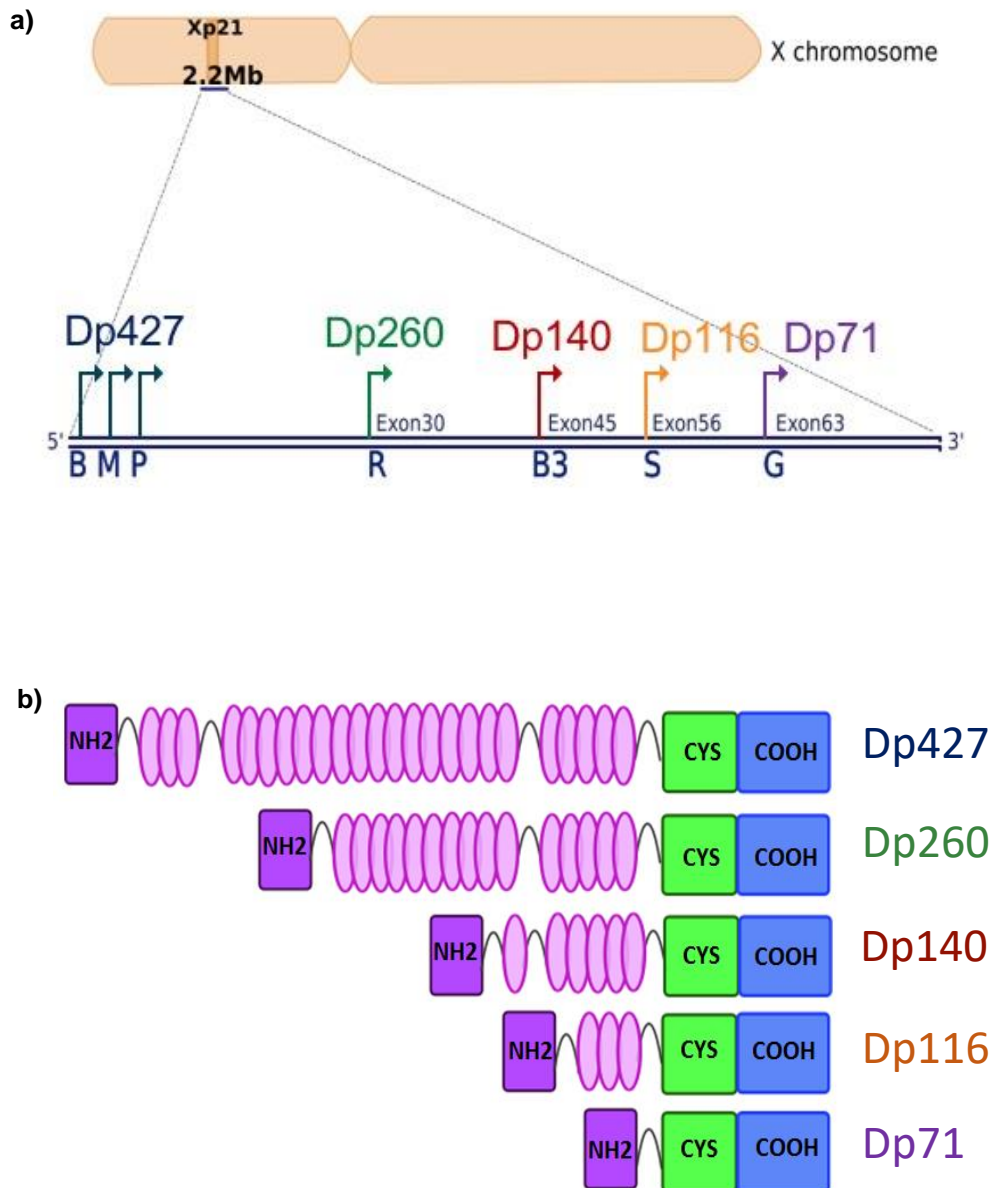


Figure 1.1. Schematic view of DMD gene and dystrophin isoforms.

a) The *DMD* gene spans 2.2Mb in Xp21. At least 8 promoters, indicated by arrows, are distributed along the gene and are responsible for the production of dystrophin isoforms (Dp). Promoters are indicated as brain (B), muscle (M), Purkinje (P), retinal (R), brain3 (B3), Schwann (S), and general (G). b) Dystrophin isoforms. Dp427 is composed of an amino-terminal domain followed by 24 spectrin-like repeats interspaced by four hinges, a cysteine rich region, and the carboxy-terminal domain.

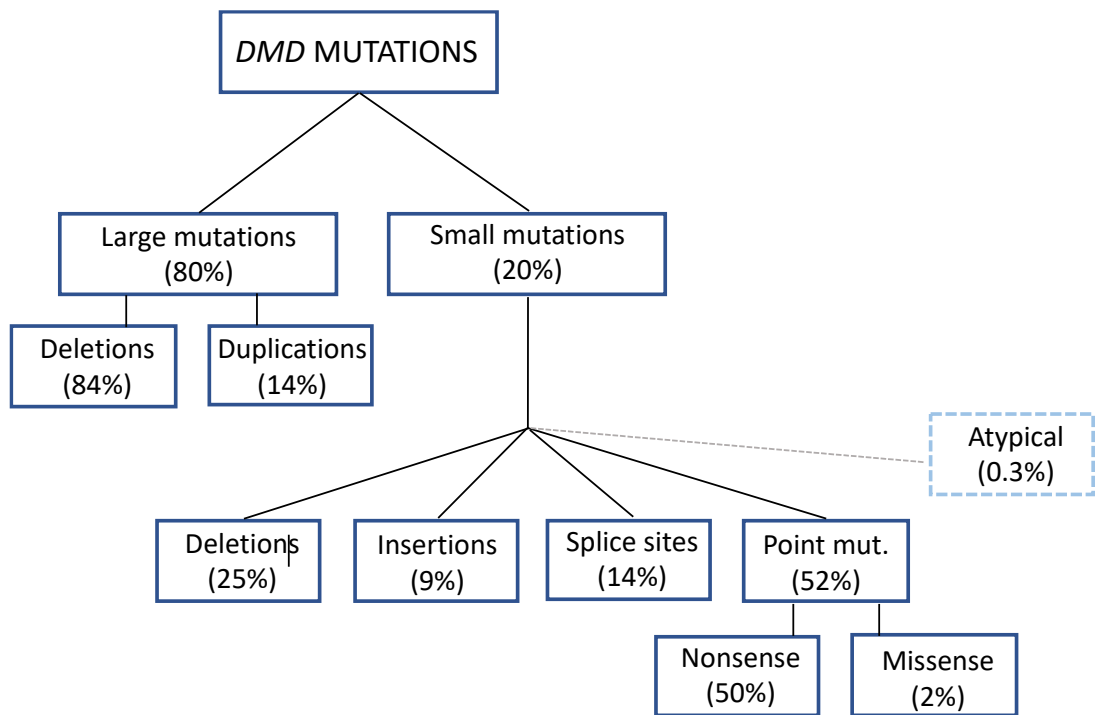


Table 1.1. Distribution of mutations within DMD gene.

1.4 DYSTROPHIN VERSATILITY: STRUCTURAL AND SIGNALLING ROLES

The transcript most frequently expressed by the *DMD* gene is the primary 14 kb-transcript composed of all 79 exons. This encodes the 427 kDa dystrophin protein of 3685 amino-acids⁸⁹ which was discovered by Kunkel and Hoffman in 1987⁹⁰.

Dystrophin represents approximately 0.002% of the total proteins present in striated muscles and localizes at the periphery of the muscle fibre, on the cytoplasmic side of the muscle membrane⁹⁰. However, dystrophin is not evenly distributed in the plasmalemma, but is enriched both at costameres⁹¹ and at the neuromuscular and myotendinous junctions^{92,93}.

Dystrophin is composed by four well defined regions, each with a specific function. These are known as actin-binding domain, rod domain, cysteine-rich domain and C-terminal domain (**Fig.1.2**).

The principal actin-binding domain (ABD1) is located at the N-terminal of the protein, where it covers a region of 240 amino-acids in between exons 2-9. This region is mainly involved in linking to the cytoskeletal actin through the two calponin homology (CH) domains CH1 and CH2, displaced in tandem⁹⁴. This domain also connects with the contractile apparatus by binding the intermediate filament protein cytokeratin 19 (K19)⁹⁵ (**Fig.1.2**).

The domain which follows the actin binding domain is composed of 24 tandem spectrin-like repeats consisting of triple α -helixes arranged in small rods of about 5 nm, known as the rod domain⁹⁶. This domain spans 2800 amino-acids, from exon 9 to exon 61 and is delimited by two proline-rich regions called hinges. Two other hinges are located within this region to allow the flexibility of the domain and confer to dystrophin the “shock absorber” property necessary for the transmission of lateral force during muscle contraction⁹⁶.

Approximately located in the middle of the rod (between repeats 11 and 14), there is a further actin binding domain (ABD2) that, through the interaction of its basic amino-acids with the acidic actin filaments, reinforces the action of ABD1⁹⁷ (**Fig.1.2**). The distal region between repeats 20 and 23 binds the intracellular microtubules^{98,99 100} (**Fig.1.2**). Repeats 17 and 22 have been recently shown to contain binding sites for α - and β -syntrophins, respectively¹⁰¹. This work by Adams *et al.* also showed that the α -syntrophin bound to repeat 17 is the only α -syntrophin that also binds to the neuronal nitric oxide synthase (nNOS) and facilitates nNOS interaction with dystrophin¹⁰¹ (**Fig.1.2**). Furthermore, the rod domain plays a role in directing dystrophin to the muscle membrane due to the ability of tryptophan residues that link with the sarcolemma phospholipids¹⁰². A recent study identified membrane-binding domains in repeats 1-3

and 10-12¹⁰³. Moreover, the polarity-regulating kinase PAR-1b binds to the repeats 8–9 and phosphorylates sites in this region, stabilizing the interaction of dystrophin with the dystrophin-associated protein complex (DAPC)¹⁰⁴.

Finally, hinge 4 is implicated in protein-protein interactions as its WW motif (found in several proteins with regulatory and signalling functions) anchors dystrophin to the sarcolemma by binding the C-terminus of β -dystroglycan^{105–107} (**Fig.1.2**).

This binding is aided by two adjacent EF hand-like motifs part of the third cysteine-rich domain which extends for 280 amino-acids and involves exons 62-69. These EF-like hands also provide a calcium binding site¹⁰⁸. The cysteine-rich domain also includes a site known as ZZ domain. ZZ domain is a zinc-finger domain with affinity to divalent cation such as Zn^{2+} ¹⁰⁹ which binds both to calmodulin (in a calcium-dependent manner) and to the adaptor protein ankyrin-B, whose function is to anchor dystrophin to the membrane¹¹⁰.

The last functional domain is the C-terminal region, which covers a region of 420 amino-acids from exons 70 to 79. It is composed of two polypeptides arranged in an α -helical coiled-coiled structure involved in protein-protein interaction. This domain provides binding sites for α -dystrobrevin and syntrophins ($\alpha 1$ and $\beta 1$)¹¹¹ (**Fig.1.2**) and contains several phosphorylation sites, important for a variety of kinases, including MAPK^{112,113}.

Generally, dystrophin is considered as a connector that structurally links the inner and outer muscle fiber environment, also known as extracellular matrix (ECM).

Intracellularly, dystrophin regulates the organization of the cytoskeletal network by interacting with all the three major filamentous structures. Dystrophin interaction with the filamentous actin (F-actin) is necessary for force generation¹¹⁴, while the association with intermediate filaments and microtubules helps to control the organization of their lattice (and consequently the localization of intracellular organelles)^{115,116}. The connection with ECM, instead, is made possible through the association of dystrophin with β -dystroglycan, a transmembrane protein that together with other transmembrane, extracellular and intracellular proteins forms the DAPC¹¹⁷. β -dystroglycan, in fact, is connected with its extracellular glycosylated isoform (α -dystroglycan), which in turn binds the ECM proteins laminin, agrin and perlecan, among others^{118–120} (**Fig.1.2**).

However, dystrophin also plays a role in the transmission of signals. This in part occurs because of its interaction with the DAPC, which includes proteins having a precise signalling role^{121–123}. Examples are the interaction of the transmembrane protein sarcospan with the extracellular $\alpha 7 \beta 1$ integrin (involved in several signalling pathways)¹²⁴, and the association of β -dystroglycan with the growth factor receptor-bound 2 (Grb2) which participates in signalling pathways involving Tyr-kinases and

integrin axes¹²⁵.

However, the involvement of dystrophin in signal transmission also derives by its association with the protein α -dystrobrevin and syntrophins. α -dystrobrevin contains numerous serine, threonine, and tyrosine phosphorylation sites targeted by kinases^{126,127}. α 1- and β 1-syntrophin act as modular adaptors that anchor signalling molecules, by virtue of their association with a large number of protein and lipid kinases, ion channels, transporters and G-protein receptors^{128–132}. Moreover, similarly to β -dystroglycan, α 1-syntrophin also binds Grb2¹²⁵.

Finally, a further contribution to the transmission of signals is given by Ca^{2+} /calmodulin kinases that act both on α 1-syntrophin and, together with other kinases, on dystrophin itself¹³³.

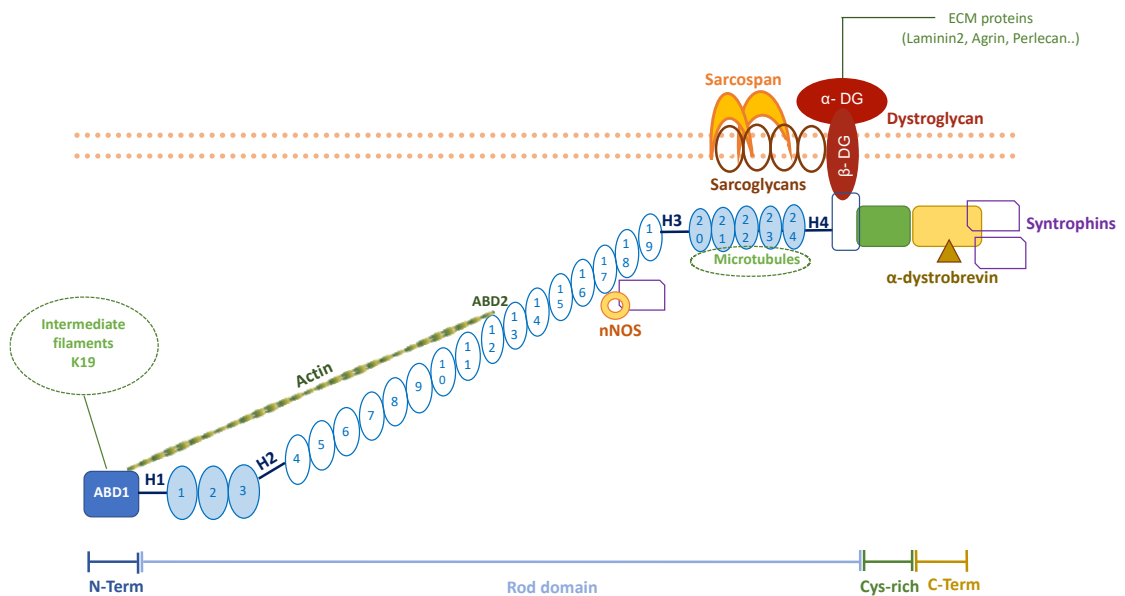


Figure 1.2. Dystrophin structure and interactions.

Dystrophin is composed by an N-terminal domain, a central rod domain divided by four hinges (H1-H4), a cysteine-rich domain and a C-terminal domain. Dystrophin binds both to cytoskeletal actin and to the transmembrane dystrophin-associated protein complex (DAPC), providing a link between the inner muscle fiber environment and the extracellular matrix. Intracellularly, dystrophin binds actin, intermediate filaments and microtubules through its N-terminal and rod domain. The rod domain also binds nNOS (through α 1-syntrophin) and β -dystroglycan through the WW domain within the hinge region (H4). The C-terminal domain binds to both α -dystrobrevin and the syntrophins α 1 and β 1. The link with the extracellular matrix (ECM) is guaranteed by the connection of β -dystroglycan with α -dystroglycan, as α -dystroglycan links the ECM protein laminin2, agrin and perlecan.

1.5 DMD PATHOGENICITY: INSIGHTS INTO THE MOLECULAR CHANGES FOLLOWING DYSTROPHIN ABSENCE

1.5.1. Faulty regeneration and development of fibrosis

In-frame mutations in the *DMD* gene typically allow the production of a mutated but partially functional dystrophin protein⁸⁸, depending on which of the above-mentioned protein-protein interactions are affected. However, in presence of mutations that either abolish dystrophin production or disrupt its stability, the connection with the proteins responsible for membrane and cytoskeleton stability is lost. Consequently, the muscle fibres are more prone to be damaged following the mechanical stress derived by eccentric muscle contraction.

Following muscle damage and degeneration, satellite cells (a specialized type of stem cells found underneath the basal lamina of each fibre) are activated to trigger muscle repair¹³⁴. When this happens, satellite cells undergo asymmetric divisions to both replenish the stem cells pool and also generate myoblasts¹³⁵. These either mature giving rise to new myofibers or repair damaged segments of existing myofibres¹³⁶. However, the sustained damage of dystrophic muscles causes the continuous activation of satellite cells, so that multiple regeneration cycles are required. The reduction of the satellite cells telomeres following each cycle contributes to cell senescence, so that the stem cells pool is rapidly exhausted¹³⁷. This leads to an impaired regeneration over the long term.

Moreover, the hostile dystrophic environment of the niche has been shown to negatively impact muscle regeneration, despite the satellite cells themselves being able to effectively regenerate after transplantation on a more permissive environment¹³⁸.

Finally, a recent study showed that the absence of dystrophin in satellite cells themselves could further affect muscle regeneration. In activated satellite cells, dystrophin is asymmetrically distributed to one side of the cell and associated with PAR-1b kinase which, acting in concomitance with the protein Pard3, regulates cell polarity¹³⁹. When dystrophin is missing, the downregulation of these polarity regulators results in polarity loss. As a consequence, the asymmetric cell divisions that give rise to the stem cell pool decrease.

The failure of the regenerative process leads to the replacement of muscle fibres with connective and adipose tissue. This creates the characteristic fibrotic pattern observed in dystrophic muscles^{140,141}.

Degenerating fibres are invaded both by chemokines, cytokines and inflammatory cells expressing high levels of inducible nitric oxide synthase (iNOS) (such as macrophages, T-lymphocytes and mast cells)^{142–145}. Extracellular fibroblasts are activated and secrete fibronectin, glycosaminoglycan and proteoglycans¹⁴⁵ which sustain the muscle fibrosis.

The replacement of muscle with fibrotic tissue further impedes the regenerative process by compromising growth factors that support cell proliferation and myogenic differentiation, as for example the transforming growth factor (TGF)- β and insulin growth factors (IGFs)^{146,147}. Increased lipogenesis is also observed consequently to the alteration of the glycogen pathway, that preferentially incorporates fructose at the expense of glucose¹⁴⁸.

1.5.2 Dysregulated intracellular homeostasis

Muscle damage has been associated with a disrupted homeostasis of ions and metabolites¹⁴⁹. Deleterious consequences have been in particular associated with an increased cytosolic concentration of sodium (Na^+) and calcium (Ca^{2+}). Na^+ increase overloads the Na^+/K^+ pumps leading to adenosine triphosphate (ATP) molecule exhaustion, while the dramatic Ca^{2+} increase results in both the activation of Ca^{2+} -dependent proteases and Ca^{2+} -channels^{150,151}.

The activation of the protease calpain eventually results in muscle necrosis¹⁵². The action of the Ca^{2+} -dependent phospholipase A_2 on the membrane phospholipids generates lysophospholipids and free fatty acids that damage the membrane¹⁵³. The homeostasis of mitochondria (where Ca^{2+} is sequestered consequently to its dramatic increase in the cytoplasm) is also affected. The mitochondrial impairment results in reduced oxidative phosphorylation, which leads to an increased production and accumulation of reactive oxygen species (ROS) which contribute to the oxidative stress reported in DMD^{154,155}. ROS are also involved in the peroxidation of membrane lipids, thus contributing to increase the membrane permeability¹⁵⁶.

The increased levels of Ca^{2+} also deregulate the Ca^{2+} ion-channels located on the sarcoplasmic reticulum and contributes to the uncoupling of the dihydropyridine and ryanodine receptors (DHPR and RYR) already observed in absence of dystrophin¹⁵⁷. Furthermore, RYR1 channels are also impacted by the cytosolic increase of the nitric oxide (NO) produced by the iNOS expressed by the inflammatory cells. NO, in fact, hypernitrosilate the RYR1 channel, which results in spontaneous release of calcium Ca^{2+} in the cytoplasm. This further impairs muscle contraction^{158,159}.

Finally, the reduction of NO derived by nNOS mislocalization causes vasoconstriction, which results in muscle ischemia¹⁶⁰.

1.6 MOLECULAR ADVANCES IN DMD TREATMENT

As a result of the abundance of the skeletal muscles and the DMD severity, a therapeutic design able to efficiently target all muscle tissue and to correct each mutation at the genomic level still does not exist. However, several approaches have been designed to slow down DMD progression. The first strategies to be conceived were acting on downstream pathways or effects secondary to the missing protein while, more recently, new approaches able to target dystrophin itself (primary defect) have started to appear. Guiraud *et al.* distinguished between “dystrophin-independent” and “dystrophin-based” strategies³, reviewed in the following paragraphs and in the below table 1.2.

Act on dystrophin?	Therapeutic strategy	Applicability	Correct DNA?
No	Steroids	All <i>DMD</i> mutations	No
	Muscle growth strategies	All <i>DMD</i> mutations	No
	Utrophin modulation	All <i>DMD</i> mutations	No
	Modulation of nNOS signalling	All <i>DMD</i> mutations	No
	Upregulation of ECM components	All <i>DMD</i> mutations	No
Yes	Aminoglycosides	Frameshift point mutations	No
	Exon-skipping	<i>DMD</i> deletions and small duplications	No
	Cell-based therapies	All <i>DMD</i> mutations	No
	Gene therapies	All <i>DMD</i> mutations	No (except engineered nucleases)

Table 1.2. Therapeutic designs for DMD treatment.

1.6.1. Dystrophin-independent strategies

Dystrophin-independent strategies are designed to combat the secondary effects derived by the absent or unstable dystrophin protein. As such, they could potentially benefit any DMD patient, independently of the underlying mutation.

a) Corticosteroids

Historically, the first drugs to be used to inhibit the inflammation originating from muscle breakdown were corticosteroids, a class of synthetic drugs able to interfere with the activation of the immune system¹⁶¹. The three most common steroid preparations used in DMD are prednisone, prednisolone (hydroxylated derivative of prednisone) and deflazacort (oxazoline derivative of prednisolone). Patients routinely treated with these

formulations^{162,163} showed improvement in muscle strength, cardiac and ventilator function as well as a delay in independent ambulation loss^{163,164}. However, even though they seem to have a cardioprotective effect¹⁶⁵, these drugs are often associated with adverse effects such as weight gain and inhibition of growth, behavioural issues, pubertal delay and bone demineralization that increases the risk of bone fractures. Corticosteroids have been also associated with cramps and myoglobinuria¹⁶⁶. Therefore, it is essential that the best regimen of drug administration, both regarding the dose and type of steroid, is identified for each patient¹⁶⁷. A phase III randomized, double-blind clinical trial suggested that deflazacort could be more effective and lead to less weight gain than prednisone/prednisolone¹⁶⁸. The recent re-analysis of the trial data confirmed that deflazacort may result in a better muscle protection and preservation of motor function, although larger prospective studies are underway¹⁶⁹. The beneficial role of these molecules was also confirmed by a recent study of young DMD patients treated for 8.5 years, which found that the daily use of glucocorticoids can improve the disease outcome¹⁷⁰.

Still, it is important to remember that the use of corticosteroids only delays muscle function decline, but does not stop the disease progression.

b) Muscle growth strategies

Another strategy to counteract the muscle wasting in DMD aims to exploit regulators of muscle mass; specifically, to modulate the action of the negative regulator of muscle growth, myostatin, and its antagonist follistatin¹⁷¹. Myostatin inhibitors have been shown to effectively ameliorate pathology in *mdx* mice^{172,173}. However, in contrast to the study based on follistatin overexpression in BMD patients¹⁷⁴, clinical trials performed in DMD patients did not show promising results and raised concerns related to adverse effects as epistaxis and enlargement of blood vessels^{175,176}. A recent study attributed these unsuccessful outcomes to the natural downregulation of the myostatin pathway in the presence of atrophying diseases (and hence in DMD)¹⁷⁷.

Parallel research lines focus instead on investigating the effect of modulation of other growth factors: increased muscle mass, resistance to fatigue and reduced muscle breakdown was observed in *mdx* mice overexpressing the insulin-like growth factor I (IGF-I), shown to promote skeletal muscle growth and repair^{178–182}. Similarly, amelioration of the disease was observed in *mdx* mice following the blockage of the bone morphogenetic protein (BMP) pathway, as this acts against the regeneration of dystrophic muscle¹⁸³.

c) Utrophin modulation

As utrophin is a ubiquitous paralog of dystrophin (with which it shares 80% homology)¹⁸⁴, it could potentially be exploited to counterbalance the effect of missing dystrophin.

Apart from in early development and in regenerating fibres, where it is associated throughout the sarcolemma, utrophin in adult muscle is localized at the neuromuscular and myotendinous junctions¹⁸⁵. The compensatory role of utrophin has been hypothesized following the observation that this protein is normally upregulated in *mdx* mice (which show a milder disease course and a better regenerative capacity compared to DMD patients¹⁸⁶), while the severity observed in double knock-out mice lacking both dystrophin and utrophin is similar to DMD patients¹⁸⁷. Despite utrophin displaying differing binding sites for actin and microtubules and failing to restore nNOS localization^{99,188,189}, its ubiquitous upregulation was shown to be safe¹⁹⁰ and was able to ameliorate both sarcolemma stability and muscle pathology in *mdx* mice^{191,192}. Both utrophin upregulation and nNO targeting could be obtained by exploiting biglycan, a protein naturally present in the extracellular matrix¹²³. Tivorsan Pharmaceuticals recently started a Phase I clinical trial to evaluate the effect of human recombinant biglycan treatment in DMD patients¹⁹³. Further studies conducted in the *mdx* mouse animal model have demonstrated that the time of utrophin upregulation is crucial, as utrophin modulation at birth was able to prevent muscle pathology itself¹⁹⁴. Kay Davies' group recently showed that utrophin upregulation in *mdx* mice also reduces mitochondrial damage and thus aids the oxidative stress diminishing the production of harmful free radicals¹⁹⁵.

Among the chemical compounds tested to mediate utrophin modulation, the anti-inflammatory and pain reliever Celecoxib improved lower limb strength and ameliorated cardiorespiratory function in *mdx* mice¹⁹⁶. Moreover, a Phase II open-label clinical trial (Summit Therapeutics) to evaluate the efficacy and safety of the utrophin up-regulator Ezutromid was undertaken in DMD patients¹⁹⁷ but, unfortunately, this study recently failed to meet its endpoints.

d) Modulation of the nNOS signalling pathway

The interaction between dystrophin and nNOS directs the localization of the NO synthesized by nNOS, whose fundamental role is to promote vasodilation and ensure oxygenation of muscles. Consequent to NO mislocalization, muscle ischemia was observed in DMD patients due to an excessive vasoconstriction¹⁹⁸. Among the different molecules tested in *mdx* mice to either boost NO-associated signalling or re-localize

nNOS at the sarcolemma^{199–201}, the commercially available vasodilator Sildenafil was able to reduce ischemic damage and ameliorate cardiac function in *mdx* mice²⁰² and was therefore tested in DMD patients. However, despite the encouraging outcome initially observed²⁰³, the use of Sildenafil was interrupted in view of adverse side effects that in some cases also included heart failure²⁰⁴.

e) Upregulation of components of the ECM

One possible way to overcome the disrupted dystrophin-mediated signalling is to exploit components of the extracellular matrix involved in the transmission of signals and stimuli between the cytoplasm and the extracellular environment. Integrins and associated laminins are implicated in numerous signalling pathways and have shown to be essential for the integrity of muscle fibres, as their loss causes muscular dystrophies^{205–207}. Hodges *et al.* observed an increase in integrin $\alpha 7$ expression in muscle fibres of both *mdx* mice and DMD patients, suggesting a potential compensatory role²⁰⁵. Interestingly, the upregulation of integrin $\alpha 7$ proved to be beneficial not only in *mdx* mice but also in double knock-out mice lacking both dystrophin and utrophin, in which the mild phenotype was accompanied by improved muscle functionality and regeneration^{208–210}. Laminin $\alpha 1$ constitutes a further appealing therapeutic candidate, as it is able to boost integrin $\alpha 7$ expression and therefore leads to the above-mentioned beneficial outcomes. However, despite laminin $\alpha 1$ showing effectiveness in ameliorating the phenotype of mice with congenital myopathies and congenital muscular dystrophies^{211,212}, it does not ameliorate the pathology in *mdx* mice²¹³.

1.6.2. Dystrophin-based strategies

Differently than dystrophin-independent strategies, not all the different dystrophin-dependent strategies can be applied to any DMD patient independently than the *DMD* mutation. These strategies are discussed below considering their suitability for duplication mutations, focus of this work.

a) Aminoglycosides

Following the discovery that aminoglycoside antibiotics are capable to revert nonsense mutations *in vitro* in yeast^{214,215}, they have been exploited as a potential therapy that may benefit patients carrying the above-mentioned mutations. These compounds can bind to ribosomes and affect their activity leading to misreading through the transcript; in this way, the ribosomes can bypass the mutant amino-acid, thus masking the resulting stop codon.

The first of these compounds used to obtain dystrophin read-through was gentamicin. In the late 90s', Barton-Davis *et al.* observed that controlled doses of this antibiotic could lead to the production of a full-length dystrophin both *in vitro* and *in vivo* in the dystrophic *mdx* mouse model²¹⁶, in which dystrophin production is aborted by a premature stop codon due to a nonsense mutation in exon 23²¹⁷.

Following optimization of the gentamicin dose to be used in patients, Jerry Mendell and his group observed some levels of dystrophin restoration and decreased levels of serum creatine kinase, marker of muscle damage, in treated DMD patients²¹⁸. However, the beneficial action of these compounds could be accompanied with an increased risk of renal failure and ototoxicity, suggesting that they might not represent the optimal therapy for this subset of patients²¹⁹.

New molecules that retained the capacity of gentamycin to induce misreading but without its severe side effects were developed in the following years. Among them, Ataluren (also known as Translarna, developed from PTC therapeutics) recently completed both a phase IIb and phase III randomized controlled clinical trials, and it has been approved as an orally administered drug for patients carrying a premature stop codon in some European countries including UK²²⁰.

However, as this study design exploits the ability of ribosomes to read through mutated bases, it can only be applied to nonsense mutations, but not to deletions or duplications.

b) Exon-skipping

The use of antisense oligomers pioneered by Ryszard Kole, laid the foundation for what nowadays is known as exon-skipping.²²¹ This molecular approach relies on the use of small oligomers 20-30 nucleotides long (also called antisense oligonucleotides or AONs) designed to be complementary to specific regions of the target pre-mRNA, generally where the mutation is located. The alignment of the pre-mRNA with the above-mentioned oligomer masks the mutated exon, that is therefore “skipped” during the pre-mRNA splicing. This allows the production on a smaller but functional version of dystrophin protein and therefore turns a severe DMD phenotype into a milder Becker-like phenotype. The excitement behind this approach, potentially applicable to 60-80% of patients with disrupted dystrophin reading frame²²², led scientists to develop multiple splice-switching oligomers targeting the most common mutational regions. Oligomers targeting different dystrophin exons have been extensively designed and used to restore dystrophin expression *in vitro*²²³ and, most importantly, *in vivo* in various animal models of DMD^{224–228}.

A lot of effort has been put in designing AONs able to skip exon 51, as this is applicable to 13-14% of patients²²⁹. Two clinical trials were performed in our unit with the aim of skipping exon 51 through the administration of two different chemicals developed respectively by Prosensa Therapeutics and Sarepta Therapeutics: the 2'O-methylphosphorothioate oligoribonucleotide (Drisapersen) and the phosphorodiamidate morpholino oligomers (Eteplirsen). The studies revealed that, despite its (limited) efficacy, the repeated subcutaneous injection of Drisapersen gave rise to significant adverse events and therefore its development was stopped by the company following a negative opinion by the Food and Drug Administration (FDA) agency. In contrast, the excellent safety profile of intravenously administered Eteplirsen and its ability to restore dystrophin protein, albeit at very low levels, led to its recent accelerated FDA approval in US^{230–232}.

Following these promising results, morpholino antisense oligomers to induce skipping of exon 53 have been developed and have entered a phase I-dose escalation clinical trial²³³. As part of the SKIP-NMD project, our group is currently involved in a Phase I/II clinical trial aimed to assess the safety and efficacy of Golodirsen²³⁴ (Sarepta Therapeutics), expecting to skip *DMD* exon 53, as 8% of *DMD* patients carry deletions amenable to exon 53 skipping. This drug has also been evaluated in a Phase III clinical trial (ESSENCE study) together with the compound Casimersen, developed for the skipping of *DMD* exon 45²³⁵.

However, this approach is only suitable for the treatment of mutations that can be corrected by skipping up to two exons²²⁹, but not for the restoration of large duplications or mutations in regulatory or N-/C-terminal dystrophin regions²³⁶.

c) Cell-based therapies

Cell therapy and gene therapy approaches, not necessarily mutually exclusive, have been considered as a potential treatment for *DMD*.

Generally, cell-therapy is meant as the administration of living healthy cells to the patient to reintroduce the missing gene product²³⁷.

The ideal candidate to use for cell-based therapies should be easily expandable *in vitro* without losing the myogenic properties and be able to be delivered systemically, as to reach the muscles body wide. Also, once transplanted, the candidate cell should be able to survive, proliferate, migrate through the muscle and also replenish the satellite cells niche.

The best option would envisage the use of cells isolated from the individual himself (autologous source). These would be expanded and genetically modified *in vitro* and

finally re-injected back into the patient. This approach would overcome the need to find an immune-compatible donor expressing the healthy gene copy.

Primary myoblasts have been considered as they can be derived from the muscle itself or obtained from induced pluripotent stem cells (iPSCs) derived from reprogrammed skin fibroblasts²³⁸. Despite the fact that myoblasts can be expanded *in vitro*, they have a limited post-transplant migration: in fact, studies showed most of these cells tend to remain at the site of injection²³⁹ or die shortly after injection into the affected muscle²⁴⁰. Also, since myoblasts have been shown to be unable to cross the endothelial wall, they cannot be systemically injected²⁴¹.

Alternatively, stem cells can be obtained from muscle (satellite cells, section 1.5.1) or from other sources such as the bone marrow and both adipose and endothelial tissues. These stem cells, in fact, showed tropism for muscle when injected into the bloodstream²⁴². Recent studies performed in mice were based on the isolation of satellite cells from single fibres isolation or fluorescence activated cell sorting (FACS) followed by their *in vivo* delivery, even though such approaches were limited to intramuscular injection^{243–245}. However, it has been hypothesised that by virtue of satellite cell plasticity, they might contribute to the generation of fibrotic scar tissue depending on the stimuli they receive following transplantation²⁴⁶.

Within the cell types that have shown to have myogenic potential, a more promising alternative to satellite cells might be represented by pericytes (mesangioblast-like cells)²⁴¹ and muscle-derived cells expressing the CD133+ glycosylated epitope. Pericytes can maintain their proliferative ability *in vitro* and, upon *in vivo* transplantation, contribute to the formation of mesoderm tissue²⁴⁷. These cells have been successfully used as allogenic cell-based therapies both in α -sarcoglycan-deficient dystrophic mice and dystrophin-deficient GRMD dogs^{248,249}. A few years ago they have also proved to be safe when delivered intra-arterially as part of a non-randomized, open label phase I/II a clinical trial in humans, but without functional improvements²⁵⁰.

CD133+ cells have been isolated from the peripheral blood and from skeletal muscle^{251,252}. Their myogenic capacity allowed them to form myotubes *in vitro* when co-cultured with myoblasts. Transplantation of CD133+ cells *in vivo* in immunodeficient *mdx* mice showed their ability to replenish the satellite cells pool and a greater regenerative capacity compared to myoblast transplantation^{253,254}. Intramuscular and systemic injection of allogenic CD133+ cells proved to be safe *in vivo* in a phase I clinical trial performed in DMD patients²⁵⁵. However, a recent study showed the reduced myogenic potential of DMD patient-derived CD133+ cells compared to healthy ones both *in vitro* and *in vivo* in immunodeficient mice injected with the human-derived

cells²⁵⁶.

d) Gene-therapies

Conventional gene-therapies aim to compensate for the genomic mutation underlying DMD by delivery of the corrected gene version to the affected cells.

Generally, these approaches rely on viral vectors derived from inactivated viruses, i.e. viruses in which the genes responsible for their virulence have been inactivated, despite maintaining their infectivity. The most commonly used viral vectors are based either on lentiviruses, adenoviruses or adeno-associated viruses (AAVs), each characterized by distinct tropism ranges and packaging capacity²⁵⁷ (**table 1.3**).

	Adenovirus	Lentivirus	AAV
Virus Size (nm)	10	80-130	18-26
Genome	dsDNA	RNA	ssRNA
Packaging Capacity	around 9 Kb	around 11 Kb	around 4.5Kb
Expression	Transient	Stable	Mostly transient
Target Cell Genome Integration	No	Yes	Rare
Muscle tropism	Low	Moderate	High
Immune response	High	Low	Very low

Table 1.3. Schematic of features distinguishing the main classes of viral vectors

The first vectors considered for gene therapies targeted to muscle tissue were those based on adenoviruses, a non-enveloped virus having a double-stranded DNA genome²⁵⁸. Recombinant adenoviral vectors have a clear tropism for the lungs, eyes and gastro-intestinal epithelium, as well as for the liver^{259,260}. However, their tropism for mature skeletal muscle is low²⁶¹. The main reason for this has been attributed to the low muscle expression of the surface Coxsackievirus-Adenovirus receptor (CAR), used for adenoviral transduction²⁶².

Moreover, these vectors have been associated with an high immunogenicity that might affect the duration of transgene expression in transduced tissues²⁶³ (**table 1.3**).

Another type of vectors considered for targeting muscle tissue is based on engineered integrating lentiviruses. Lentiviruses belong to the *Retroviridae* family and are enveloped particles whose genome is represented by a single-stranded RNA molecule²⁶⁴. Lentiviral vectors have a larger packaging capacity compared to adenoviral vectors (11 kb versus 9 kb approximately)²⁶⁴. The main concern about viral delivery of full-length dystrophin is the extremely large size of its cDNA (around 14 kb), beyond the capacity of all viral vectors²⁶⁵. However, recently our group developed a strategy that

successfully allowed dystrophin cloning into a lentiviral vector and dystrophin restoration *in vitro* in DMD myoblasts²⁶⁶.

Studies conducted *in vivo* in *mdx* mice treated by intramuscular and systemic injection of lentiviral vectors showed that transduction of muscle was only moderate^{267,268}. The administration of a lentiviral vector in neonatal mice led to highly efficient transduction of muscle, with concomitant targeting of satellite cells in both wild-type and *mdx* mice^{269–271}. This feature is particularly interesting considering the integrating nature of these vectors, as it would guarantee a permanent expression of wild-type dystrophin in the stem cells pool. However, the practical limitations of their use in the neonatal period and concerns regarding their integration into the genome (which may result in carcinogenesis)²⁷² have hampered the progress with this strategy.

Further viral vectors that combine a low immunogenicity with a wide tropism and a generally episomal transgene expression are represented by the recombinant adeno-associated viruses (rAAVs)²⁵⁷. These vectors have a smaller size compared to the above-mentioned viral vectors, and they have a single-stranded DNA genome. The main advantages of these vectors are their ability to sustain transgene expression for long time and the fact that they can be delivered systemically²⁷³. In skeletal muscle, AAV transduction was shown to last years in striated muscles²⁷⁴. Furthermore, most AAV serotypes have a tropism towards skeletal and cardiac muscle tissue^{275,276}. However, the main limitation of these vectors is their limited packaging capacity (around 4.5 kb), which prevents the possibility to express transgenes as large as full-length dystrophin (**table 1.3**).

Meanwhile, the observation of a BMD patients in whom a considerable portion of the *DMD* gene (46%) was deleted²⁷⁷ resulted in the development of a partially functioning miniaturized dystrophin. This feature was exploited to overcome the size limitation imposed by the AAVs capacity. These molecules, known as mini- and microdystrophin^{278–280}, lack a considerable part of the rod domain but retain the essential functional regions that guarantee dystrophin functionality (i.e. the domains binding actin and β -dystroglycan). Generally, mini-dystrophins can only be packaged into lentiviral vectors as they retain a bigger portion of the full-length dystrophin cDNA. Pfizer recently started a Phase 1b clinical trial (NCT03362502) to investigate the efficiency and safety of intravenous administration of a mini-dystrophin formulation in ambulant DMD patients²⁸¹. The expression of mini-dystrophin transgenes via rAAVs would require the splitting of the transgene in two overlapping rAAVs, so that, *in vivo*, their recombination would restore the entire micro-dystrophin sequence²⁸². Micro-dystrophins, instead, only contain the N-terminal actin-binding domains, from two to three hinge regions, some spectrin repeats, and the cysteine-rich region that together

with the WW-domain binds β -dystroglycan (section 1.4). Due to their extremely reduced size, micro-dystrophin can be fitted into AAVs for *in vivo* delivery. Recently, treatment with micro-dystrophin has been shown to be safe and efficiently increases the levels of dystrophin protein both in dystrophic dogs (local and intravenous administration) and mice (intravenous administration), restoring muscle function^{283,284}.

Open label Clinical Trials from Sarepta (phase II, NCT03769116) and Solid (phase I/II, NCT03368742) are currently ongoing to test the efficiency of micro-dystrophin and its safety following its systemic intravenous delivery^{285,286}.

However, the extensive engineering of these molecules could potentially create new epitopes that could be recognized as non-self and thus elicit an immunogenic response.

While some of these treatments can mitigate the clinical symptoms of Duchenne and slower the disease progression, none of them provide a permanent correction of the causative mutation.

However, in the last two decades, the discovery of specific enzymes (called nucleases) engineered to introduce specific genetic alterations in a variety of cells, have opened the path to the promising “genome editing” therapeutic approach. Genome editing holds a great potential, as is based on the manipulation of genomic sequence with the ultimate aim to permanently correct DNA mutations.

1.7 THE EVOLUTION OF ENGINEERED NUCLEASES

Engineered nucleases are hybrid enzymes composed by two domains: one domain is designed to bind DNA in a sequence-specific fashion, while the second domain cleaves DNA in a nonspecific way. Therefore, engineered nucleases can potentially target any genomic locus of interest, where they can mediate a double-strand break. The introduction of such a cut stimulates the cell intrinsic repair pathways that repair double-strand breaks that occur during each cell cycle²⁸⁷ and thus maintain the integrity of the genetic information²⁸⁸.

These are the error prone non-homologous end joining (NHEJ) pathway and the high-fidelity homology directed repair (HDR) pathway. Precise cycle-specific molecular partners dictate which repair pathway will be prioritized.

Generally, HDR initiates following the resection of the 5' strand at the site of break by means of nucleases and helicases^{289,290}. This results in the generation of 3' single-stranded DNA (ssDNA) overhangs²⁹¹. Thanks to the later association with the RAD51 recombinase that generates a nucleoprotein filament, the ssDNA can invade the sister double-stranded DNA duplexes in which a region of homology is present and use them as a repair template in a scar-less fashion²⁹² (**Fig.1.3**).

Conversely, in NHEJ repair mechanisms, the two cleaved ends are held together by the heterodimeric protein Ku70-Ku80. The association of DNA with the Ku70-Ku80 complex in turn recruits the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs), a signalling serine/threonine kinase that acts consequently to cellular stress²⁹³.

If the double-strand break results in ssDNA ends, these are trimmed by the exonuclease activity of the nuclease Artemis²⁹⁴. The ligation of the blunt DNA ends finally occurs following the recruitment of the ligation complex, formed by the DNA ligase 4, its co-factor X-ray cross-complementation group 4 (XRCC4), and the XRCC4 like factor (XLF)/Cernunnos²⁹⁵. Through this final step, the break site is repaired^{296–298} (**Fig.1.3**).

HDR and NHEJ pathways are responsible for different genomic alterations when applied to gene editing²⁹⁹. HDR requires the presence of an identical, or almost identical, sequence to be used as repair template. It occurs only at defined cell cycle stages (during late S phase or G2) and is normally exploited either to cause or correct a specific mutation in the genome, or substitute a specific DNA stretch (**Fig.1.4**). Conversely, NHEJ predominantly occurs during the G1 phase (though it can occur throughout the cell cycle) and involves the simple ligation of the two DNA ends. NHEJ often results in the formation of unfaithful repair products as it introduces at the break site either substitutions, translocations or small (typically 1-50 bp) deletions and/or insertions, “indels”³⁰⁰. As such, it is generally used to disrupt or restore the reading frame of a mutated gene (**Fig.1.4**).

To date, four different classes of engineered nucleases have been designed, each defined by a distinct way of identifying and cleaving their genomic targets. Genome editing started with the identification of meganucleases (MNs), followed by the zinc-finger nuclease (ZFNs) and the transcription activator-like effector nucleases (TALENs) up to the most recent: clustered regularly interspaced short palindromic repeats (CRISPR) (**Fig.1.5**).

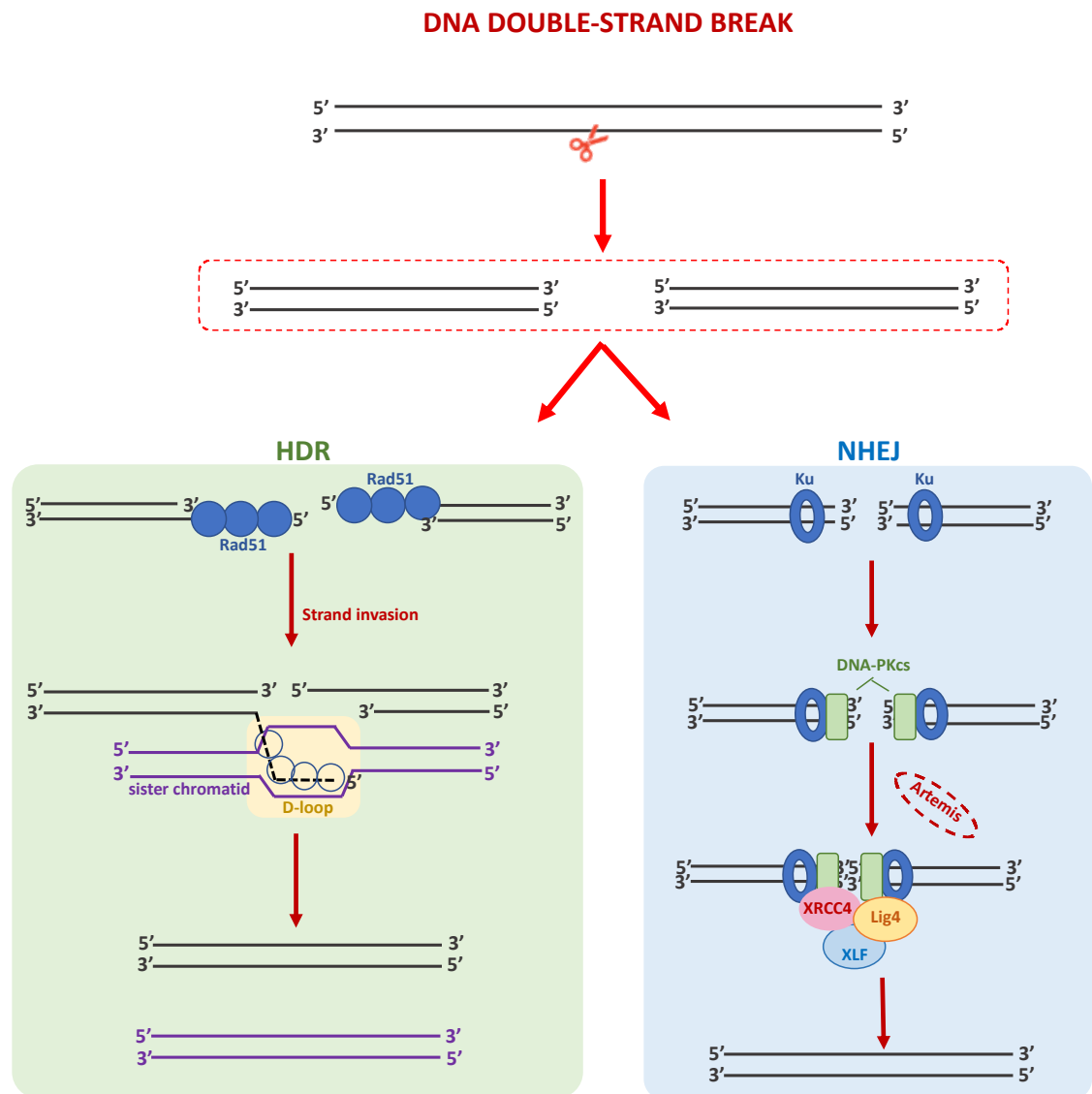


Figure 1.3. Schematic of DNA-repair pathways.

Double-strand DNA breaks are generally repaired following either HDR or NHEJ pathways. HDR pathway (left panel) requires processing of DNA ends to generate ssDNA. Once single stranded DNA associate with Rad51 protein, the nucleoprotein filament looks for a region of homology on the sister chromatid and creates a displacement loop (D-loop). The sister chromatid allows DNA repair working as a template. NHEJ repair (right panel) does not involve DNA ends processing. However, single stranded DNAs that may result from the DNA break are processed by the nuclease Artemis. This pathway requires the association of cleaved DNA ends with the heterodimeric protein Ku, which recruits DNA-PKcs. The DNA ends ligation is finally triggered by the DNA ligase 4, in association with XRCC4 and XLF.

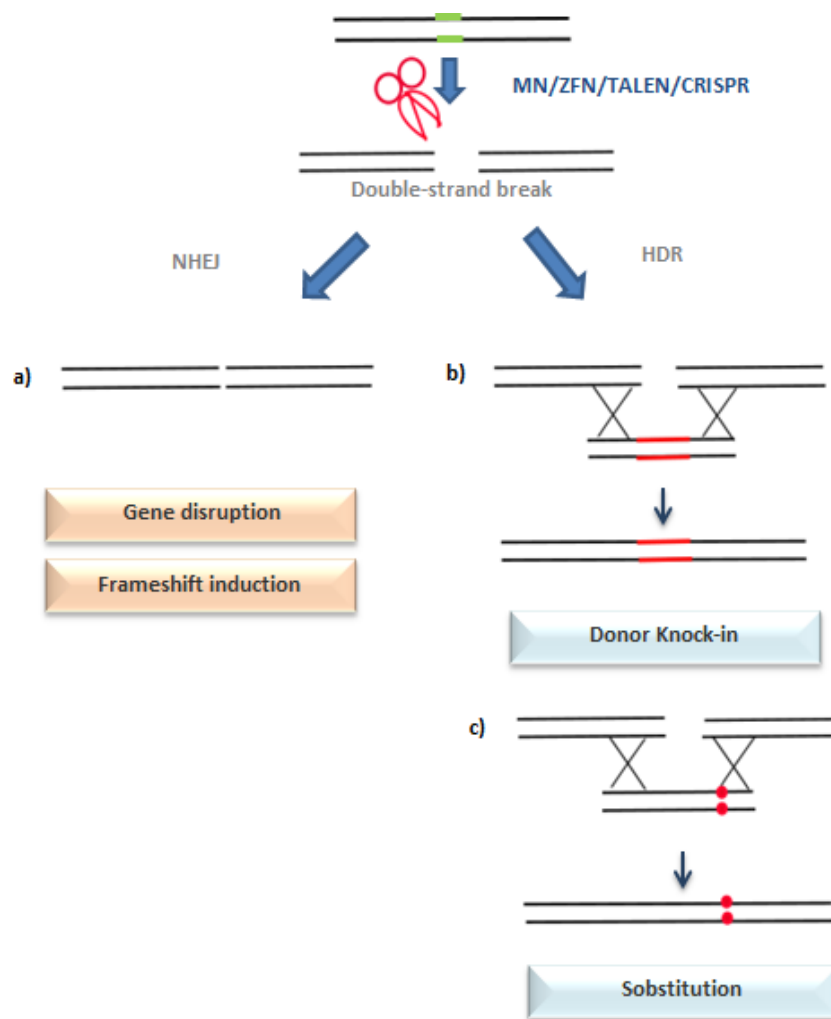


Figure 1.4. Genome editing using engineered nucleases.

Upon introduction of a double-strand breaks into the targeted genome region, intrinsic repair mechanisms become activated. a) In the absence of donor DNA template, the break is repaired by the NHEJ repair mechanism, which re-ligates cleaved ends. This usually results in deletions of one or more bases (indels) and can therefore be exploited to introduce a frameshift mutation and lead to gene disruption. b) The HDR pathway requires the exchange of DNA sequences from the donor DNA to the genomic DNA targeted for repair. Introduction of precise alterations can only occur in the presence of a donor DNA containing the desired sequence (targeted correction or knock-in) or c) containing single base pair mutations.

1.7.1 Meganucleases

Meganucleases, also known as homing endonucleases, define a large group of endogenous restriction enzymes found in several organisms, both prokaryotes and eukaryotes^{301,302}.

Meganucleases can be classified into five major families, each characterized by specific sequences and structural motifs³⁰³. The largest and well characterized meganucleases family is the LAGLIDADG family. This name derives by the presence of one or two copies of the LAGLIDADG motif that makes them function, respectively, as a single chain protein or a homodimer^{304,305}. Independently of their molecular structure, all members of this family arrange in a similar $\alpha\beta\alpha\beta\alpha$ fold. The LAGLIDAD motif placed centrally to α -helixes is surrounded at both sites by β -sheets, which are involved in the interaction with target DNA.

Meganucleases recognize genomic regions from 12 bp up to 40 bp long that are supposed to occur rarely in any given genome^{306,307} (**Fig.1.5.a**). This feature makes them the most precise nuclease class among the four mentioned above.

Despite their high precision, a major drawback of using meganucleases as editing tools is that the target gene must contain the target sequence for the cleavage to be successful. Moreover, the repertoire of naturally existing meganucleases is limited, thus further reducing the capability to target them to a desired gene in a given location³⁰⁸.

With the purpose of enriching the choice of targets and improving the cleavage efficiency, chimeric meganucleases have been successively produced both by either introducing small modifications in the sequence-recognition motif or by fusing or associating them with enzymatic domains derived from different meganucleases^{309–313}.

Meganucleases have been designed to efficiently target intronic dystrophin regions where they introduced indels, confirming their suitability for *DMD* gene editing³¹⁴. A different work showed that meganucleases successfully restored (through NHEJ) the reading-frame of a construct expressing the mutated cDNA of dog dystrophin. This approach allowed to obtain dystrophin restoration both *in vitro* in patient-derived cells and *in vivo* in *mdx* mice³¹⁵.

1.7.2. Zinc-finger nucleases

The development of zinc-finger nucleases derived from studies performed in early 90s' by Chandrasegaran *et al.* on the natural type IIS restriction enzyme *FokI*. He attributed the DNA binding and cleavage activity of this enzyme to two distinct and separable domains and observed that the domain responsible for cutting the DNA lacks specificity³¹⁶. A few years later, the same research team observed that the substitution of the original DNA recognition domains with a different one could affect FokI target specificity, so that this enzyme could be redirected towards custom genomic locations^{317–319}.

Among the DNA recognition domains tested, a set of three Cys2His2 zinc-fingers appeared to be particularly promising: each finger was able to inserting an α -helix into the major groove of the DNA double-helix and binding a DNA triplet³²⁰. These tailor-made chimeric proteins were therefore named zinc finger nucleases (ZFNs).

Most ZFNs are composed of three or four zinc finger motifs fused to a monomeric FokI domain, allowing the specific recognition of a variety of sequences from nine to twelve nucleotides long³²¹. Moreover, since FokI requires dimerization to be catalytically active³²², two juxtaposed monomeric ZFNs are used, ideally with a 5-6 base-pair long spacer region among them³²³ (**Fig.1.5.b**). This increases the length of the target site, which is therefore expected to be very rare or even unique in the genome, and consequently improves the cleavage specificity. As the tandem array of Cys2His2 zinc finger motifs in the DNA binding domain can be altered, ZFNs can be tailored to identify a wider set of DNA targets compared to MNs.

ZFNs have been successfully applied to different organisms to pursue diverse genome editing goals including gene disruption, insertion and correction of pathogenic mutations^{324–331}. However, the high cost required to assemble efficient ZFN has been a limiting factor for this approach. Furthermore, the generation of a ZFN array is indeed based on the combination of smaller zinc-finger modules whose affinity to target DNA is strictly dependent on surrounding modules and on the chromatin accessibility of the target region. The resulting "context dependence" often interferes with a successful assembly³³².

Ousterout *et al.* designed ZFN to skip *DMD* exon 51 in DMD-derived myoblasts, thus restoring the *DMD* reading frame and resulting in a functional protein. His approach was validated *in vitro* and *in vivo* in immunodeficient mice transplanted with the edited patient-derived cells³³³.

1.7.3. Transcription activator-like effector nucleases (TALENs)

The development of TALENs as editing tools followed the recent discovery of virulence factors naturally synthesized in the phytopathogenic bacteria of the *Xanthomonas* genus, named transcription activator-like effectors (TALE)³³⁴. Since they resemble eukaryotic transcription factors, TALEs can affect the plant's gene expression and lead to the development of disease symptoms³³⁴. TALEs are characterized by a central domain of tandem repeats able to bind single DNA nucleotides³³⁵. As such, TALEs became immediately appealing to scientists who exploited this hallmark to develop the customizable genome targeting tool known as TALENs³³⁶.

TALENs share with ZFNs the chimeric structure derived by the fusion of a DNA-binding domain and a monomeric FokI domain. However, in TALENs the zinc-fingers modules are replaced by a tandem array of up to 20 modules of 33-35 amino acids each. Within each module, two central hypervariable amino acids in positions 12 and 13 (known as repeat-variable di-residues or RVDs) are involved in the recognition and binding of a single DNA nucleotide³³⁷, conferring to this nuclease the highest targeting resolution compared to the other nucleases developed so far. Similar to ZFNs, TALENs work as juxtaposed dimers and their optimal conformation requires a spacer length of 12-21 bp. Each TALEN recognizes a region of 15-20 bp, further reducing the off-target effects risk compared to ZFNs³³⁸(**Fig.1.5.c**). Even though the cloning methods tested to produce functional TALENs revealed to be laborious and time consuming, the advantages offered by TALENs made them thrive in a short period of time^{336,339}.

As the other nucleases classes, also TALENs have been studied in relation to DMD; as an example, TALENs succeeded in restoring the disrupted *DMD* gene reading-frame in dermal fibroblasts derived from DMD patients carrying multi-exon deletions³⁴⁰.

However, this technology presents a disadvantage when considering its use for therapeutic applications: the size of a TALENs cDNA is around 3kb (versus 1kb of ZFN cDNA). Therefore, it is harder to be packed in viral vectors generally employed for TALENs delivery and expression into cells as AAVs (whose capacity is around 5 kb)³⁴¹.

1.7.4. Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9

a) [The discovery of CRISPR system](#)

Clustered regularly interspaced short palindromic repeats (CRISPR) was identified for the first time in the late 1980s by the research group led by Atsuo Nakata from Osaka University. While trying to clone the *iap* gene from the bacteria *Escherichia coli*, this

team accidentally observed in the bacterial genome a series of genomic repeats interrupted by shorter spacer DNA of unknown function^{342,343}.

A few years later, a study performed by Van Soolingen *et al.* in the Netherlands revealed the presence of genomic clusters of interrupted repeats in several *Mycobacterium tuberculosis* strains, with different interrupting sequences for each strain^{344,345}. Independently, Mojica *et al.* identified sequences of about 30-bp regularly repeated within the archaeal genome of *Haloferax mediterranei* and *Haloferax volcanii*^{346,347}.

In the late 1990s, the advent of the sequencing technology allowed scientists to obtain genome sequences of living organisms and further highlighted how widespread these regions were across a considerable number of microbial organisms^{348,349}. To mitigate the ambiguity due to the many names given to these sequences, Mojica and Jansen coined the acronym CRISPR, that became widely used³⁵⁰. To date, CRISPR DNA loci, have been found in approximately 40% of sequenced eubacterial and 90% of archaeal genomes³⁴⁸.

The function of CRISPR loci remained puzzling until 2005, when three independent research groups observed that the spacer CRISPR DNA was gathered from invading phages and plasmid DNA^{351–353}. Moreover, the work of Barrangou *et al.* showed that, following the insertion of viral sequences derived from the invader organism in the CRISPR locus, the bacteria *Streptococcus thermophiles* became resistant to further bacteriophage infection³⁵⁴. One year later, CRISPR-mediated prevention of horizontal plasmid transfer was also reported by Marraffini *et al.*³⁵⁵. These observations suggested that CRISPR represents a prokaryotic adaptive immune system to defend against invading phages and plasmids.

The molecular mechanism of the CRISPR system was finally elucidated by the studies published by Brouns *et al.*³⁵⁶ and Garneau *et al.*³⁵⁷. These scientists showed that the CRISPR system performs a targeted cleavage of invading nucleic acids through a specific CRISPR-associated nuclease, guided by precise RNA-molecules transcribed from the CRISPR locus itself. Target DNA is specifically chosen by virtue of its proximity to a 2-5 nucleotide motif called protospacer adjacent motif (PAM)^{358,359}.

Once cleaved, the foreign DNA is then incorporated in the CRISPR locus, increasing its size. Since genetic elements from each invader are successively incorporated in the CRISPR locus of the host genome, their disposition within the CRISPR locus represents a chronological memory of occurred infections.

b) Heterogeneity of CRISPR loci

CRISPR loci are characterized by a set of homologous genes located in close proximity to the CRISPR sequences, named CRISPR-associated (cas) genes³⁵⁰, followed by an A/T-rich region (known as leader) and a CRISPR-array (**Fig.1.6**).

Cas genes code for RNA-binding proteins, helicases, nucleases and polymerases needed for the interaction and cleavage of the target DNA³⁵⁰. The leader region which flanks the cas genes acts as a promoter for the downstream CRISPR-array³⁶⁰. This is formed by a variable number of repeated and partially palindromic sequences spaced by short stretches of individual DNA sequences (protospacers) derived from the invader genome. Partially palindromic sequences can be from around 23 and 47 base pairs long, while the length of protospacers varies from 20 to 70 base pairs.

Depending on the number and on the biochemical role of the required Cas proteins, the CRISPR system is catalogued in two main classes (I and II), each with a total of three main CRISPR types³⁶¹. Subtypes can be also identified within each CRISPR type^{361,362}. CRISPR systems belonging to class I (i.e. CRISPR type I, III and IV) rely on multiple Cas proteins to trigger the molecular events that result in DNA cleavage³⁶³.

Conversely, CRISPR-Cas systems belonging to class II (i.e. CRISPR-Cas type II, V and VI) use a single Cas protein to recognize and destroy the target DNA.

Each CRISPR class is listed in the below **table 1.4**, together with identified CRISPR types and subtypes.

CRISPR Class	CRISPR Type	CRISPR Subtype
I	I	I-A, I-B, I-C, I-D, I-E, I-F**, I-U
	III	III-A, III-B**, III-C, III-D
	IV	IV**
II	II	II-A, II-B, II-C**
	V	V-A, V-B, V-C, V-D, V-E, V-U1, V-U2, V-U3, V-U4, V-U5
	VI	VI-A, VI-B1, VI-B2, VI-C

Table 1.4. Classification of CRISPR system.

One of the most extensively studied type II CRISPR systems is the one derived from the bacteria *Streptococcus pyogenes*, belonging to the subtype IIA. Cas9 nuclease is the signature protein of this system (therefore indicated as CRISPR/Cas9).

Its locus is characterized by a non-coding trans-acting RNA (tracrRNA) located upstream of four Cas genes, but on the opposite DNA strand³⁶⁴ (**Fig.1.7**). The acquisition of the foreign sequence requires the action of all the four gene products and the tracrRNA.

Other than the above-mentioned Cas9, Cas genes include Cas1, Cas2 and Csn2 genes. Cas1 and Cas2 code for integrases necessary for the integration of foreign DNA into the CRISPR locus³⁶⁵. Csn2 gene, the signature gene for the type II A CRISPR systems, participates to this process presumably by binding the double-stranded DNA³⁶⁶. Experiments performed by Heler *et al.* demonstrated that the specific action of Cas9 is crucial for labelling the target DNA and recruiting the integrases Cas1 and Cas2³⁶⁷.

c) CRISPR in action

The type II CRISPR system isolated from *Streptococcus pyogenes* carries out targeted DNA double-strand breaks according to defined sequential steps, summarized below and shown in **Fig.1.7**.

When the cell is invaded for the second time by a pathogen, the leader sequence of the CRISPR locus acts as a promoter to transcribe the CRISPR-array into a premature crRNA transcript (pre-crRNA).

The noncoding trans-acting RNA (tracrRNA) is co-transcribed on the opposite strand of the CRISPR locus. TracrRNAs hybridize with the complementary short palindromic sequences contained in the pre-crRNA³⁶⁸ repeated sequence. Such hybridization recruits Cas9 (which stabilizes this interaction) and the endogenous RNase III, forming a complex that cleaves the premature transcript into mature crRNAs³⁶⁸. Since the cut occurs in the same site in all the hybridized regions, all the crRNAs are formed by a common part derived from the repeats together with a unique sequence corresponding to the specific spacer DNA³⁶⁶. The reliance on RNase III potentially explains why type II systems are limited to bacteria, as this enzyme is not generally found in archaea³⁶⁹.

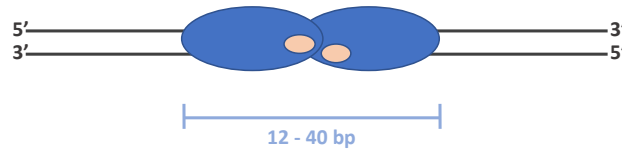
Finally, the mature tracrRNA:crRNA complex probes the foreign DNA to find its complementary regions adjacent to the protospacer adjacent motif (PAM) sequence which, in case of *Streptococcus pyogenes*, is represented by the trinucleotide 5'-NGG-3'³⁷⁰. When an appropriate PAM is identified, the nuclease Cas9 is recruited to reconstitute the so called interference complex. Cas9 transiently binds to the DNA and

initiates its unwinding starting from the first 10-12 nucleotides proximal to the PAM and moving distally^{371,372}. This DNA region is known as the seed sequence. Once the unwinding is completed, Cas9 can pair with the seed sequence together with the rest of the crRNA sequence. This generates both an RNA:DNA hybrid (called R-loop) formed by the crRNA spacer sequence and its complementary DNA sequence (the target strand) and a displaced ssDNA (the PAM-containing, non-target strand)³⁷³. Inefficient pairing causes the quick release of Cas9, which continues sampling other DNA sequences³⁷⁴. Conversely, if the pairing is successful, Cas9 cleaves both strands of the target DNA at a precise location (3 base pairs upstream of the PAM site) creating a blunt end double strand break.³⁷⁵

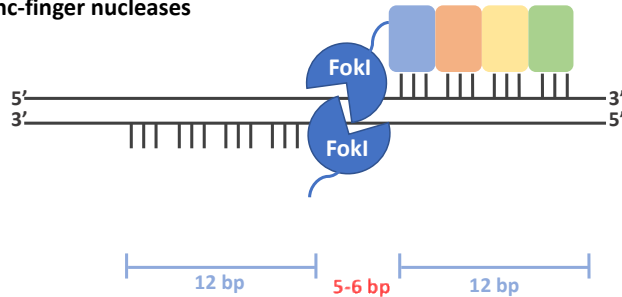
Biochemical and structural studies showed that the *S. pyogenes* Cas9 protein is composed by two lobes linked by two linker segments. The lobes are known as the nuclease (NUC) lobe and the recognition (REC) lobe. The interaction of these two domains form a positively charged groove that accommodates the RNA:DNA hybrid. The alpha-helical REC lobe is the domain that mostly interacts with the RNA components^{375–378} and is the most variable region of Cas9 protein; in fact several nucleotide variations reported here have been associated to the variable size encountered in different Cas9 orthologs.³⁷⁶

The nuclease lobe instead is dedicated to the cleavage of target DNA. Two defined domains within this lobe mediate the cleavage³⁷⁹: the HNH-like nuclease domain cleaves the target DNA strand, while the complementary DNA strand is cleaved by the RuvC-like nuclease domain. The RuvC-like domain also contains a C-terminal PAM-interacting domain, and is therefore the domain responsible for directing the entire nuclease to the target genomic region.

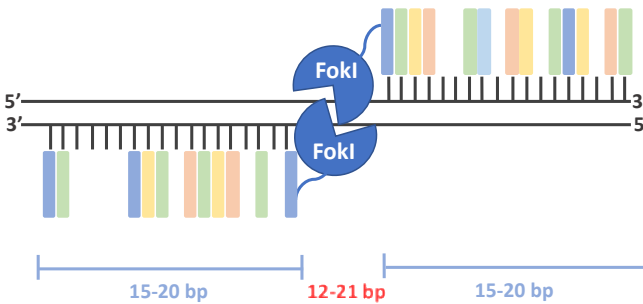
a) Meganucleases



b) Zinc-finger nucleases



c) Transcription activator-like nucleases



d) CRISPR/Cas9

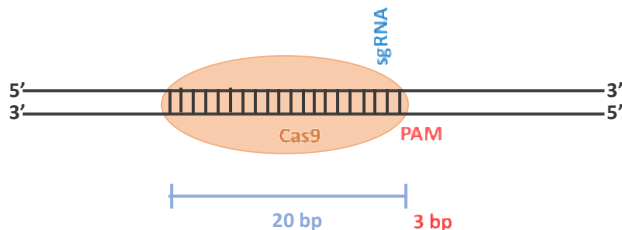


Figure 1.5. Schematic of the four nucleases classes.

a) Meganucleases are endogenous enzymes recognizing specific nucleotide sequences of 12-40 bp. They work either as monomers or dimers. b) Zinc-finger nucleases are composed by three to four zinc-finger modules fused to a FokI nuclease domain. Each zinc-finger recognizes 3 nucleotides. Zinc-finger nucleases work as juxtaposed couples separated by 5-6 nucleotides. c) Transcription activator-like (TALE) nucleases are composed by up to twenty TALE modules fused to a FokI nuclease domain. Each TALE recognizes 1 nucleotide. TALE nucleases work as juxtaposed couples separated by 12-21 nucleotides. d) CRISPR/Cas9 is composed by a sgRNA molecule and a Cas9 nuclease. Specific nucleotides within the sgRNA can pair to DNA sequences located upstream of trinucleotides motif called PAM.

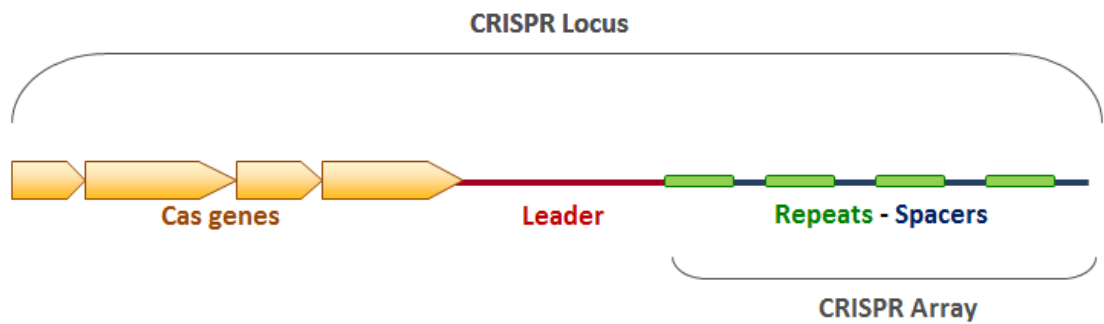


Figure 1.6. CRISPR locus in a bacterial chromosome.

The locus includes a CRISPR array, composed of a series of alternating spacers and palindromic direct repeats, between 21 and 47 bp in length. Spacer length is constant but, being derived from foreign internalized DNA, it is variable in sequence. Foreign DNA is recognized, cleaved and added to the array as new spacer thanks to Cas proteins, encoded by CRISPR-associated genes (Cas genes). The array is transcribed via the promoter included in the upstream leader sequence

***Streptococcus pyogenes* SF370 CRISPR locus 1**

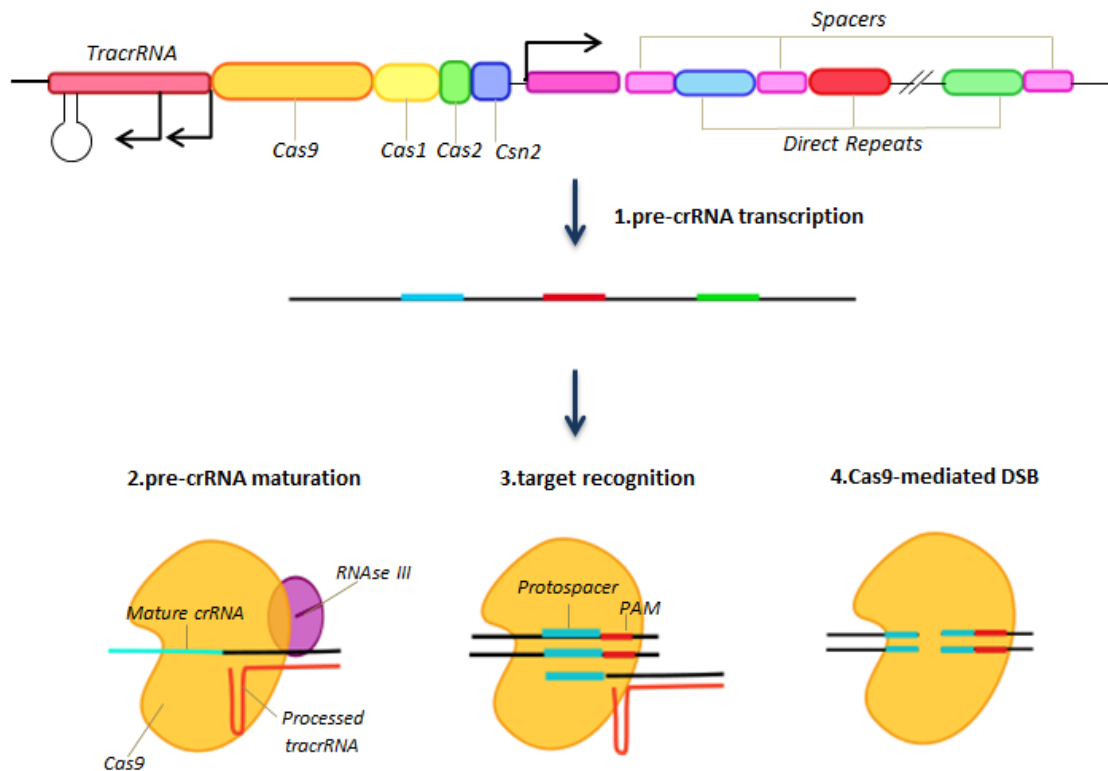


Figure 1.7. CRISPR/Cas type II system of *Streptococcus pyogenes*.

1) The transcription of the CRISPR locus generates a precursor RNA (pre-crRNA) from the CRISPR-array. 2) TracrRNA pairs with homologous crRNA sequences and recruits Cas9 and RNase III that cleave the pre-crRNA giving rise to mature crRNAs. 3) If foreign DNA from which crRNA derives infect again the cell, crRNA:tracrRNA complex can pair with sequences proximal to PAM. 4) Cas9 recruited at the genomic site targeted by the crRNA:tracrRNA complex cleaves the foreign DNA via double-stranded break.

1.8 CRISPR/CAS9 AS GENOME EDITING TOOL

Since the CRISPR/Cas type II of *Streptococcus pyogenes* can act using only Cas9, mature crRNA and a tracrRNA, the expression of these three essential components has been initially exploited for the development of a new RNA-guided genome editing system³⁸⁰ (**Fig.1.5.d**). Further improvements in the system were achieved by fusing the crRNA containing the target DNA sequence and the complementary tracrRNAs, so that the number of components required to mediate a targeted DNA cleavage is reduced. This chimeric structure, called single guide RNA (sgRNA or gRNA), allows to combine both the target specificity of the crRNA and the scaffolding properties of the tracrRNA in one single molecule³⁸⁰ (**Fig.1.8**).

sgRNA is composed of 80 nucleotides, of which the 20 nucleotides located at the 5' end are complementary to the target DNA sequence, while the remaining ones fold in a hairpin structure which mimics the complex formed by the interacting crRNA-tracrRNA. This structure allows a better sgRNA binding orientation and aids in Cas9 recruitment³⁸¹. Since the only strict requirement for a correct sgRNA targeting seems to be the presence of a PAM recognition motif downstream of the target sequence (**Fig.1.8**), implementation and scaling of the system have focused on the delivery of multiple sgRNAs in the same cells, so that multiplexed editing can be achieved³⁸².

Cas9 nuclease, in turn, has been refined through codon optimization and the addition of a nuclear localization signal (NLS) to accomplish heterologous expression in mammalian cells³⁸⁰ (**Fig.1.8**).

Exploiting the minimal type II CRISPR system, scientists successfully accomplished genome editing in a variety of experimental model systems, including bacteria³⁸³, plants³⁸⁴, animal models^{385–387} and mammalian cells^{388–390}.

Like other nuclease classes, the ability to preferentially exploit one of the two intrinsic cell repair pathways following double-stranded DNA breaks potentially expands the applicability of the CRISPR/Cas editing tool to any type of DNA mutation. Moreover, the ease of sgRNA customization makes CRISPR/Cas the preferential system for targeted *ex vivo* (and, above all, *in vivo*) genome editing for the treatment of inherited diseases, such as DMD, that are caused by a wide mutational spectrum³⁸⁸.

But questions remain concerning the specificity of CRISPR/Cas9 endonucleases and their potential off-target effects throughout the genome. Since sgRNAs are designed to recognize DNA stretches of 20 bp, these could occur more than once in the genome. Several algorithms have been designed to predict off-target effects³⁹¹. However, these do not take into consideration genomic variability among individuals, which could give

rise to unexpected “de novo” off-target sites. Therefore, despite its striking potential, optimization of CRISPR/Cas9-based genome editing is essential.

Considering its ability to introduce double-strand breaks in specific genomic loci, the CRISPR/Cas system has also been exploited for the generation of animal models of diseases^{392–395}. Furthermore, the Cas9 nuclease component underwent further engineering that allowed it to accomplish a variety of functions including modulation of gene expression³⁹⁶, epigenetic control³⁹⁷ and visualization of genomic loci³⁹⁸.

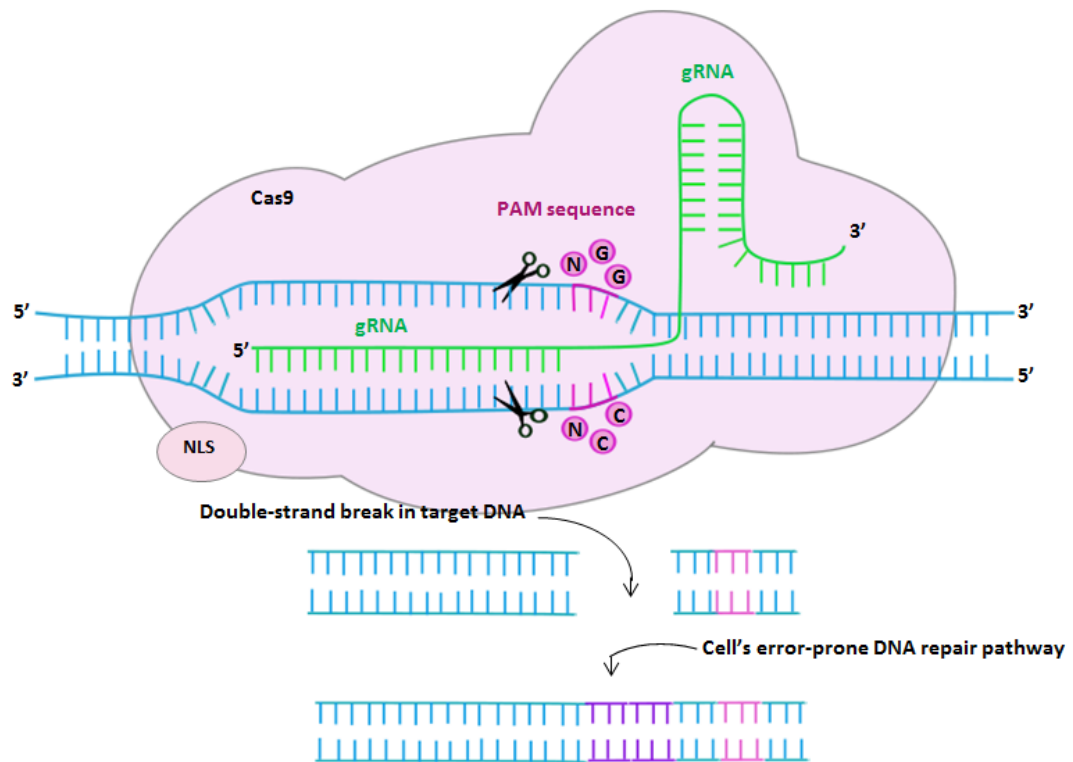


Figure 1.8. CRISPR/Cas9-mediated editing.

A 20 bp targeted genomic sequence located upstream of a 5'-NGG-3' sequence (PAM) is recognized by an RNA molecule 70 nt long called single guideRNA (gRNA or sgRNA). The first 20 nt of the gRNA are involved in pairing with DNA, while the remaining fold in an hairpin structure that recalls the Cas9 nuclease. In the figure, the Cas9 carries a nuclear localisation signal (NLS). Cas9 mediates a double strand cut, which can be repaired by HDR or by NHEJ. The latter is error-prone, leading to insertion/deletions in the restored template.

1.9 AIM OF THE PROPOSED WORK

I have used CRISPR/Cas9 genome editing as proof-of-concept for the development of a new therapeutic strategy aimed at **using a single CRISPR/Cas9 nuclease** to repair both small and large multi-exon duplications within *DMD* gene, which cause DMD.

This project arose from the consideration that any nuclease target sequence selected within the duplicated gene region will be present twice. Therefore, any CRISPR/Cas9 nuclease designed to target a duplicated region would create two double-strand breaks, one double-strand break at each repetition of the target locus. The use of a single nuclease will be potentially safer therapeutically, as the nuclease dose would be reduced compared to an approach based on two separate nucleases targeted to sites flanking the duplicon.

The starting hypothesis is that, following double-strand breaks, the cellular NHEJ intrinsic mechanism will repair the damage, eliminating the duplicated stretch of DNA between the repeated nuclease recognition site. Following this strategy, the duplication should therefore be precisely corrected, allowing the production of a wild type protein (**Fig.1.9**).

To avoid frame-shift mutations that could result from the indels that commonly originate at the NHEJ-repaired breakpoints, this study design is built on the use of CRISPR/Cas9 nucleases directed at an intronic region within the duplication. Specifically, I chose *DMD* intron 9 as a target, as this intron is included within the duplicated region in several available patient-derived cell lines and is duplicated in a range of patients whose mutation falls within the mutational hotspot of *DMD* exons 2-20¹.

The effectiveness of this hypothesis has been tested by:

1. Design of CRISPR/Cas9 nucleases to intron 9 of *DMD* using online tools
2. Determination of Cas9-mediated cleavage efficiency in HEK 293T cells
3. Testing of viral and non-viral methods to deliver CRISPR/Cas9 to patient-derived cells and identification of the most efficient approach
4. Confirmation of dystrophin repair in treated cells at the DNA, RNA and protein level.

This project will provide the basis to build future studies aimed at applying the CRISPR platform to correct duplications occurring throughout *DMD* gene and to repair

dystrophin in *in vivo* models of *DMD* duplications.

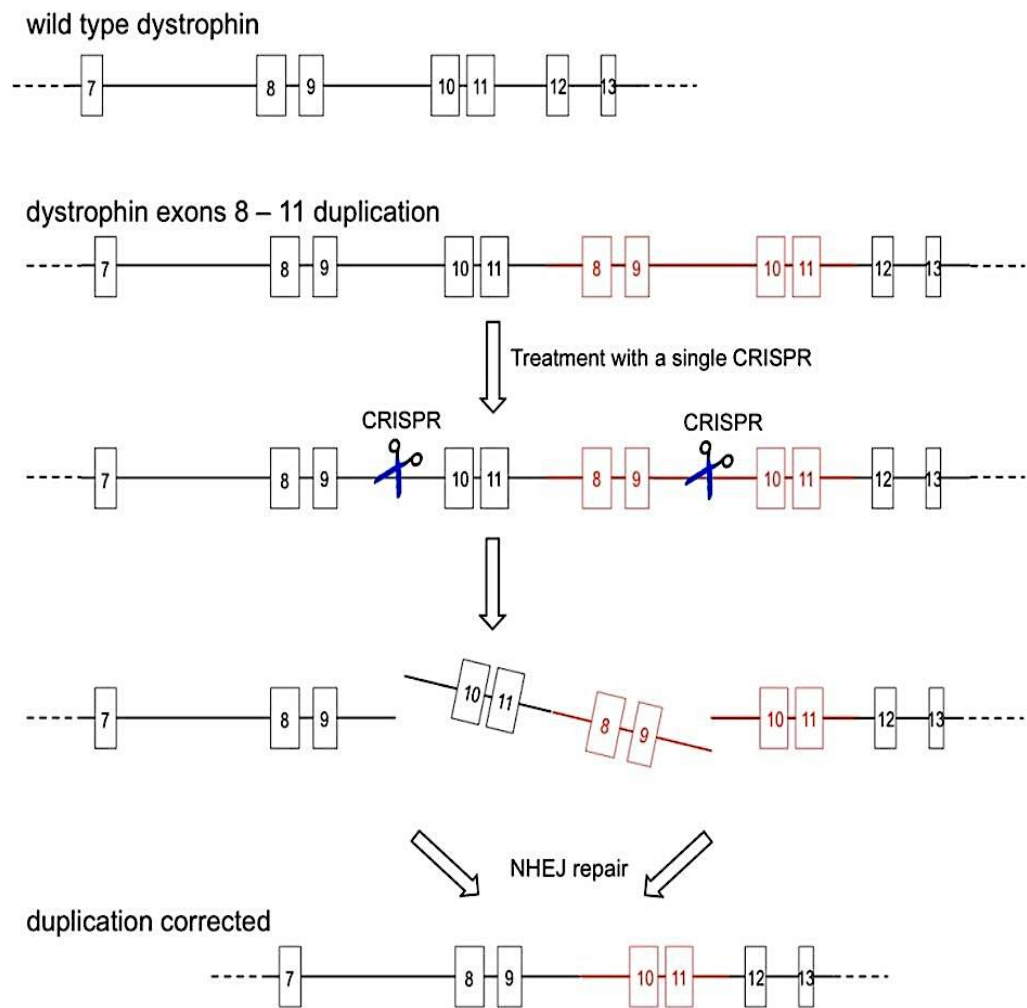


Figure 1.9. Schematic representation of CRISPR/cas9 mediated duplication repair in *DMD* gene.

A nuclease is targeted to *DMD* intron9, included in duplicated regions of cells available through the MRC Centre for Neuromuscular Diseases Biobank. The target site will be present twice in mutated *DMD* gene. Following nuclease-mediated cleavage, repair via NHEJ will eliminate the DNA segment between the two sites and the duplication will be corrected.

Chapter 2. Materials and Methods

2.1 MOLECULAR CLONING

2.1.1. Bacterial manipulation

a) Preparation of LB agar plates

For each 10 cm² dish, 30 ml of Luria Bertani (LB) agar was used. 1.5% LB agar plates were prepared by dissolving 15 g Bacto-agar (Sigma, Cat n° A5306) in 1.0 L of distilled water and sterilized by autoclaving for 1 hour at 120°C. As all the plasmids carried either an ampicillin- or kanamycin-resistance cassette, once the agar mix was cooled to 50°C in a water-bath, the antibiotic expressed by the bacterial plasmid was added at 100 µg/ml (Ampicillin) or 50 µg/mL (Kanamycin) (Sigma-Aldrich Cat n° A5354 and K1377). 10 ml of the mixture was then poured into the plates and left to solidify at room temperature.

b) Preparation of LB broth

LB broth was prepared by dissolving 15 g of LB broth powder containing 10 g/L tryptone, 10 g/L NaCl and 5 g/L yeast extract (Sigma, Cat n° L3522) in 1.0 L distilled water, then sterilized by autoclaving for 1 hour at 120°C. Once the LB broth was cooled to 50°C, the selected antibiotic was added at the concentration indicated in section 2.1.1.a.

c) Bacterial plating and growth

Bacterial colonies were grown by streaking a 10 µl tip previously inserted into bacterial stocks on LB agar plates containing the antibiotic expressed by the bacterial plasmids. Plates were incubated overnight at 37° and the following day, three bacterial colonies were selected and inoculated into 3-5 ml of LB broth supplemented with the appropriate antibiotic at the concentration given above (section 2.1.1.a.). Selected colonies were amplified by overnight incubation in an orbital shaker set at 225 rpm and 37°C. The following day, plasmid DNA was extracted from the bacterial growths (section 2.1.1.f) and used for the transformation of chemically competent *E.coli* bacteria.

d) Heat-shock transformation of chemically competent bacteria

Transformation occurred in the recombination-deficient chemically competent *StbI3* *E.coli* bacteria (One Shot® *StbI3*TM Life Technologies, Cat. n° C7373-03). This strain was selected as, being a recombinase-deficient strain, it lowers the chance of

homologous recombination of the Long-Terminal Repeats (LTRs) present in lentiviral plasmids³⁹⁹. Initially, Stbl3 bacteria stored at -80°C were thawed on ice for 10 minutes and incubated with 1-10 µl of plasmid DNA for 30 minutes on ice. Stbl3 cells were then heat shocked at 42°C for 15 seconds for opening the bacterial membrane pores allowing the entry of the vector, and incubated on ice for 2 minutes to reseal the pores and trap the DNA⁴⁰⁰. 250 µl of super optimal broth with catabolite repression (S.O.C) medium (ThermoFisher Scientific, Cat.n°15544034) was added to the transformed bacteria, that were then plated on LB agar plates supplemented by the appropriate antibiotic (section 2.1.1.a). Single colonies were isolated and amplified by overnight orbital shaking at 225 rpm at 37°C in 3-5 ml of LB broth supplemented with antibiotic at the above concentration. DNA extraction was then performed as indicated in section 2.1.1.f. to confirm the plasmid identity via Sanger sequencing (UCL Sequencing Facility, UCL).

e) Plasmid amplification

Bacterial growths derived from transformation were either used for extracting plasmid DNA (section 2.1.1.f) or further amplification. Amplification was performed by transferring 3-5 ml of bacterial culture into larger volumes (from 100 ml to 250 ml) of LB broth supplemented by antibiotic and overnight incubation in an orbital shaker at 225 rpm and set at 37°C.

f) Extraction of plasmid DNA from small volumes of bacterial cultures

QIAprep Spin Miniprep kit was used to extract plasmid DNA from 1 to 5 ml of bacterial cultures. Initially, bacteria were harvested following centrifugation for 10 minutes at 6800 g in a table-top centrifuge at 4°C. The supernatant was discarded and the pellet was resuspended in 250 µl of pre-chilled P1 lysis buffer containing RNase A (100 µg/µl), until no pellet remained. 250 µl of buffer P2 were added and the tube was gently inverted 6 times and incubated for 5 minutes at room temperature to complete the lysis reaction. The lysate was precipitated by adding 350 µl of N3 buffer and vigorously inverting the Eppendorf tube 6 times. Tubes were then centrifuged for 10 minutes at 17,000 g in a table-top microcentrifuge at room temperature. All the subsequent centrifugation steps were performed under the same conditions. This step was necessary to isolate the unwanted debris from the supernatant containing the plasmid DNA, which was transferred to a QIAprep 2.0 spin column. Following a centrifugation necessary for DNA binding, the silica column was washed with 750 µl of PE buffer and underwent a centrifugation to remove any residual ethanol. Elution in a clean Eppendorf tube was achieved by the addition of 30-50 µl of sterile water at pH 7.5 to the centre of the column followed by 1-minute incubation at room temperature and centrifugation.

Eluted plasmid DNA was then quantified as detailed in section 2.3.2 and stored at -20°C.

g) Extraction of plasmid DNA from large volumes of bacterial cultures

The QIAGEN Plasmid Maxi Kit was used when plasmid DNA had to be extracted from bacterial culture of volumes of 250-300 ml. Bacteria were centrifuged for 30 minutes at 6800 g at 4°C. The harvested pellet was isolated and resuspended in 10 ml of cold P1 lysis buffer containing RNase A (100 µg/µl) in a 50 ml Falcon tube. 10 ml of P2 buffer was then added to the Falcon tube which was inverted 6 times and left at room temperature for 5 minutes, when 10 ml of pre-chilled buffer P3 were added. The tube was inverted 6 times, incubated on ice for 5 minutes and centrifuged for 20 minutes at 6800 g in a centrifuge at 4°C. The resulting supernatant was then transferred to a QIAGEN-tip 500 column previously equilibrated by letting 10 ml of QBT buffer flow through it by gravity. Once all the supernatant was filtered through it, the column was washed twice with 30 ml of QC buffer and finally transferred to a new 50 ml Falcon used as a collection tube for DNA elution, which occurred by flow-through of 15 ml of QF buffer. Eluted DNA was precipitated upon mixing with 10.5 ml isopropanol at room temperature and a 1-hour centrifugation performed at 6800 g in a centrifuge set at 4°C. The supernatant was carefully removed and the pellet was washed with 5 ml of 70% ethanol. A further centrifugation step of 30 minutes at 6800 g was then carried out at room temperature. The DNA pellet was air-dried for 10 minutes at room temperature and dissolved in 250-300 µl of TE buffer at pH 8.

2.1.2. List of vectors

a) LentiCRISPRv1 vector

The integrating LentiCRISPRv1 plasmid (Addgene 49535, **Fig.2.1**) from Feng Zhang's laboratory was used to express CRISPR/Cas9 in HEK 293T cells and primary fibroblasts⁴⁰¹. sgRNA expression was driven by the U6 promoter, while Cas9 expression was controlled by the core subunit of the elongation factor 1α (EF1α) core promoter. The plasmid also carried an ampicillin and puromycin resistance cassette, responsible for the bacterial and cell selection, respectively.

b) pL-CRISPR.EFS.GFP vector

The integrating pL-CRISPR.EFS.GFP plasmid from Benjamin Ebert's laboratory (Addgene 57818, **Fig.2.2**) was used to express CRISPR/Cas9 in primary myoblasts⁴⁰². sgRNA expression was driven by the U6 promoter, while bicistronic Cas9 and enhanced green fluorescent protein (EGFP) expression was controlled by the core subunit of the

EFS (EF1 α short) promoter. The plasmid also carried an ampicillin resistance cassette, responsible for the bacterial selection.

c) LentiCRISPR-EGFP sgRNA5 vector

The integrating LentiCRISPR-EGFP sgRNA5 plasmid (Addgene 51746, **Fig.2.3**) from Feng Zhang laboratory expressed a sgRNA (sgRNA5) targeted to a sequence of the EGFP protein and the Cas9 nuclease⁴⁰¹. The same promoters as for LentiCRISPRv1 (section 2.1.2.a.) drove the expression of CRISPR/Cas9 components. The plasmid also carried an ampicillin and puromycin resistance cassette, responsible for the bacterial and cell selection, respectively.

As EGFP protein is not expressed within the human genome, this vector served as a negative control for the Cas9-mediated cleavage in HEK 293T and primary fibroblasts.

d) pLJM1-EGFP vector

The integrating pLJM1-EGFP plasmid (Addgene 19319, **Fig.2.4**) from David Sabatini's laboratory expressed EGFP via the constitutive cytomegalovirus (CMV) promoter, together with an ampicillin and puromycin resistance cassettes responsible for the bacterial and cell selection, respectively⁴⁰³.

As cells expressing EGFP indicate successful transduction, this transgene was used as a positive control transgene to monitor the ability of lentiviral particles which express it to transduce primary fibroblasts.

e) psPAX2 vector

PsPAX2 plasmid (Addgene 12260, **Fig.2.5**) from Didier Trono's laboratory expressed the HIV-1 packaging components gag and pol, controlled by the CAG hybrid promoter derived by the CMV enhancer fused to the chicken beta-actin promoter, as well as an ampicillin resistance cassette whose expression is driven by the constitutive CMV promoter.

f) pCMV-VSV-G vector

pCMV-VSV-G transgene (Addgene 8454, **Fig.2.6.**) from Bob Weinger's laboratory expressed the G-protein of the vesicular stomatitis virus (VSV-G) by means of the constitutive CMV promoter, together with an ampicillin resistance cassette under the control of the ampicillin resistance promoter⁴⁰⁴.

g) pLV-U6g-EPCG vector

Commercially available lentiviral particles purchased from Sigma-Aldrich expressed the pLV-U6g-EPCG lentiviral transgene (**Fig.2.7**). Other than a puromycin antibiotic resistance cassette, the vector co-expressed sgRNA2, Cas9 and EGFP, allowing me to choose the best selection method for the cell type to be used. Similarly to the LentiCRISPRv1, LentiCRISPR-EFGP sgRNA5 and pL-CRISPR.EFS.GFP lentiviral vectors, sgRNA was expressed by the constitutive U6 promoter, while the expression of Cas9 and the flanking cassettes was driven by the truncated human elongation factor-1 α (tEF1 α) promoter.

h) U6gRNA-CMVCas9-GFP vector

U6gRNA-CMVCas9-GFP expressing sgRNA0 and sgRNA2 (from here on named CRISPR0 and CRISPR2 σ) were purchased from Sigma-Aldrich (**Fig.2.8**). sgRNA expression was driven by the constitutive U6 promoter. CRISPR2 σ expressed sgRNA2, while CRISPR0 expressed sgRNA0 which recognizes a sequence that is not present in the human genome and was therefore a negative control. The plasmid also harboured an EGFP-selection cassette, co-expressed with Cas9 by the CMV promoter.

i) pEGFP-C2 vector

The pEGFP-C2 vector was purchased from Clontech (**Fig.2.9**). EGFP expression was driven by the constitutive CMV promoter. Neomycin and Kanamycin resistance were also expressed by the simian virus (SV) 40 promoter.

j) pCMV-GFP vector

pCMV-GFP plasmid (Addgene 11153) (**Fig.2.10**) from Connie Cepko's laboratory expressed EGFP by means of the constitutive CMV promoter and ampicillin via the ampicillin-resistance promoter⁴⁰⁵.

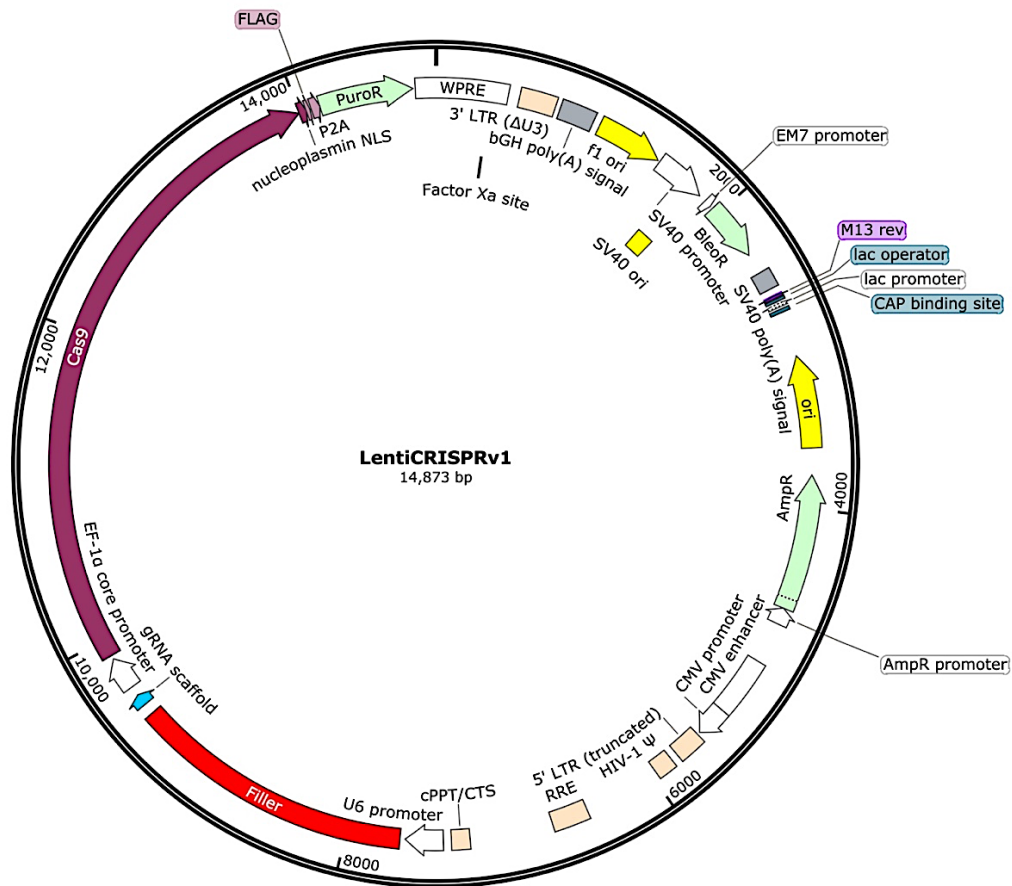


Figure 2.1. Schematic of lentiviral LentiCRISPRv1 plasmid.

The LentiCRISPRv1 plasmid expressed sgRNA by the U6 promoter and Cas9 by the EF1 α core promoter. The antibiotics expressed were puromycin (co-expressed with Cas9) and ampicillin. LTRs were responsible for the vector integration into the host genome.

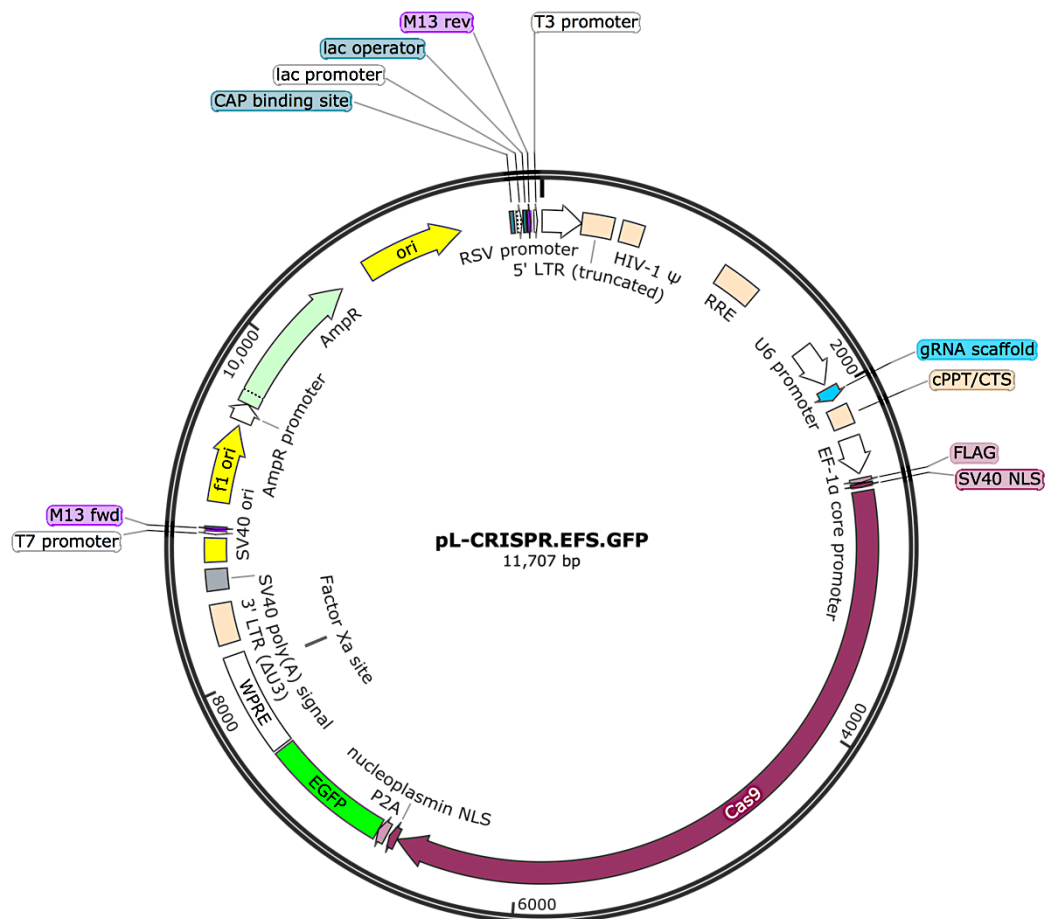


Figure 2.2. Schematic of pL-CRISPR.EFS.GFP plasmid.

The integrating pL-CRISPR.EFS.GFP plasmid expressed sgRNA by the U6 promoter and Cas9 by the EFS promoter, which also drove bicistronic EGFP expression. The antibiotic expressed was ampicillin. LTRs were responsible for the vector integration into the host genome.

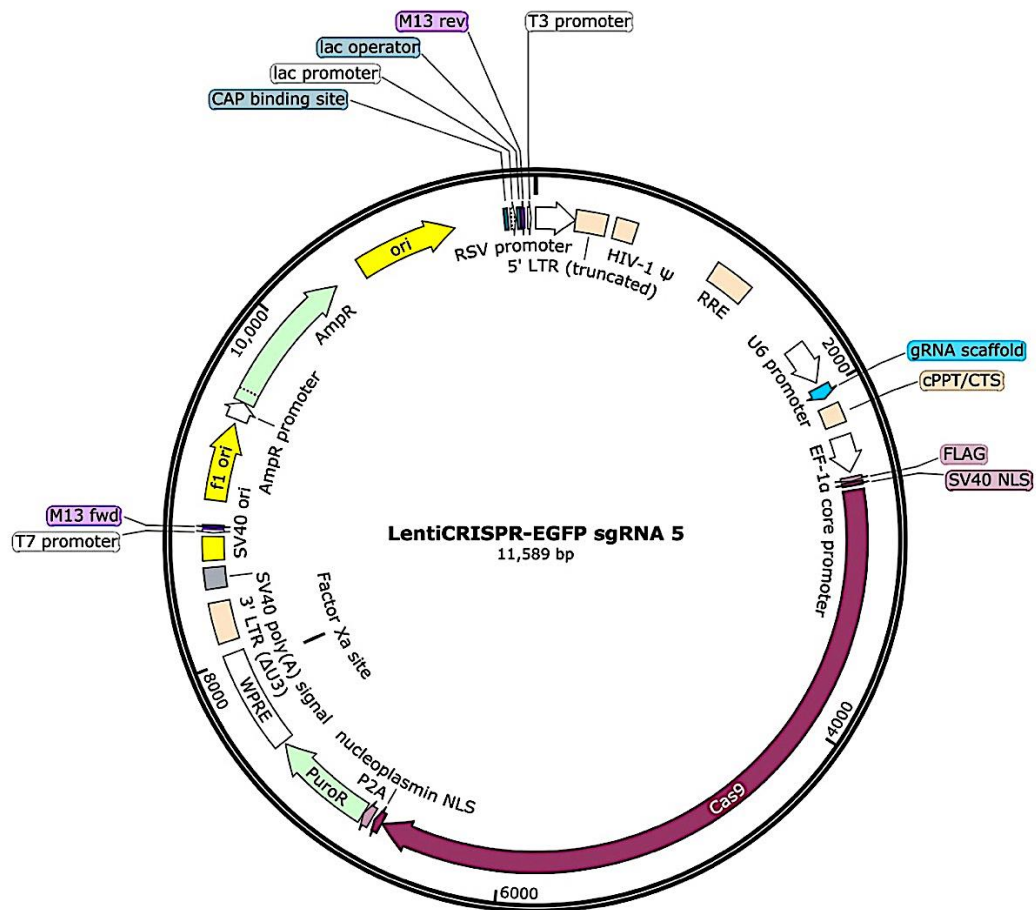


Figure 2.3. Schematic of LentiCRISPR-EGFP sgRNA5 plasmid.

The integrating LentiCRISPR-EGFP sgRNA5 plasmid expressed sgRNA by the U6 promoter and Cas9 by the EF1 α core promoter. Similarly to LentiCRISPRv1, the antibiotic expressed were puromycin (co-expressed with Cas9) and ampicillin. LTRs were responsible for the vector integration into the host genome.

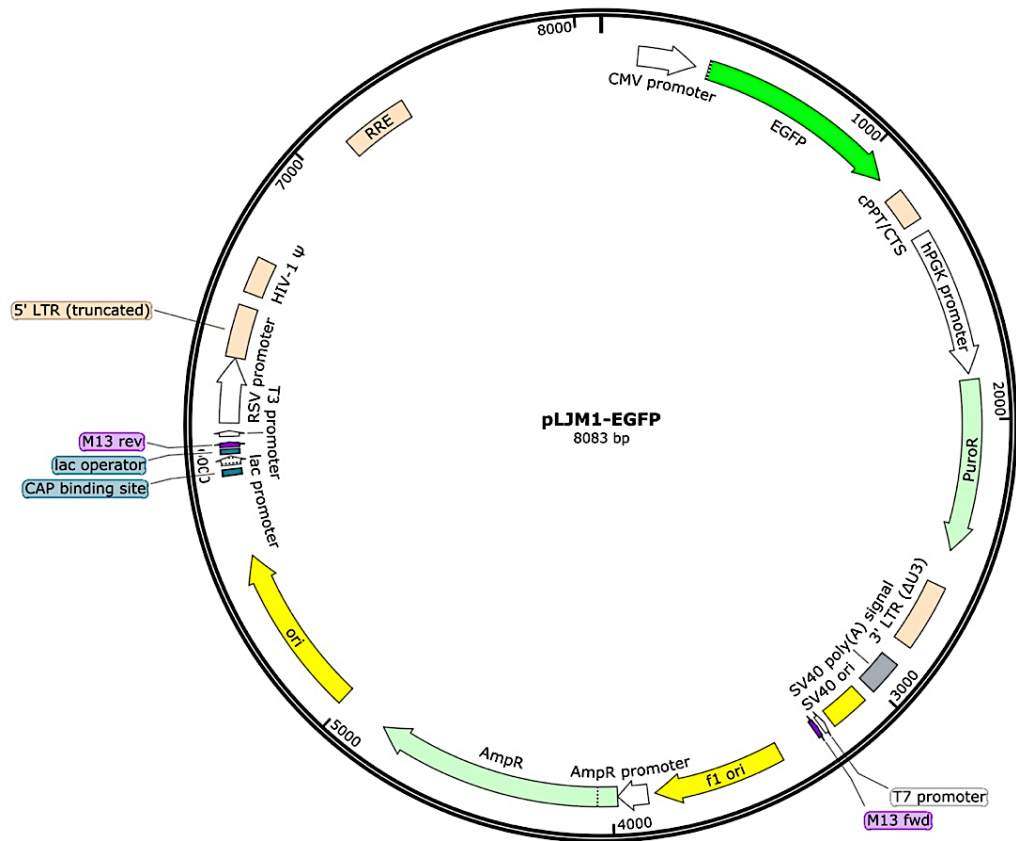


Figure 2.4. Schematic of pLJM1-EGFP.

The integrating pLJM1-EGFP plasmid expressed EGFP by the CMV promoter and the antibiotics ampicillin and puromycin (driven by the human phosphoglycerate kinase (PGK) promoter). LTRs were responsible for the vector integration into the host genome.

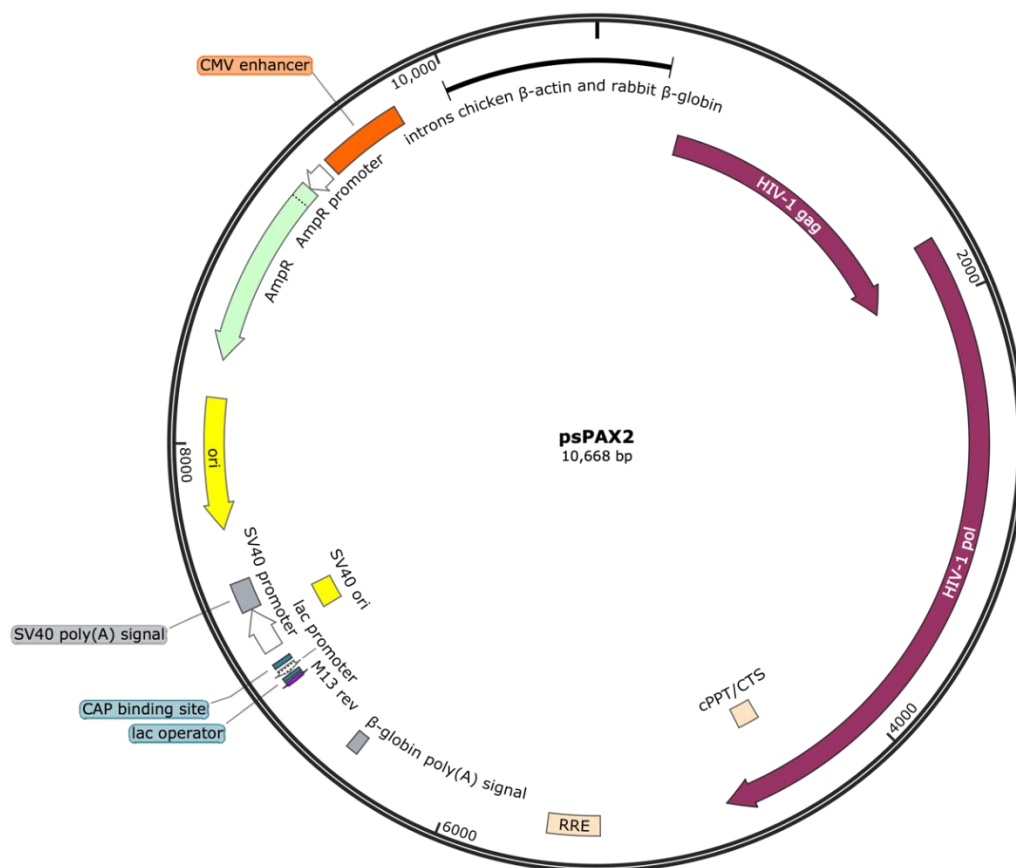


Figure 2.5. Schematic of psPAX2 packaging plasmid.

psPAX2 plasmid expressed the packaging components (HIV-1 gag and pol) by the hybrid CAG promoter and the antibiotic ampicillin.

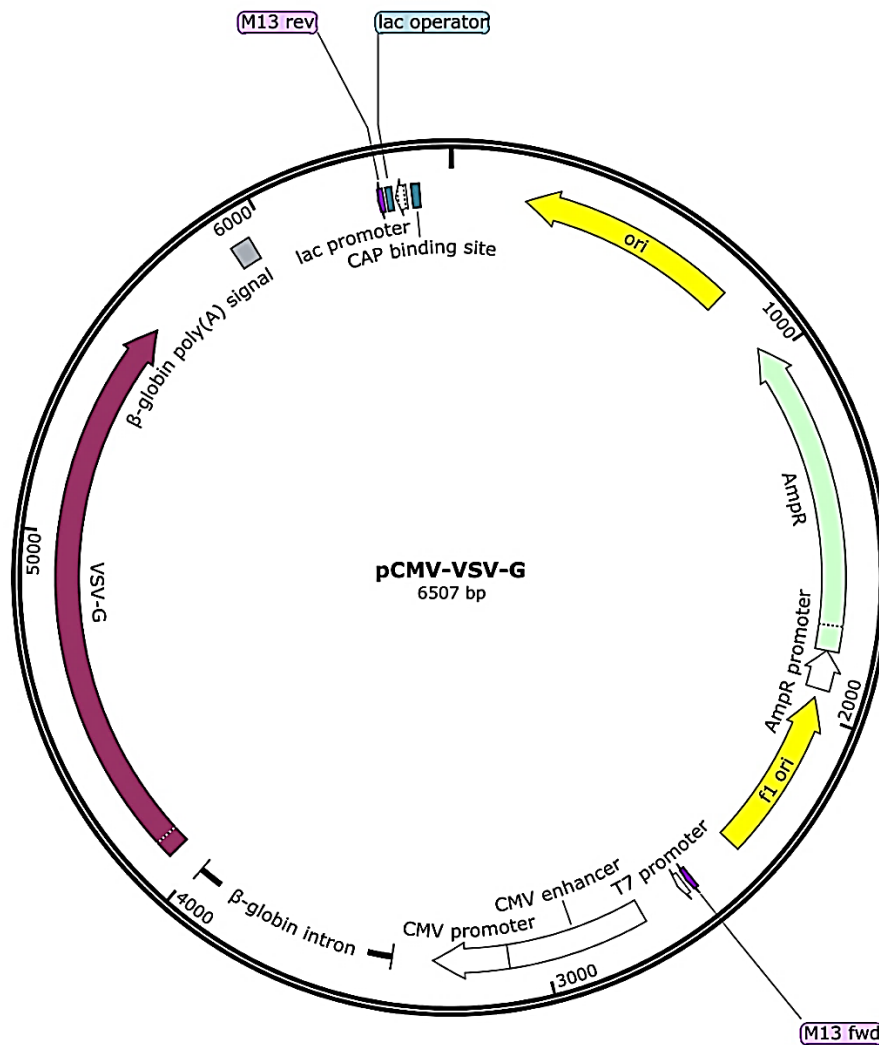


Figure 2.6. Schematic of pCMV-VSV-G plasmid.

pCMV-VSV-G plasmid expressed the G-protein of vesicular stomatitis virus envelope component by the CMV promoter. Ampicillin was also expressed.

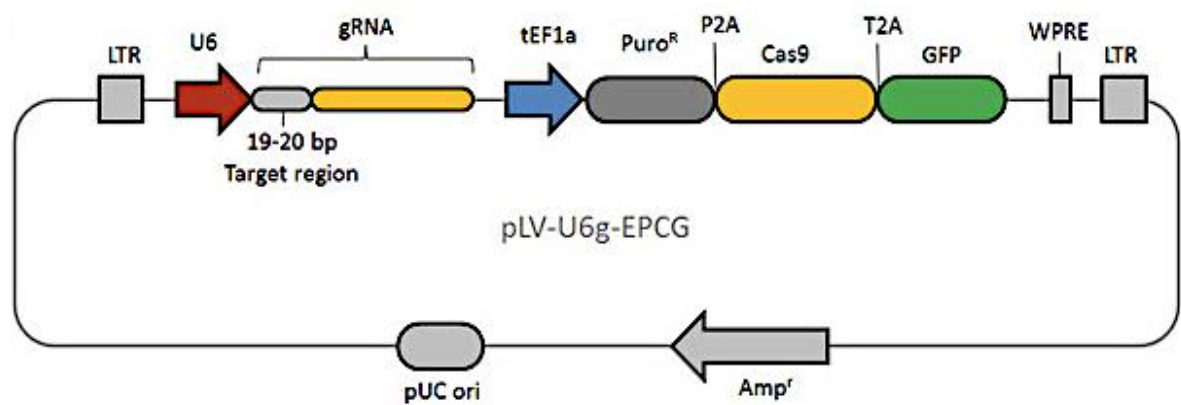


Figure 2.7 Schematic of the pLV-U6g-EPCG lentiviral vector.

The integrating pLV-U6g-EPCG vector expressed sgRNA by the U6 promoter and co-expressed Cas9 and EGFP together with puromycin by the tEF1 α promoter. Ampicillin was also expressed. LTRs were responsible for the vector integration into the host genome.

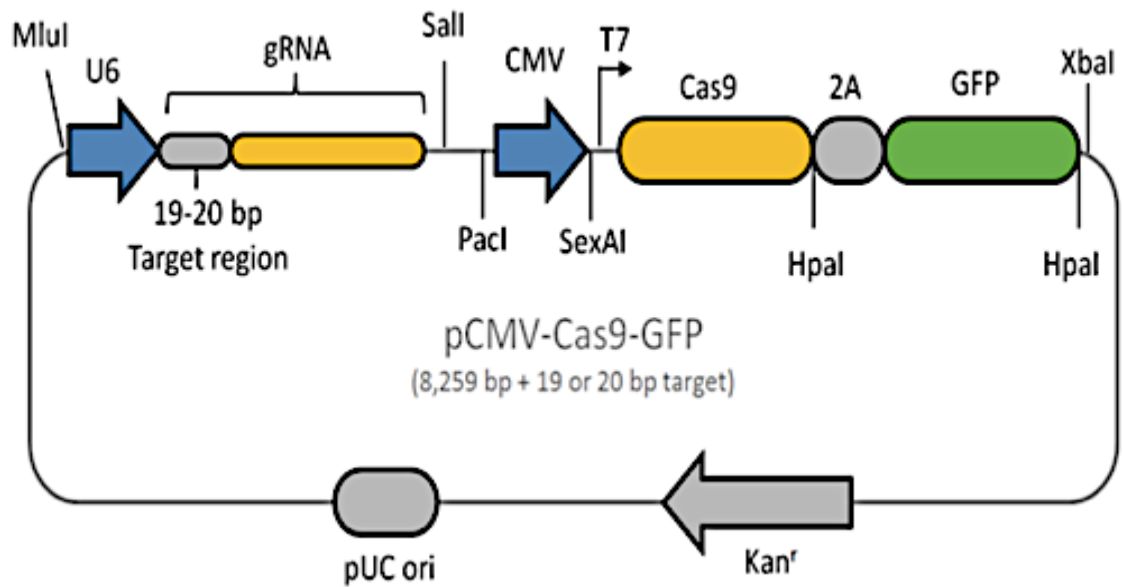


Figure 2.8 Schematic of the pCMV-Cas9-GFP vector.

The pCMV-Cas9-GFP vector expressed sgRNA by the U6 promoter and co-expressed Cas9 and EGFP by the CMV promoter. Kanamycin was also expressed.

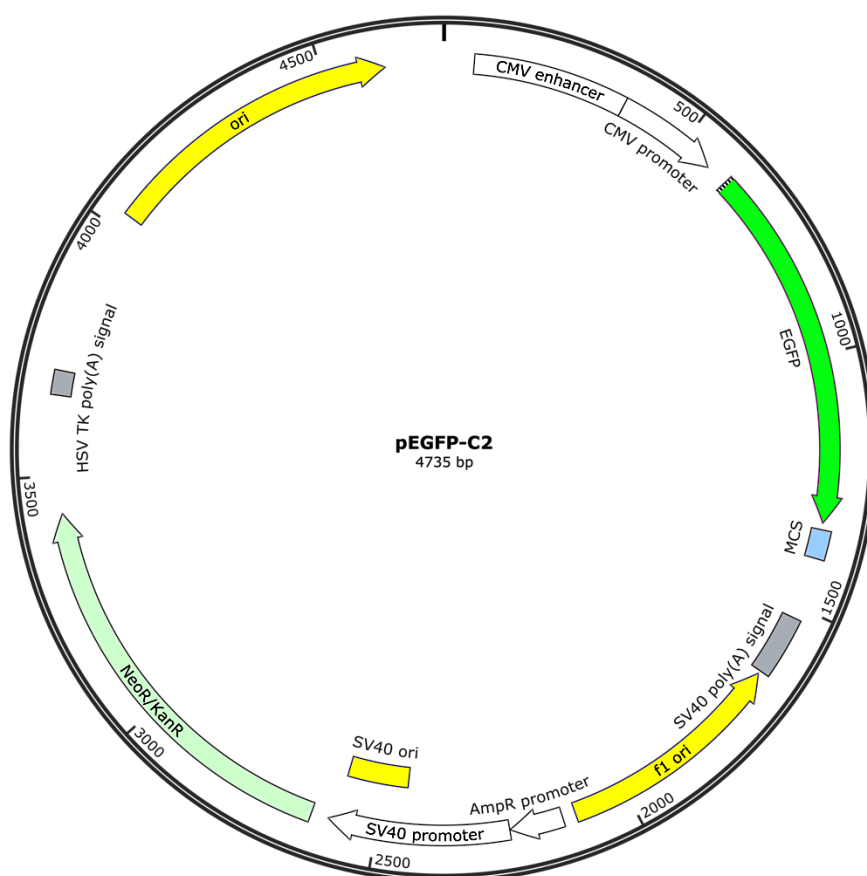


Figure 2.9. Schematic of the pEGFP-C2 vector.

The pEGFP-C2 vector expressed EGFP by the CMV promoter. Neomycin and kanamycin resistance were also expressed.

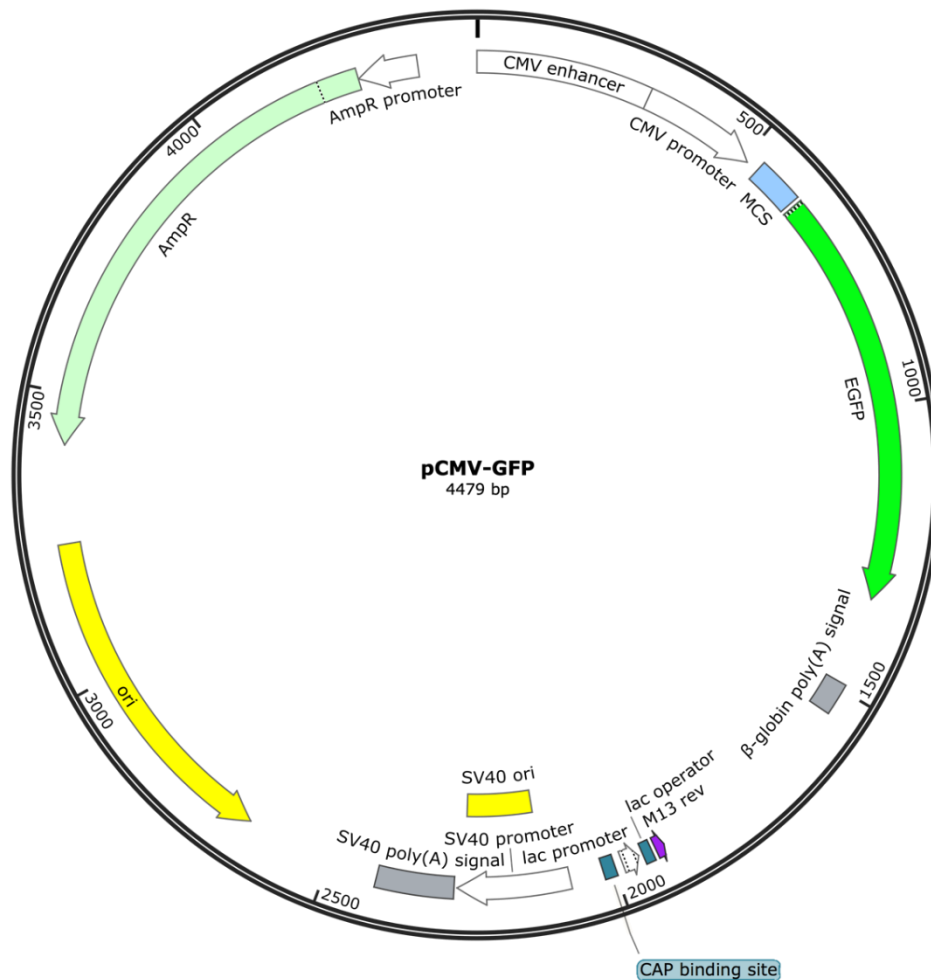


Figure 2.10. Schematic of the pCMV-GFP vector.

The pCMV-GFP vector expressed EGFP by the CMV promoter. Ampicillin resistance was also expressed.

2.1.3 Plasmid verification

Restriction digest was used to confirm the identity of the plasmids described in section. 2.1.2. Different sets of restriction enzymes (New England BioLabs) (**Fig. 2.11**) were used for the specification of each vector (apart for pLV-U6g-EPCG, as it was purchased as lentiviral particles).

psPax2 and pCMV-VsVg plasmids were digested with the dual cutter EcoRI, while CRISPR0 and CRISPR2 σ were processed with XhoI and BamHI restriction enzymes, according to the manufacturer's protocol.

LentiCRISPRv1, LentiCRISPR-EGFP sgRNA5, pLMJ1-EGFP and pL-CRISPR.EFS.GFP vectors carrying LTRs were digested with AflIII dual cutter, to ensure that recombination among LTRs did not occur⁴⁰⁶.

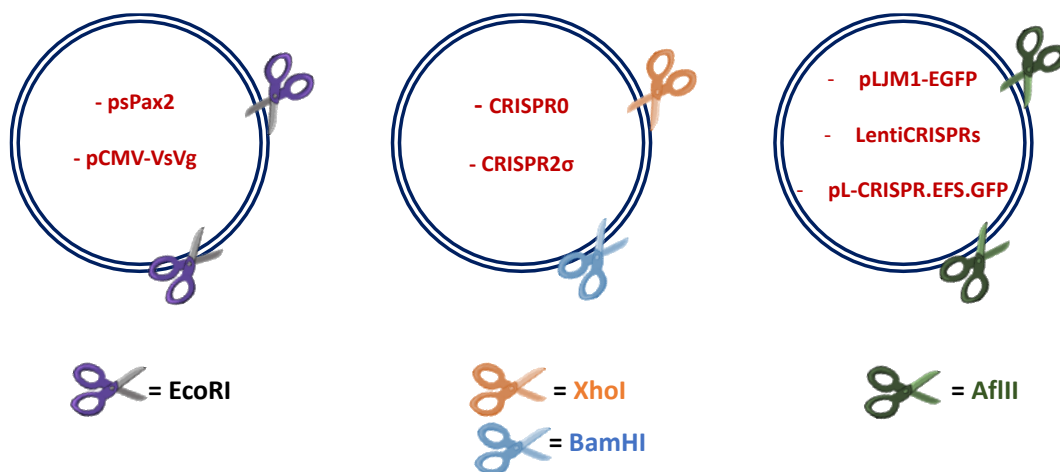


Figure 2.11 Schematic of restriction digest.

The expected molecular size for each restriction digest is:

- psPax2 digestion = 4,374 kb and 6,329 kb
- pCMV-VsVg digestion = 1,572 kb and 4,791 kb
- LentiCRISPRv1 digestion = 7,927 kb and 3,661 kb
- LentiCRISPR-EGFP sgRNA5 digestion = 5,030 kb and 9,843 kb
- pLMJ1-EGFP digestion = 3,656 kb and 4,427 kb
- pL-CRISPR.EFS.GFP digestion = 8,045 kb and 3362 kb
- CRISPR0 and CRISPR2 σ digestion = 2,745 kb and 5,513 kb

2.1.4 Bacteria storage

Bacteria were stored at -80°C, in 50% (v/v) glycerol (Sigma) and LB broth supplemented by the appropriate antibiotic (ampicillin or kanamycin).

2.1.5 Generation of functional CRISPR/Cas9 vectors

a) Identification of CRISPR/Cas9 genomic target sequences and sgRNAs.

CRISPR/Cas9 target sequences were identified within intron 9 of the DMD gene using the bioinformatics web-tool developed by Feng Zhang's laboratory at Massachusetts Institute of Technology (<http://crispr.mit.edu/>) which required input sequences of 23-500 bp and to specify the studied organism. The full genomic sequence of the human DMD intron 9 was obtained from the Ensembl Genome Browser (www.ensembl.org/) and four regions spanning from 176 to 225 bp throughout the intron were selected on the basis of the absence of known polymorphisms. The algorithm returned a list of sgRNAs, ranked according to the presence of the -NGG- PAM sequence. sgRNAs score was inversely correlated to their likelihood of targeting unwanted sites. For each of the four genomic inputs, the sgRNA with the highest score (considered the best sgRNA) was selected and then cloned into commercially available lentiviral vectors as indicated in section 2.1.5.d.

b) Digestion of lentiviral plasmids expressing CRISPR/Cas9

LentiCRISPRv1 vector pXPR-001 (Addgene 52961) was digested with the restriction enzyme BsmBI in order to remove the 2 kb filler in the vector's backbone and allow sgRNAs cloning, according to the protocol provided by Zhang's laboratory⁴⁰⁷, detailed below.

5 µg of the purified LentiCRISPRv1 vector were digested for 30 minutes at 37°C with 3 µl of the Fast Digest BsmBI (Fermentas) mixed with 3 µl of 10X Fast Digest buffer, 3 µl of the FastAP alkaline phosphatase (ThermoFisher Scientific) (which de-phosphorylates the digested plasmids ends preventing them from re-ligation) and water in a total volume of 60 µl. 10 µl of SYBR® Safe DNA stain 6X (Invitrogen) were added to the 60 µl reaction to allow the visualization of the digestion products following the exposure of the 1% agarose gel to the UV light from the transilluminator (UVP Bio-Imaging System). A 1 kb molecular marker (New England BioLabs) was used to determine the size of the digestion products, expected to be approximately 11 kb and 2 kb. The higher molecular weight band was gel purified by using the QIAquick Gel Extraction Kit (QIAGEN) according the manufacturer's instructions (section 2.3.6.b).

These experimental steps were also used for the digestion of the pL-CRISPR.EFS.GFP

plasmid.

c) sgRNAs oligos design

As the BsmBI digestion results in a staggered cut that generates cohesive ends, pairs of sgRNA oligos were designed with additional sequences outside their complementary region to ease their ligation with the plasmid overhangs. For each guide, two complementary oligonucleotides were designed so to have one of the following extra sequences, represented by 5'CACCG and 3' CAAA on the forward and reverse strand, respectively, as shown below:



To avoid self-cleavage, the sgRNA sequences did not encompass the PAM motif.

Oligos designed for each sgRNA are indicated below; capital letters correspond to the target sequence, while lowercase letters refer to the overhangs.

- sgRNA 1 primers:

1a. caccgACTACCCCGTCTCTATTACT

1b. TGATGGGGCAGAGATAATGAcaaa

- sgRNA 2 primers

2a. caccgGTACCTCAACAAGAAAGTTA

2b. CATGGAGTTGTTCTTTCAATcaaa

- sgRNA 3 primers

3a. caccgGACCATTGTTGGGTGTCGTCTT^L_{SEP}

3b. CTGGTAAACCCACAGCAGAAcaaa

- sgRNA 4 primers

4a. caccgCCTTCATTAAGGCATCGCTC^L_{SEP}

4b. GGAAGTAATTCCGTAGCGAGcaaa

d) sgRNAs cloning into the LentiCRISPRv1 vector

sgRNA oligos were annealed and phosphorylated to allow their ligation into the digested LentiCRISPRv1 vector. This was carried out by mixing 1 µl of each 100 µM sgRNA oligos with 10 µl of water, 0.5 µl of T4 PNK Enzyme (NEB Cat. n° M0201S) and 1 µl of 10X T4 Ligation Buffer (NEB) which, unlike the PNK associated buffer, is supplemented with ATP, which facilitates the phosphorylation process. The mixture was placed in the thermocycler (Applied BioSystem), where the reaction took place starting from a temperature of 37°C for 30 minutes, 95°C for 5 minutes and a decrease down to 20°C, at 5°C/minute. Annealed and phosphorylated oligo duplexes were then diluted 1:200 in sterile water. Ligation in to the digested vector occurred by mixing 1 µl of annealed oligos and 5 µl of the linearized LentiCRISPRv1 with 11 µl of pure water and 1 µl of Quick Ligase (NEB Cat. n° M2200S). The reaction mix was then incubated for 10 minutes at room temperature. Ligation products were used to transform recombinase-deficient One Shot® Stbl3 *E. coli* bacteria (section 2.1.1.d.).

These experimental steps were also used for the sgRNA cloning into the digested pL-CRISPR.EFS.GFP vector.

2.1.6 Production of integrating lentiviral particles

For each lentiviral transgene, the day before transfection, HEK 293T cells were plated into 12 100 mm culture dishes (Corning®) at a density of 6×10^6 cells per plate, in Dulbecco-MEM medium supplemented with 1% v/v glutamine and 10% FBS.

After 24 hours of incubation at 37°C and 5% CO₂, plasmids necessary for the assembly of lentiviral particles were transfected into the cells by means of the 3 µg/µl FuGENE®6 transfection reagent, using a DNA:FuGENE®6 ratio of 1:3. Transfected plasmids included the transgene (16 µg/plate), psPax2 plasmid (Addgene 12260, 12 µg/plate) and pCMV-VSVg plasmid (Addgene 8454, 4 µg/plate), expressing the packaging and envelope components. PsPax2 expresses both packaging and structural components (including the Gag, Pol, Rev and Tat genes) required for the integration and expression of the viral vector into the host cell⁴⁰⁸. pCMV-VSVg instead specifies the G-protein of the vesicular stomatitis virus (VSVg), which is the heterologous envelope component that determines the target specificity of the viral vector and, due to its high stability and wide tropism⁴⁰⁹, is often used to pseudotype lentiviral vectors. Transgenes were represented by either LentiCRISPRv1, LentiCRISPR-EGFP sgRNA5, pL-CRISPR.EFS.GFP or pLJM1-EGFP.

Initially, 3 µl of FuGENE®6 were added to 3 ml of OptiMem and incubated for 5 minutes at room temperature. The mix of psPax2, pCMV-VsVg and selected transgene was then added to the OptiMem- FuGENE®6 and incubated at room temperature for further 30 minutes, following which they were added dropwise to the HEK 293T cells (**Fig.2.12**). Transfected cells were incubated at 37°C and 5% CO₂ to allow the assembly and release of lentiviral particles into the culture medium.

At days 3 or 4 post-transfection, cell culture medium was collected and filtered through a 40 µm filter (Corning®) to remove cells and debris. Filtered media, containing lentiviral particles, was placed in 25 x 83 mm polyallomer centrifuge tubes (Beckman Coulter) and centrifuged for 2h at 60000 x g at 4°C in a Sorvall Discovery 90SE centrifuge. Following ultracentrifugation, supernatant was discarded and viral particles were collected and resuspended in 200 µl of Opti-MEM®. Generated lentiviral particles were aliquoted and stored at -80.

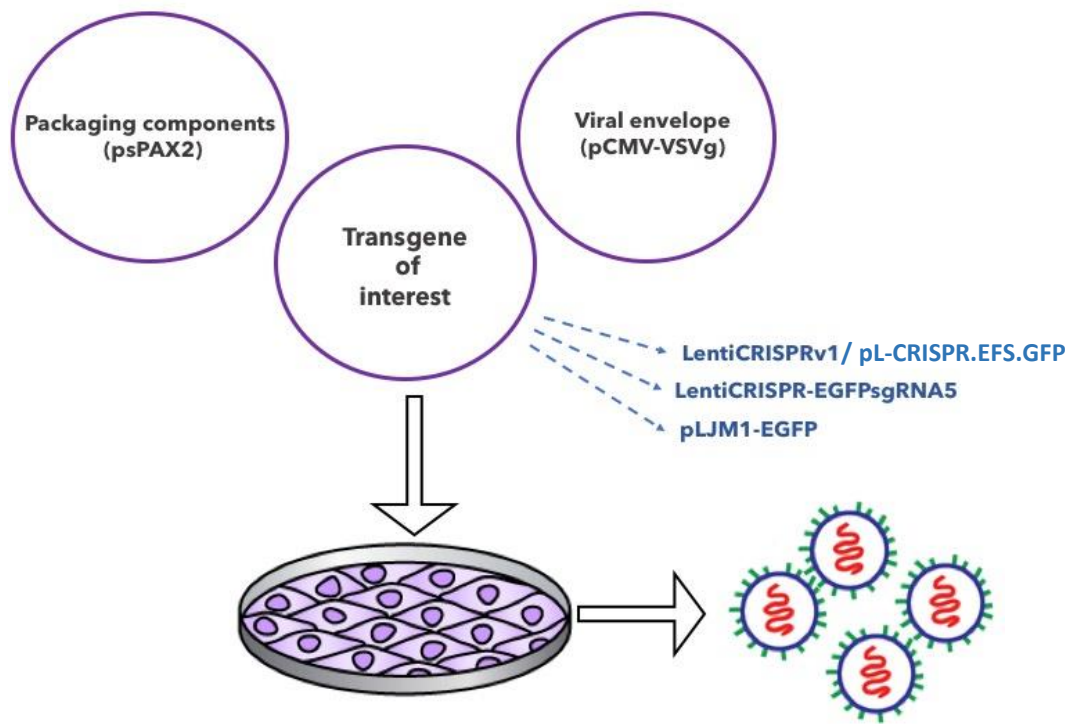


Figure 2.12 Schematic of lentiviral particles production.

Three independent plasmids expressing the packaging, envelope and transgene components were co-transfected into HEK 293T cells. Packaging and envelope components are specified by the psPAX2 and pCMV-VSVg plasmids, respectively. Defined transgene plasmids were used for the production of separate lentiviral particles. These are represented by the LentiCRISPRv1, pL-CRISPR.EFS.GFP, LentiCRISPR-EGFP expressing sgRNA5 (not targeting a genomic region) and the pLJM1-EGFP vector, expressing the EGFP protein. Transfected HEK 293T assembled lentiviral particles and released them into the culture medium, from which they were harvested.

2.1.7 Titration of generated lentiviral particles

Titration was performed by quantifying the number of integrated lentiviral copies in target cell genomes by qPCR, a method which exploits specific wavelengths to excite a specific dye (named SYBRGreen) able to fluoresce when intercalated into double stranded DNA. The intensity of the fluorescent signal in target regions is therefore indicative of the number of lentiviral particles integrated into the transduced cells⁴¹⁰.

Initially, 1×10^5 HEK 293T cells were seeded in each well of a 6- well plate, and transduced after 24 hours with increasing volumes (1 μ l, 5 μ l, 10 μ l, 20 μ l) of lentiviral particles expressing each of the transgenes (i.e. LentiCRISPRv1, LentiCRISPR-EGFP sgRNA5, pL-CRISPR.EFS.GFP or pLJM1-EGFP). Culture medium (section 2.2.4) was replaced the following day and cells were maintained in culture until day 5, when they were washed with 500 μ l phosphate buffer saline (PBS), detached from the plate with 500 μ l of trypsin (section 2.2.8), collected and centrifuged at 1000 g for 4 minutes at room temperature. Genomic DNA was extracted from the resulting pellet (section 2.3.1), quantified (section 2.3.2) and diluted in sterile water to 10 ng/ μ l.

The following primers were designed to recognize the HIV-1 Long Terminal Repeats integrated in the genome and to the human albumin (hALB) gene.

HIV LTR-FW: 5'-AGCTTGCCTTGAGTGCTTCAA-3'

HIV LTR-REV: 5'-AGGGTCTGAGGGATCTCTAGTTACC-3'

hALB-FW: 5'-GCTGTCATCTCTTGTGGGCTGT-3'

hALB-REV: 5'-ACTCATGGGAGCTGCTGGTTC-3'

As it is constitutively expressed in the human genome, hALB was used as a reference gene to normalize the qPCR results⁴¹⁰.

The pRRL.sin.cppt.PGK.GFP.Alb plasmid (pAlb), which contained an albumin sequence (acting as a human genome target) and an HIV-1 packaging sequence (serving as the lentiviral target), was used to generate two standard curves, that were run alongside the samples. Each standard curve, run in duplicate, consisted of serial 10-fold dilutions of pAlb, to generate a dilution series corresponding to plasmid copy numbers ranging from 10^7 copies/ μ l to 10^3 copies/ μ l. Dilutions were derived from a pAlb aliquot containing 10^{10} copies/ μ l, obtained by diluting 1 μ l of the original plasmid (2.9 μ g/ μ l) into 36.1 μ l of pure water.

qPCR was performed mixing 0.5 μ l of the 10 μ M forward and reverse oligos mix with 10 μ l of the genomic DNA derived from transduced cells, 12.5 μ l of the Power SYBRGreen PCR Master Mix (Cat n° 4368577) and water, to make a final volume of

25 µl. Negative controls were uninfected genomic DNA and water (blank). Each sample was run in triplicate.

qPCR was run at 10 minutes at 95°C, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. The number of viral genomes/ml in the starting volume of lentiviral particles was then determined by analysing the qPCR data, which were retrieved by the StepOne software and exported in an Excel file.

qPCR data analysis was based on a specific parameter also known as threshold cycle (Ct) value. Ct value corresponds to the amplification cycle in which, for both the DNA standards and samples, a fluorescent signal over the background is detected. The average Ct values derived from standards were plotted against the base 10 logarithm (log₁₀) of the copies of lentiviral particles for each standard.

For each standard curve an equation was generated, from which the slope and intercept values were derived.

The mean Ct values of each sample amplified with HIV-LTR and hALB primers were then put into the appropriate equation, subtracting the unknown value from the intercept value and dividing the product by the absolute number of the intercept. By using the POWER function, the log₁₀ values were converted back into standards, providing the separate titre value for HIV-LTR and hALB copies.

Once the number of LTR copies per hALB copies was determined, it was multiplied by two (to obtain the number of viral copies per host allele) and further multiplied by the number of infected cells to obtain the total number of viral genomes present in the analysed samples. The resulting product, further divided by the volume of transducing lentiviral particles (expressed in ml), was expressed as viral genomes/ml (vg/ml) and was representative of the number of viral genomes present in the starting volume.

2.2 CELL CULTURE

2.2.1. Ethics statement

The work carried out in this project involved the manipulation of human cells obtained from the MRC Centre for Neuromuscular Diseases BioBank. Ethical approval and consent for research have been obtained to facilitate pharmacological, gene and cell therapy trials in neuromuscular disorders (REC reference number 06/Q0406/33) and to allow the use of cells as a model system to study pathogenesis and therapeutic strategies for neuromuscular disorders (REC reference 13/LO/1826), in compliance with national guidelines regarding the use of human-derived cell lines for research. Written informed consent was obtained from each patient's guardian.

2.2.2. List of human cell lines

Each of the cell lines from the MRC Centre for Neuromuscular Diseases BioBank is indicated in the table below, together with information about the specific genomic mutation. Primary fibroblasts and myoblasts were derived from skin and muscle biopsies, respectively. Experiments were performed in primary cells which underwent a maximum of 10 passages. DUP2i cell line is the result of DUP2 myoblasts immortalization, which was carried out by Dr. Vincent Mouly⁴¹¹.

Cell line	ID	Dystrophin mutation
Immortalized embryonic kidney	HEK 293T	N.A. (wild-type dystrophin)
Primary fibroblasts	DYS+	N.A. (wild-type dystrophin)
Primary fibroblasts	DUP1	Duplication exons 5-11(out-of frame)
Primary myoblasts	DUP2	Duplication exons 3-16 (in-frame)
Immortalized DUP2 myoblasts	DUP2i	Duplication exons 3-16 (in-frame)
Induced pluripotent stem cells	DUP3	Duplication exons 3-41 (in-frame)

Table 2.1. Human cell lines.

2.2.3. Matrigel preparation and plate coating

Matrigel basement membrane matrix (BD 354234) was defrosted on ice and diluted to 1 mg/ml in chilled Dulbecco's modified Eagle's medium (DMEM). Gentamicin was added (final concentration of 50 µg/ml). Aliquots were prepared and stored at -20°C. Prior to their use, aliquots were defrosted overnight at 4°C. The volume of Matrigel needed to cover the surface was added to each plate, which was then placed at 37°C for 30-60 minutes. 500 µl, 3 ml and 5 ml Matrigel were used for one well of a 6-well plate, T75 Cornwell flask and T175 Cornwell flask, respectively. Excess Matrigel was then removed from the plate prior to seeding of cells.

2.2.4. Culture of human embryonic kidney (HEK 293T) cells

HEK 293T cells were seeded at a density of 30-40% (25000 cells/cm²) and passaged (section 2.2.8) when reaching a confluence of 70%. Culture medium was Dulbecco-MEM containing 2mM GlutaMax (Invitrogen 35050038) and 10% (v/v) foetal bovine serum (FBS) (PAA-A15151). Except when specifically required from the experimental conditions, 1% (v/v) of a solution containing penicillin (100 U/ml) and streptomycin (100 µg/ml) was also added. Cells were maintained at 37°C and 5% CO₂.

2.2.5. Culture of human fibroblasts

Primary fibroblasts were seeded at a density of 30-40% (10000 cells/cm²) and passaged (section 2.2.8) when reaching a confluence of 60%. Culture medium was prepared as for the section 2.2.4. Cells were maintained in an incubator set at 37°C and 5% CO₂.

2.2.6. Culture and differentiation of human myoblasts

Primary and immortalized myoblasts were seeded at a density of 20% (5000 cells/cm²) and passaged every other day (section 2.2.8) when reaching a confluence of 40%. Culture medium was complete Skeletal Muscle Growth medium (PromoCell, Cat n° C-23060) containing 3mM GlutaMax (Invitrogen), 10% (v/v) FBS and 40 µg/ml Gentamicin (Sigma, Cat n° G127210mg). Cells were maintained at 37°C and 5% CO₂.

Terminal differentiation of human myoblasts was achieved by culturing myoblasts with a confluence of 70% with the M2 differentiation medium. This was composed by DMEM (MegaCell, Cat n° M3942), 2% (v/v) FBS, 1X non-essential amino-acids, 2mM Glutamine, 0.5mM, β -mercaptoethanol and basic fibroblast growth factor (bFGF, also known as FGF2) (5ng/ml).

2.2.7. Culture and differentiation of human induced pluripotent stem cells (hiPSCs)

3x10⁴ cells/cm² hiPSCs were seeded onto 0.1mg/ml Matrigel-coated plates (section 2.2.3.) in mTeSR1 medium (Stem Cell Technologies, Cat n° 85870) supplemented with 10µM Y-27632 Rho kinase (ROCK)-inhibitor (Sigma, Y0503). Cells were passaged every other day (section 2.2.8.) and maintained at 37°C and 5% CO₂.

Myogenic differentiation of iPSCs requires their culture in well-defined media (whose composition is indicated in the **supplementary tables**) for a period varying from 24 to 31 days (**table 2.2**).

Myogenic differentiation							
	Day 0	Days 1-3	Days 4-6	Days 7-8	Days 9-12	Days 13-24	Days 25-28
Medium	mTeSR1	DICL	DICLF	DKHIFL	DKI	DKHI	Passage

Table 2.2. Media used for inducing myogenic differentiation in hiPSCs.

iPSCs cells seeded as indicated above represented the day 0 of the differentiation protocol. Cells were then incubated with DICL medium (**appendix I**) for 3 days (days 1-

3), which was changed every day. From day 4 to day 6, cells were cultured with DICLF medium (**appendix I**), while DKHIFL medium (**appendix I**) was used during day 7 and day 8 and replaced by DKl medium (**appendix I**) until day 12. Media used until day 12 were changed daily. From day 12 onwards, cells were cultured with DKHI medium (**appendix I**) for up to day 25-28. DKHI medium was changed every other day. Growing cells in each well of a six-well plate were passaged using 1 ml collagenase (Thermo Fisher Scientific) and 100 μ l dispase (Thermo Fisher Scientific). Pax7-positive satellite-like cells (myogenic progenitors) started to appear from day 21. These were identified by qPCR (section 2.3.13.a) and sorting of the ErbB3 myogenic marker (section 2.2.14.b).

2.2.8. Cell passaging

HEK 293T cells, primary fibroblasts and primary myoblasts were passaged by means of trypsin-ethylenediaminetetraacetic acid (EDTA). Medium was removed from the cells, which were then washed with sterile 1X PBS for 1 minute at room temperature to remove residual FBS. The volumes of PBS varied from 500 μ l to 5 ml depending on the culture plate's size: 500 μ l, 3 ml and 5 ml PBS were used for one well of a 6-well plate, T75 Corning flask and T175 Corning flask, respectively. PBS was then discarded and cells were incubated with 0.05% of trypsin-EDTA (Life Technologies Cat n° 25300062) for 10 minutes at 37°C and 5% CO₂ to allow the detachment of the cells from the culture surface. The trypsin-EDTA volumes were the same as the PBS volumes used in the previous step. The action of the trypsin was then stopped by the addition of growth medium containing 20% FBS. Detached cells were transferred to a Falcon tube and centrifuged for 5 minutes at room temperature, at 500 g.

When culturing iPSCs, trypsin-EDTA was replaced by TRYPLE™ Express reagent, a gentler reagent, so to avoid the disruption of iPSCs embryoid bodies.

2.2.9. Count of viable cells

Cell pellets (section 2.2.8) were resuspended in 1 ml growth medium and 10 μ l of cell suspension were then mixed with an equal volume of trypan blue dye (Sigma, Cat n° T6146). A 10 μ l aliquot was then added to the Countess™ chamber slide (Life Technologies). This was inserted in the Countess™ automated cell counter (Life Technologies), which returned the number of viable cells/ml.

2.2.10. Cryopreservation of cells

Cells were cryopreserved in freezing medium, prepared by adding 10 % (v/v) dimethyl sulphoxide (DMSO) (Sigma) to FBS. 0.5-1.0x10⁶ cells/ml in freezing medium were

transferred to cryovials (Dutscher, Cat n° 377224) and placed in the “Mr Frosty™” freezing container whose external chamber was filled with isopropanol. Cells were stored in Mr Frosty either at -80°C freezer (short-term storage, up to 1 week) or at -196°C in liquid nitrogen (long-term storage).

2.2.11. Cell transfection by using commercially available reagents

a) Lipofectamine transfection of HEK 293T cells

Lipofectamine 2000 transfection in HEK 293T cells was carried out to evaluate the efficiency of CRISPR/Cas9 nucleases. 1.50×10^5 HEK 293T cells were plated in each well of a 6-well plate with DMEM medium supplemented with 10 % FBS (section 2.2.4). Each plasmid expressing CRISPR/Cas9 was transfected in the cells using Lipofectamine 2000 (LF2000®), at a ratio of 1 µg DNA to 2 µl LF2000. Transfection efficiency was measured by including a parallel transfection of a GFP-expressing vector (pEGFP-C2) (**Fig.2.9**). After 24 hours, transfection medium was replaced with fresh culture medium. Three days post-treatment, cells were trypsinised (section 2.2.8), and centrifuged for 5 minutes at 500 g.

b) Lipofectamine transfection of primary myoblasts

The day prior to the transfection, 2×10^5 cells were seeded in each well of a six-well plate so as to be 70-80% confluent on the following day (10000 cells/cm²). Transfection complexes were prepared as follows: both Lipofectamine 2000 and DNA were added, separately, to each Eppendorf containing 250 µl of serum-free OptiMem (Thermo Fisher Scientific). Each complex was gently mixed and incubated for 5 minutes at room temperature. DNA-OptiMem mix was gently added to the diluted Lipofectamine 2000 and incubated at room temperature for 25-30 minutes. Meanwhile, cell culture medium (section 2.2.6.) was replaced by 2 ml of serum-free OptiMem. Assembled DNA-lipid complexes were then added dropwise to each well containing cells and incubated at 37°C and 5% CO₂ for 4 hours. Finally, transfection medium was removed and replaced by the Skeletal Muscle Growth medium (PromoCell). Cells were then incubated at 37°C and 5% CO₂. Transgene expression was evaluated after 48 hours.

2.5 µg DNA was used in combination with different amounts of Lipofectamine 2000 as suggested by the manufacturer's protocol, to identify the best experimental condition. Lipofectamine volumes tested were 6 µl, 9 µl, 12 µl and 15 µl.

c) TurboFect transfection of primary myoblasts

The day prior to the transfection, 2×10^5 cells were seeded in each well of a six-well plate, so as to be 70-90 % confluent on the following day. 2 µg DNA was diluted to 400

µl in serum-free OptiMem and mixed thoroughly by vortexing for 30 seconds. TurboFect was then added to the diluted DNA and well mixed by vortexing. Different volumes of transfection reagent (4 µl, 6 µl and 8 µl) were tested. Each complex was then incubated at room-temperature for 15-20 minutes and added dropwise to each well. The plate was gently rocked to distribute the mixture evenly and was finally incubated overnight at 37°C and 5% CO₂. Transgene expression was evaluated after 48 hours.

d) GeneJuice transfection of primary myoblasts

2 x 10⁵ cells were seeded in each well of a six-well plate the day before transfection. 6µl of GeneJuice reagent were added to 100 µl of serum-free OptiMem medium, thoroughly mixed by vortexing and incubated for 5 minutes at room temperature. 2 µg DNA was added to the mix and gently mixed by pipetting. Following a second incubation at room temperature for 15 minutes, the mix was then added dropwise to the plate, which was rocked gently to allow even distribution of the transfection mix and incubated overnight at 37°C and 5% CO₂. Transgene expression was evaluated after 48 hours.

2.2.12. Cell transfection by using nuclear electroporation

a) Nuclear electroporation of human fibroblasts and myoblasts - Amaxa®

DUP1 fibroblasts and DUP2 myoblasts (**table 2.1** in section 2.2.2), were electroporated according to the protocol provided by Amaxa (cells:DNA ratio of 1.5 x 10⁶:2.5 µg). Following trypsinisation, cells were harvested and centrifuged for 5 min at 500 g in a table-top centrifuge at room temperature. The supernatant was discarded and the pellet was resuspended in 100 µl of the solution V provided with the nucleofection kit. 3 µg of CRISPR/Cas9-GFP plasmid (**Fig.2.1.2.h**) were then added and the mix was transferred to the provided electroporation cuvette, which was immediately placed into the Amaxa device, where the electroporation was performed. The electroporation program P022 was chosen, as preliminary tests performed in the laboratory on human fibroblasts showed it to result in the highest transfection efficiency (about 30%), whilst maintaining high cell viability. Upon completion of the electroporation, the cuvette was released from the machine and its content was transferred to a 10 cm² plate containing pre-warmed culture medium, by using the small plastic Pasteur pipette provided with the kit. The plate was then incubated overnight at 37°C and 5% CO₂.

b) Nuclear electroporation of human myoblasts - NEON®

DUP2 and DUP2i myoblasts carrying the in-frame duplication of exons 3-16 were electroporated with the NEON device. 5 x 10⁵ cells were resuspended in 9 µl of buffer R provided with the kit. 1 µg of the highly pure CRISPR/Cas9-GFP plasmid (**Fig.2.1.2.h**)

was added to the suspension. Specific NEON tips containing a gold-plated electrode were used to transfer the 10 µl mix to the NEON electroporation tube. A second electrode placed at the bottom of the electroporation tube allows the transfer of the electric field from the electrode located in the tip. The electroporation tube was filled with 3 ml of the electrolytic buffer E and placed into the electroporation station which generates the electric pulse. Number of pulses, duration and voltage intensity are defined for each cell type. Electroporation parameters were as follows: 1 pulse, 20 milliseconds, 1400 V. Electroporation tip was finally removed from the electroporator and its content was released in a well of a 12-well plate filled with 1.5 ml pre-warmed culture medium.

c) Nuclear electroporation of human induced pluripotent stem cells (hiPSCs) - NEON®

DUP3 iPSCs (**table 2.1** in section 2.2.2) were electroporated with the NEON device as detailed in section 2.2.12.b. 5×10^5 cells were prepared for the electroporation step (2.2.12.b.). Three different electroporation parameters were tested:

-1 pulse, 20 milliseconds, 1400 V;

-2 pulses, 30 milliseconds, 1050 V;

-1 pulse, 30 milliseconds, 1100V.

Following electroporation, cells were seeded in each well of a 6-well Matrigel-coated plate which was pre-filled with 2 ml of pre-warmed iPSCs culture medium without antibiotic and supplemented with the Rho kinase (ROCK) inhibitor (Sigma, Cat n°Y0503) to increase the viability of the cells⁴¹².

2.2.13. Cell transduction

a) Lentiviral transduction

Cells to be transduced with integrating lentiviral vectors were seeded at 70-80% confluence (25000 cells/cm²). The following day, different volumes of lentiviral particles were added to the culture medium, rocking the plate to ensure homogeneous distribution. Cells were then incubated overnight at 37°C and 5% CO₂. The following day, the medium containing the lentiviral particles was discarded, cells were washed using PBS and cultured in growth medium for 48 hours, following which the transgene expression was assayed.

b) Adenoviral transduction

Cells to be transduced with adenoviral vectors expressing MyoD were seeded to be 70-80% confluent on the following day (25000 cells/cm²). Different volumes of adenoviral particles at 50, 100 or 200 MOI were added to the culture medium, rocking the plate to ensure homogeneous distribution. Cells were then incubated at 37 °C and 5% CO₂ and, after 4 hours, the medium containing the adenoviral particles was replaced by growth medium. Cells were incubated at 37°C and 5% CO₂ for another 48 hours, following which the transgene expression was assayed.

MyoD-expressing fibroblasts were finally seeded to be 70% confluent (50000 cells/cm²) and cultured for 7-9 days with high-glucose Dulbecco-MEM, 1% penicillin (100 U/ml) and streptomycin (100 µg/ml) solution and 2% (v/v) horse serum to induce terminal differentiation. Two other differentiation media were tested. The former was composed by high-glucose Dulbecco-MEM, 1% penicillin (100 U/ml) and streptomycin (100 µg/ml) solution without any serum, while the latter included high-glucose Dulbecco-MEM, 1% penicillin (100 U/ml) and streptomycin (100 µg/ml) solution and 1% (v/v) FBS.

2.2.14. Fluorescence Activated Cell Sorting (FACS)

a) GFP sorting

The pool of cells containing GFP-positive cells were washed with PBS and trypsinised (section 2.2.8). Cells were centrifuged at 500 g for 5 minutes at room temperature, washed with 500 µl of PBS and the pellet resuspended in 300 µl PBS and transferred into the appropriate Falcon® round-bottom polystyrene FACS tubes (Thermo Scientific). Samples were then placed on ice and brought to the FACS facility (<https://www.ucl.ac.uk/child-health/core-scientific-facilities-centres/flow-cytometry-core-facility>), where trained personnel performed the FACS by either FACSaria III, FACSCalibur or MoFlow XDP Cell sorter. Sorted cells were then incubated overnight at 37°C and 5% CO₂.

b) Sorting of iPSC-derived myoblast-like cells expressing ErbB3

Muscle precursor cells were adjusted to a density of 0.5-1x10⁶ in 100 µl of growth medium. 1 µl of ErbB3-APC antibody (Cat. n° FAB3481A) was added to the cell suspension and incubated for 30-60 minutes in the dark at 4°C. 900 µl of PBS were then added to the mix, which was centrifuged for 3 minutes at 300 g at room temperature. The resulting pellet was re-suspended in 0.5 ml of fresh growth medium and passed through a 40 µm cell strainer (Corning®), prior to FACS sorting. Non-stained cells were used as a negative control to set the sorting gate.

2.3. GENERAL MOLECULAR BIOLOGY PROCEDURES

2.3.1. Extraction of genomic DNA from human cell lines

Genomic DNA was extracted HEK 293T cells as well as from fibroblasts and myoblast cell lines derived both from healthy individuals and DMD patients by using the QIAgen DNeasy Blood & Tissue Kit. As soon as reached 80-90% confluence, cells seeded with a density of 10000 cells/cm² in one well of a 6-well plate were passaged (section 2.2.8), transferred into a sterile 1.5 ml Eppendorf tube and centrifuged for 5 minutes at 500 g at room temperature. The medium was then removed and the pellet was resuspended in 200 µl PBS and 20 µl proteinase K were added to digest contaminant proteins. 200 µl AL buffer were then added and the tube was vortexed for 30 seconds to ensure homogenous mixing and incubated at 56 °C for 10 minutes. 200 µl of 100 % ethanol was added and vortexed as above. The sample was then loaded onto a DNAeasy Mini spin column placed within a 2-ml collection tube and centrifuged for 1 minute at room temperature to allow DNA binding to the column. The filter was discarded and the collection tube was changed. Two centrifugation steps at 6800 g for 1 minute and 20000 g for 3 minutes were performed after the addition of 500 µl of AW1 and AW2 buffer, respectively. A new collection tube was used for each step. Finally, 200 µl of AE elution buffer was pipetted onto the column transferred to a 1.5 ml collection tube and incubated for 1 minute at room temperature. DNA was eluted by a centrifugation step of 1 minute at 6800 g.

2.3.2. DNA and RNA quantification

Both plasmid/genomic DNA and total RNA were quantified by means of the NanoDrop1000 Spectrophotometer (Thermo Scientific), according to the manufacturer's instructions.

2.3.3. Primers design

Primers were designed by using the primer3 on-line tool (primer3.ut.ee/) and synthesized by Sigma-Aldrich. All primers were provided as desalted stocks which were diluted in pure water to a concentration of 100 µM and stored at -20°C. When needed, they were further diluted in sterile water to a concentration of 10 µM.

2.3.4. Touchdown PCR

The genomic region surrounding the CRISPR/cas9 target sites was amplified with the high fidelity Q5 DNA polymerase. This was accomplished by means of touchdown PCR (tdPCR), a technique based on stepwise decrements in annealing temperature to

improve the specificity as well as yield of PCR reaction, reducing the production of multiple amplicons⁴¹³. Together with the primer melting temperature (T_m), the choice of the annealing temperature is a critical factor in regards to PCR success, as it could shift a faithful target amplification towards the generation of multiple misprimed products; in fact, annealing temperatures consistently lower or higher than T_m increase the chance of having nonspecific products impeding the amplification of desired DNA regions⁴¹⁴. The annealing temperature used by the tdPCR assumes values higher than the estimated T_m , which are progressively decreased along with the amplification cycles. The chance of nonspecific binding is therefore contained because of the initially higher annealing temperature, and further reduced two-folds per cycle with the reaching of more permissive annealing temperature values.

Primer couples specific for each CRISPR/Cas9 target site were named as LentiCRISPR (LC) 1-4 and used at 10 μ M:

LC1 forward primer = 5'-GACCTCAGCAAAGTCCCTCT-3'

LC1 reverse primer = 5'-ACCTCACCCACCCATCTAAA-3'

LC2 forward primer = 5'-TCAATGGCAGGTGATATCGC-3'

LC2 reverse primer = 5'-GGTTGTTCCGTATAGTTGGCC-3'

LC3 forward primer = 5'-TGCTAGATCAAGTGGGAGTTCT-3'

LC3 reverse primer = 5'-ACTCAGAGCTATTAGGAGGGT-3'

LC 4 forward primer = 5'TCCCTCCGTTGAAATCACCA-3'

LC 4 reverse primer = 5'-CTGTAGGGGCAGTGAAAGCT-3

The expected molecular size for each amplicon is indicated below:

LentiCRISPR1 amplicon = 710 bp

LentiCRISPR2 amplicon = 450 bp

LentiCRISPR3 amplicon = 630 bp

LentiCRISPR4 amplicon = 565 bp

TdPCR is generally carried out with two cycling phases: the first phase involves a starting cycle that begins with an annealing temperature of $T_m + 10^\circ\text{C}$ followed by 10-15 cycles in which the annealing temperature is progressively decreased to an

annealing temperature of 55°C. The second phase is 20-25 amplification cycles set at the 55°C annealing temperature.

The tdPCR started with 2 minutes at 98°C to allow DNA denaturation followed by 26 amplification cycles, each of 10 seconds at 98°C, 30 seconds at 68°C (-0.5°C per cycle) and 30 seconds at 72°C. Another 9 amplification cycles, each of 10 seconds at 98°C, 30 seconds at 56°C and 30 seconds at 72°C were then added to complete the protocol.

5 µl of the resulting amplicons were run on a 1% agarose gel (section 2.3.5) which was then exposed to the UV light of a transilluminator (UVP Bio-Imaging System). If only the expected amplicon was detected, DNA fragments were then purified with QIAquick PCR purification Kit (QIAGEN) (section 2.3.6.a.). If instead multiple non-specific products resulted from the tdPCR, the QIAquick Gel Extraction Kit's protocol (QIAGEN) was followed (section 2.3.6.b.).

2.3.5. Preparation of agarose gels and gel electrophoresis

Agarose gels were made by mixing TBE buffer with UltraPure™ agarose powder (Thermo Fisher Scientific). Different pore sizes of the agarose network were created depending on the molecular size to be visualized. Generally, 1% agarose gels were used; these were prepared adding 0.5 g or 1 g agarose to 50 ml or 100 ml TBE buffer, respectively. In some instances, 2% agarose gels were used. TBE and agarose powder mixture was boiled in a microwave and SYBR® Safe gel staining reagent (Invitrogen) was added in a ratio 1:10000 before it had cooled. Upon stirring, the liquid compound was poured into the gel cassette containing the appropriate gels comb (BioRad) and left at room temperature until the gel had solidified, at which point was placed in the electrophoresis apparatus filled with TBE 1X. Gels were run at constant voltage (110 V) for 45 minutes.

2.3.6. DNA purification

a) DNA purification from PCR

PCR purification was performed as detailed in the QIAquick PCR Purification Kit (QIAGEN). 5 volumes of PB buffer per volume of PCR reaction were mixed with the PCR samples. As PB buffer was supplemented with pH indicator (at a ratio of 1:250), a change in colour from yellow to orange or violet indicative of altered pH (higher than the optimal pH 7.5) was corrected by adding 10 µl of 3 M sodium acetate at pH 5. The mixture was loaded onto a QIAquick column (provided with the kit) already placed in a sterile collection tube and underwent the experimental steps described in section 2.3.6.b.

b) DNA purification from agarose gel

Gel extraction was performed according to the QIAquick Gel Extraction Kit protocol (QIAGEN). The band to be excised from the agarose gel was first visualized with a blue-light transilluminator (UVP Bio-Imaging System) and excised from the gel using a clean scalpel blade (Fisherbrand™ SD1270100). The gel fragment containing the DNA of interest was transferred to a clean 1.5 ml or 2 ml Eppendorf tube and weighed on an analytical balance (Sartorius). Three volumes of QG buffer were added to 1 volume of gel (assuming that 100 mg is approximately equal to 100 µl). The mixture was incubated at 50°C for 10 minutes during which it was repeatedly vortexed every 3 minutes to facilitate this solubilizing step. Buffer QG confers a yellow colour to mixture, which is indicative of a pH less than or equal to 7.5 (necessary for DNA binding to the silica column in the next step). If the melted product turned orange or violet (indicating pH higher than 7.5), 10 µl of 3 M sodium acetate was added to the mixture which was mixed and/or further incubated at 50°C until its colour returned to yellow. Before loading onto the silica column, 1 volume of isopropanol was added to improve the yield of DNA whose size was outside the range of 500-4000 bp. The preparation was then transferred to a QIAquick spin column placed within a 2-ml collection tube and centrifuged for 1 minute at 17900 g in a table-top microcentrifuge set up at room temperature. Speed, time and temperature were kept constant for all the centrifugation steps. The flow-through was discarded and the column was washed with 750 µl of active PE buffer. Two separate centrifugation steps were then needed to remove the flow-through and the excess buffer which, containing ethanol, may affect the purity of the eluate. The column was then transferred to a sterile 1.5 ml Eppendorf collection tube and 30-50 µl of sterile water (pH 7.5) were added at the centre of the column. Following a 1-minute incubation at room temperature, the column was centrifuged for 1 further minute allowing DNA elution. DNA concentration was finally assessed via the NanoDrop spectrophotometer (section 2.3.2).

2.3.7. T7 nuclease assay

The ability of vectors expressing CRISPR/Cas9 to cut genomic DNA was evaluated by means of the T7 endonuclease (T7E1) mismatch detection assay, one of the available mismatch-based techniques⁴¹⁵ adopted to detect indels that generally derive upon NHEJ-mediated repair⁴¹⁶. The assay exploits specific nucleases able to cleave DNA heteroduplexes at the mismatch site originated when, following repeated denaturation/renaturation cycles, wild-type DNA pairs to DNA carrying indels (as for CRISPR/cas9 targets) (**Fig.2.13**).

To perform the T7 endonuclease assay, 200 ng of purified tdPCR product was mixed

with water to a final volume of 17.1 µl and 1.9 µl of 10X buffer 2 (NEB). The mixture was then heated to 95°C and gradually cooled down in thermocycler to 25°C to allow denaturation and re-annealing of the amplicons. Pre-T7 annealing was performed as follows: 10 minutes at 95°C (ramping 50%), 1 minute at 85°C (ramping 8%), 1 minute each 75°C, 65°C, 55°C, 45°C, 35°C (ramping 8%) and 25°C.

1 µl of T7 endonuclease was then added to each sample and incubated for 15 minutes at 37°C. The reaction was stopped by the addition of 2 µl of 0.25M EDTA and run on a Novex® TBE gel as detailed below (section 2.3.8).

2.3.8. TBE gel electrophoresis

5 µl of each T7 reaction product were mixed with water up to a final volume of 8 µl and 2 µl TBE Hi-Density Sample Buffer (5X) and resolved on a high-sensitivity 4-20% Novex® TBE gel (Life Technologies) at 200 V for 45-60 minutes. The TBE gel was then removed from its cassette and stained for 15 minutes with SybrGold (Life Technologies) diluted 10000X in 200 ml TBE. SybrGold staining was performed at room temperature on an orbital shaker in the dark. The gel was visualized with the Gel Doc XR+ system which allows acquisition of high-resolution images.

2.3.9. Determination of CRISPR/Cas9 cleavage efficiency

T7 assays run on amplicons derived from CRISPR/Cas9 targeted genomic sites result in multiple bands: the fragment with the higher molecular weight corresponds to the full-length non-cleaved product, while the two smaller fragments refer to the 'arms' on each side of the cleaved CRISPR/Cas9 target sequence. The indel frequency at the cleavage site correlates with LentiCRISPR efficiency. T7 (and consequently LentiCRISPR) efficiency was determined by visualizing the enzyme-cut versus uncut fragments when reaction products are run on a gel.

The efficiency of CRISPR/Cas9 genomic cleavage efficiency was quantified by exporting in Excel format the data from the densitometric band analysis, which was done by means of the Fiji software. The following equation was used to calculate the percentage of cleavage for each CRISPR/Cas9 vector:

CRISPR/Cas9 cleavage efficiency (%) = Upper cleaved band/ (Upper cleaved band + Full-length band area value) x 100

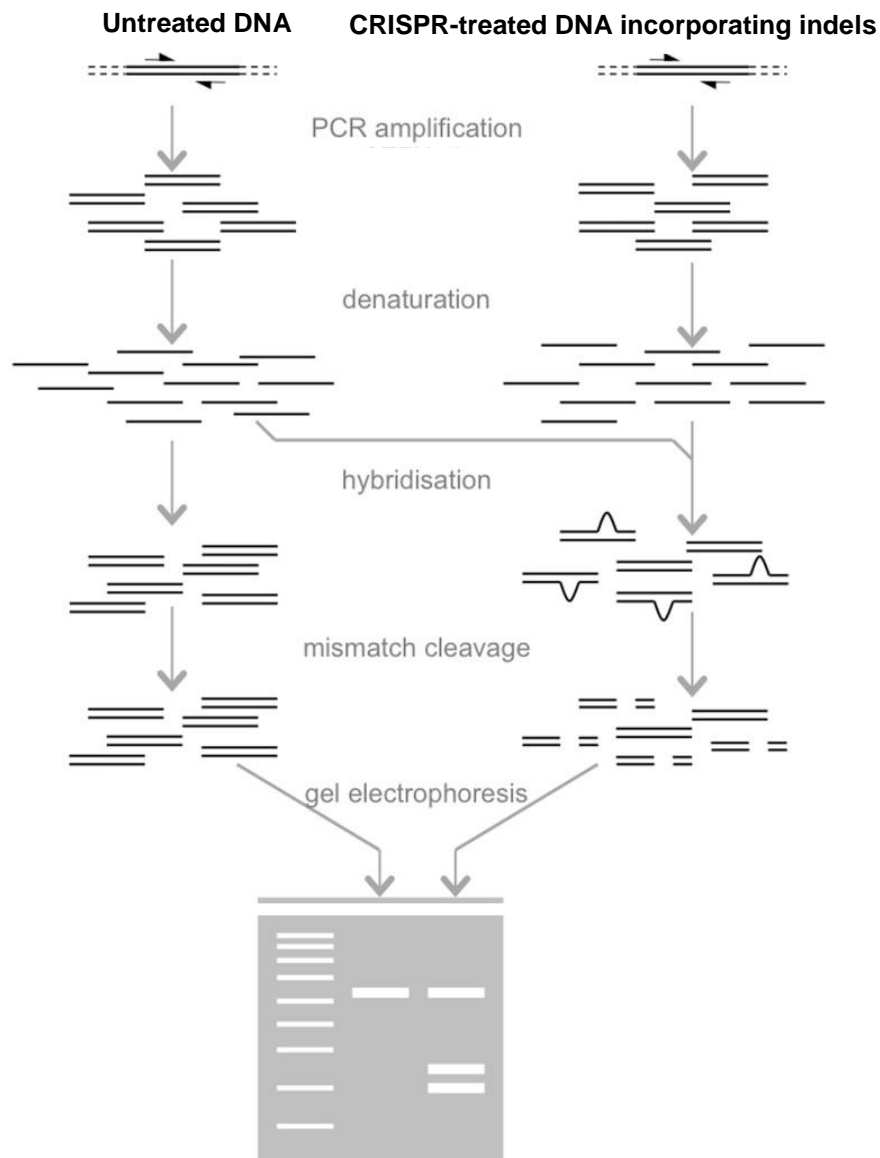


Figure 2.13. Schematic of mismatch nuclease assay to determine the efficiency of CRISPR cleavage.

A tdPCR reaction was performed on genomic DNA using primers flanking the CRISPR target site in intron 9. Amplicons were denatured, re-annealed and digested with the base pair mismatch-sensitive T7 endonuclease, which cleaves duplexes at the site of any indels. CRISPR cleavage efficiency is then assessed by visualization of mismatch enzyme-cut versus -uncut fragments on agarose gel.

2.3.10. Extraction of RNA from human cell lines

RNA was extracted from patient-derived cells according to the protocol provided with the RNeasy Mini kit from QIAgen. Cells grown in a 6-well plate (table section 2.2.2) were used for RNA extraction when they reached 70% confluence. Culture medium was aspirated and 350 µl of the RLT lysis buffer supplemented with β-mercaptoethanol was added to the cells. 10 µl of β-mercaptoethanol was added per ml of RLT buffer. Following a 5 minute incubation at room temperature, 350 µl of 70% ethanol was added to the lysate to ease RNA binding to the silica membrane. The mixture was homogenized by vigorous pipetting and loaded onto a MiniElute Spin column which was centrifuged for 1 minute at 8000 g at room temperature. The flow-through was discarded and the column was washed with 700 µl of RW1 buffer. Following a further centrifugation as described above, the column was placed in a new collection tube and washed twice with 500 µl of RPE buffer. The filtrate was discarded and the column was centrifuged at 18000 g for 1 minute at room temperature to remove any traces of ethanol. Finally, the column was transferred to a sterile RNase-free 1.5 ml Eppendorf tube and 30 µl of RNase-free water was placed onto the centre of the column which was left at room temperature for 1 minute. RNA elution occurred following one last centrifugation step at 10000 g for 1 minute. Eluted RNA was then immediately placed on ice and its yield and purity was assessed by using the NanoDrop spectrophotometer (section 2.3.2).

2.3.11. Retro-transcription

a) Retro-transcription (SuperScript™ VILO™ cDNA Synthesis kit)

The eluted RNA was purified from contaminating DNA. 1 µg of RNA was diluted to a final volume of 8 µl with water, mixed with 1 µl of 10X DNase Buffer and 1 µl of DNase (Sigma) and incubated at room temperature for 15 minutes, before termination of the reaction with 1 µl of 1.5M Stop solution and incubation at 70°C for 10 minutes. To generate cDNA, 10 µl of the reaction product was then mixed with 4 µl of 5X Vilo (ThermoFisher Scientific) reaction mix and 6 µl of ultrapure water to set up the reverse-transcription reaction. This involved the incubation of the mix at 25°C for 10 minutes, at 42°C for 60 minutes and at 85°C for 5 minutes.

b) Retro-transcription (High-Capacity RNA-to-cDNA™ Kit)

cDNA to be analysed by qPCR was obtained by the extracted RNA by means of the High-Capacity RNA-to-cDNA™ Kit (Thermo Fisher Scientific, Cat. n° 4387406). The retro-transcription protocol was set up by mixing 500 ng – 1 µg RNA (made up to 9 µl with RNase-free water) with 10 µl of the 2X RT Buffer mix, 1 µl of the 20X RT Enzyme

mix and water up to a final volume of 20 µl. The mix was briefly centrifuged and incubated at 37°C for 30 minutes. The reaction was then stopped by increasing the temperature to 95°C for 5 minutes and the resulting cDNA stored at -20°C until used for qPCR.

2.3.12. Reverse-transcriptase (RT)-PCR

5 µl of cDNA obtained by the retro-transcription Vilo kit was used to set up a 50 µl RT-PCR reaction. Primers, whose sequence is indicated below, were designed to target exon 1 and exon 20 and thus amplify the entire duplicated dystrophin region. Q5 high-fidelity polymerase was used to amplify targets using an elongation time of 40 seconds per 1kb of amplicon. RT-PCR extension was performed at 72°C with an extension time of both 2 and 4 minutes, so that both restored and mutated dystrophin transcripts could be amplified.

DysEx1-Forward primer: 5' GAGGACACATTGCAAGCACA 3'

DysEx20-Reverse primer: 5' TAGTGATGGCTGAGTGGTGG 3'

The reaction was run at 98°C for 30 seconds and then for 35 cycles at 98°C for 10 seconds, 65°C for 20 seconds, and 72°C for 2 minutes (or 4 minutes depending on the transcript to be amplified). The final extension was carried out at 72°C for 2 minutes.

2.3.13. Quantitative PCR

a) One-step qPCR

One-step qPCR was set up starting from RNA derived from hiPSCs induced towards myogenic differentiation. For each well of the qPCR plate, the reaction Master Mix (final volume 6.25 µl) was made by combining 6 µl of SYBRGreen reaction mix (BioRad) with 0.15 µl iTaq RT enzyme and 0.5 µl of each primer (whose sequences are indicated below) at a concentration of 10 µM each. 5.25 µl of RNA was added to each well. Different master mixes were prepared for each of the following primer pairs:

Pax6-Forward primer = 5'-AGTGAATCAGCTCGGTGGTGTCTT-3'

Pax6-Reverse primer = 5'-TGCAGAATTCGGGAAATGTGCAC-3'

Pax7-Forward primer = 5'-CCCCCGCACGGGATT-3'

Pax7-Reverse primer = 5'-TATCTTGTGGCGGATGTGGTTA-3'

Myogenin-Forward primer = 5'-TACCAGGAACCCCGCTTCTA-3'

Myogenin-Reverse primer = 5'-CTCGTAGCCTGGTGGTTCG-3'

Myosin heavy chain-Forward primer = 5'-AGTCATGGCGGATCGAGAGA-3'

Myosin heavy chain-Reverse primer = 5'-CAGTCACCGTCTTGCATTCT-3'

SDHA-Forward primer = 5'-TGGGAACAAGAGGGCATCTG-3'

SDHA-Reverse primer = 5'-CCACCACTGCATCAAATTCATG- 3'

The qPCR was set up as shown in **Fig.2.14** and run at 95°C for 3 minutes and followed by 40 cycles of 10 seconds at 95°C and 1 minute at 60°C.

b) Two-step qPCR

Two-step qPCR was set up starting from RNA derived by CRISPR/Cas9-treated myoblasts induced towards terminal differentiation (section 2.2.6). For each well of the qPCR plate, the reaction master mix was set up by mixing 10 µl of the Takyon qPCR ROX SYBR 2X MasterMix dTTP blue (Cat n° UF-RSMT-B0701), 2 µl of each 10 µM primer (forward and reverse) and 3.5 µl of ultrapure water in a final volume of 17.5 µl.

Primer pairs were:

DysEX8-9-Forward primer = 5' TTGCCAAGGCCACCTAAA 3'

DysEX8-8- Reverse primer = 5' TCTCTCATATCCCTGTGCTAGA 3'

DysEX20-Forward primer =5' TGGATCGAATTCTGCCAGTT 3'

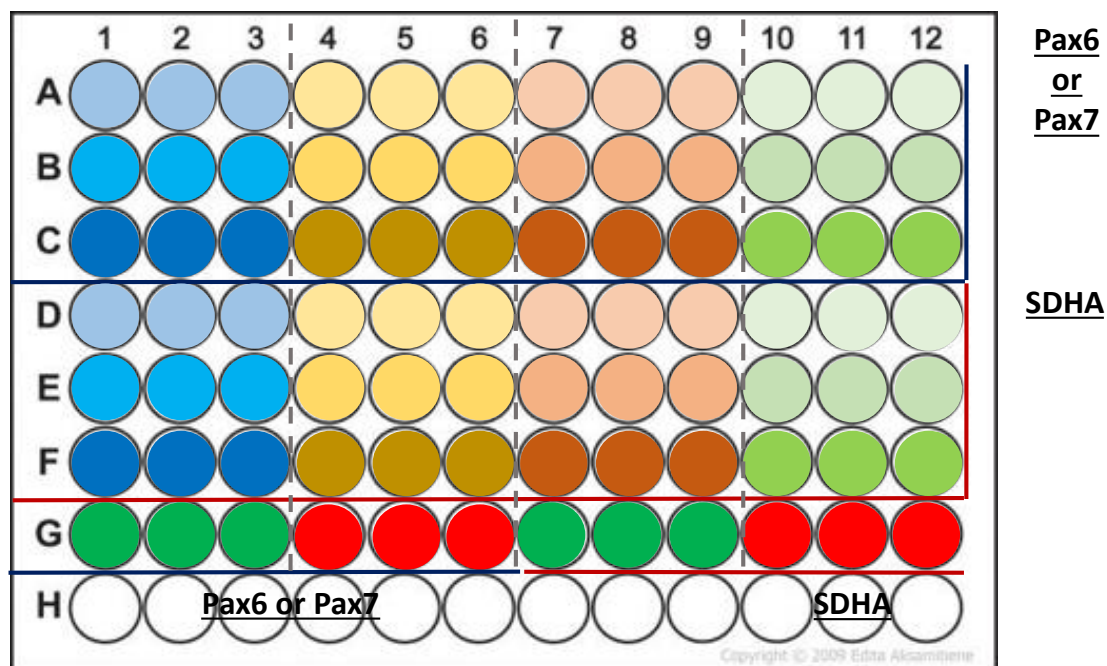
DysEX20-Reverse primer = 5' GCTCCAATTGTTGTAGCTGATTAT 3'

GAPDH-Forward primer = 5' TTGAGGTCAATGAAGGGGTC 3'

GAPDH-Reverse primer = 5' GAAGGTGAAGGTCGGAGTCA 3'

For the generation of a standard curve, cDNA was diluted 1:1, 1:2, 1:4, 1:8, 1:16, 1:32 in water, whereas if the comparative $\Delta\Delta C_t$ method was used (section 2.3.14), cDNA was diluted so that the resulting amount present in each well was 10 ng.

The qPCR was then run as follows: 95°C for 3 minutes, and 40 cycles of 10 seconds at 95°C and 1 minute at 60°C.



Legend:

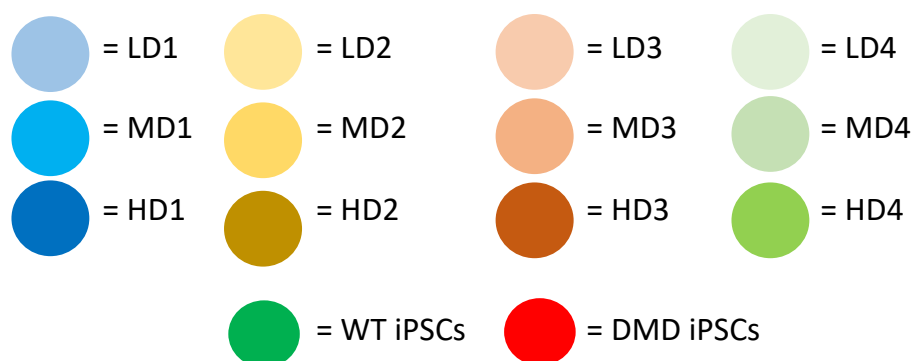


Figure 2.14. Quantification of neuronal and myogenic markers in iPSCs induced towards myogenic differentiation: qPCR plate set-up.

RNA extracted from each replicate of DUP3 iPSCs induced to differentiate into the myogenic lineage was used to set-up a qPCR plate designed to monitor the level of expression of either Pax6 (neuronal specification) or Pax7 (myogenic cells). Gene expression data were normalized against the constitutive succinate dehydrogenase (SDHA) gene expression. LD1-4, MD1-4, HD1-4 represented each replicate of RNA extracted from cells seeded at low, medium and high density, respectively. WT-iPSCs and DMD-iPSCs represented two control cell lines derived from a wild-type and DMD individual, respectively, and maintained at the iPSC stage. Different qPCRs were performed for each set of cells treated with either 3 μ M or 6 μ M of CHIR99021.

2.3.14. $\Delta\Delta\text{Ct}$ comparative qPCR analysis

qPCR data were retrieved from the StepOne device and exported in Excel format for data analysis. First, the mean Ct value of untreated samples was obtained for each target gene. The Ct values of each untreated and treated samples were then divided by the mean Ct value (generating the ΔCt value). This step was done for each of the target genes. $\Delta\Delta\text{Ct}$ was obtained by dividing the ΔCt of the gene of interest and the ΔCt of the normalizer gene. The relative expression fold changes were determined by calculating the $2^{-\Delta\Delta\text{Ct}}$ values.

2.3.15. Cell fixation and permeabilization

Cells fixation was done by removing the cell culture medium and incubating the cells with pre-warmed 4% paraformaldehyde (previously prepared dissolving 40 g paraformaldehyde powder in 800 mL heated PBS solution 1X and stored at -20°C) for 15 minutes at room temperature. 4% paraformaldehyde was then removed and appropriately discarded in a bottle into the fume hood, which was then sent to chemical waste. Cells were washed for 3 minute intervals with PBS, at room temperature. PBS was removed and cells were permeabilized by incubation at room temperature for 5 minutes with 0.5 % Triton-X.

2.3.16. Myogenicity assay

4×10^5 myoblasts were seeded in each well of a 8-well chamber-slide previously coated with 1 mg/ml Matrigel (section 2.2.3) in a final volume of 250 μl of Skeletal Muscle Growth Medium (PromoCell)/well. Cells were incubated overnight at 37°C and 5% CO_2 . The following day, culture medium was removed and replaced by 250 μl of M2 differentiation medium (section 2.2.6). At day 7, the medium was removed and cells were fixed and permeabilized (section 2.3.15). Triton-X was then removed and cells washed 2 times with PBS as above. PBS was replaced with 250 μl 10% goat serum, and incubated for 30 minutes at room temperature to block non-specific sites. Desmin immunostaining was then performed as indicated in section 2.3.17.a. Chambers were removed from the slide and carefully placed on top of a 25 mm coverslip (FisherbrandTM), on which a line of fluoromount (Dako mounting medium, Cat. N° S3023) was previously added, avoiding the formation of air bubbles. Myogenicity was assessed by microscope analysis of the slide by measuring the ratio between the total number of desmin-positive cells with more than 3 nuclei, and the total number of DAPI stained nuclei. This value, when multiplied by 100, represents the fusion index⁴¹⁷.

2.3.17. Cell immunostaining

a) Desmin immunostaining

Desmin primary antibody (Dako, mouse anti-human Cat. n° M0760) was diluted 1:100 in 10% goat serum and PBS and 200 µl of antibody was added to each well and incubated for 60 minutes at room temperature. DAPI (Thermo Fisher Scientific) was also added (1:10000) and incubated together with desmin antibody to stain the nuclei in each cell. After 3 PBS washes of 10 minutes each, at room temperature, cells were incubated with 200 µl of secondary antibody/well (Invitrogen, goat anti-mouse conjugated with Alexa488, Cat. n° A11029) prepared in PBS (1:100 dilution). Chamber-slides were wrapped in foil during this step to avoid photobleaching of the fluorochrome.

b) MyoD immunostaining

MyoD immunostaining was performed on fixed and permeabilized cells (section 2.3.15) by using the Anti-MyoD polyclonal primary antibody (SantaCruz, 1:50 dilution). Cells were washed with PBS (3 x 10 minutes, at room temperature) and were incubated in the dark with the goat anti-rabbit secondary antibody conjugated with Alexa-594 (Invitrogen, Cat. n° A-11012), diluted 1:200 in PBS.

2.3.18. Phalloidin staining

Primary human fibroblasts were seeded in each well of a 6-well plate and after 24 hours they were fixed and permeabilized as indicated in section 2.3.15. Following 1 further PBS wash of 10 minutes, cells were incubated for 90 minutes at room temperature with phalloidin-iFluor conjugated with Alexa-594 (Abcam, Cat. n° 176757) and washed again with PBS (three times, 10 minutes each).

2.3.19. Protein extraction

Proteins were extracted from differentiated patient-derived myotubes obtained upon culturing 80% confluent myoblasts (25000 cells/cm²) with differentiation medium (section 2.2.6.) for 9 days. Culture medium was removed and replaced with 200 µl of NHC lysis buffer (4 M urea, 125 mM Tris pH 6.8 and 4% SDS) supplemented with 1X Complete Protease Inhibitor Cocktail Tablets (Roche). If higher amounts of protein were needed, cells grown either in a 10 cm dish or in a T75 flask (Corning®) were trypsinised (section 2.2.8), resuspended in the above complete lysis buffer at a ratio of 1:1 and left on ice for 10 minutes. Samples were collected, boiled for 3 minutes and centrifuged at 14,000g for 10 minutes at 4°C. Each supernatant was collected and stored at -80°C until further processing.

2.3.20. Protein quantification

Protein quantification was carried out according to the Pierce® BCA protein assay (ThermoScientific Cat. n° 23250) developed for compatibility with the NHC reducing agent.

A standard curve with serial dilutions of the bovine serum albumin (BSA) protein was set up as follows:

Tubes	Diluent (NHC)	BSA (2000 µg/ µl) Stored at room temperature	Concentration (µg/ µl)
A	0 µl	25 µl stock	2000
B	15 µl	75 µl stock	1500
C	50 µl	50 µl stock	1000
D	50 µl	50 µl tube B	750
E	50 µl	50 µl tube C	500
F	50 µl	50 µl tube E	250
G	50 µl	50 µl tube F	125
H	75 µl	50 µl tube G	25
I	50 µl	0 µl	0

Each sample of the curve was vortexed and centrifuged down before preparing the following tube.

25 µl of each was then transferred to a new Eppendorf tube labelled, as above, from A to I. Meanwhile for each protein sample to be analyzed, triplicates were prepared by mixing 1.25 µl of protein with 23.25 µl of NHC lysis buffer, up to a total of 25 µl. 25 µl of a buffer named reconstitution buffer, previously diluted 1:1 in ultrapure water and vortexed for 30 seconds, were then added to each of the newly labelled tubes of the standard curve samples and each triplicate sample, up to a total volume of 50 µl. All Eppendorf tubes were vortexed, briefly centrifuged at 8000 g at room temperature and incubated at 37°C for 15 minutes. Before the end of the incubation, reagent A and B provided with the kit were mixed at a ratio of 1 of A and 50 of B and vortexed briefly. 1 ml of the mixture was added to each protein standard and sample and incubated at 37°C for 30 minutes and cooled for 5 minutes at room temperature before measuring the absorbance values. Protein concentration was measured by transferring the mixture to a plastic cuvette (Sigma, Cat. n° 759150) which was then inserted into the BioPhotometer (Eppendorf).

2.3.21. Protein electrophoresis

50 – 100 µg of protein samples were diluted in water up to a final volume of 26 µl and were mixed with 10 µl of the 4X NuPAGE® LDS Sample loading buffer (Cat. n° NP0007) and 4 µl of the 10X NuPAGE® Sample Reducing agent (Cat. n° NP0009). These were denatured by boiling for 10 minutes and loaded on a precast NuPAGE® Novex Tris-Acetate 3-8% gradient gel (Invitrogen, Cat. n° EA03785) placed in an electrophoresis apparatus (BioRad) filled with 800 ml of the diluted 20X NuPAGE® Tris-Acetate SDS Running buffer (Cat. n° LA0041). 10 µl of the HiMark pre-stained protein standard (Life Technology, Cat. n° LC5699) and the Odyssey One-Color protein molecular weight marker (Licor, Cat. n° 928-40000) were included as a reference molecular weight to analyze the protein of interest. 500 µl of the NuPAGE® Antioxidant (Cat. n° NP0005) were then added to the buffer inside the inner chamber prior to electrophoresis, which was done on ice. The gel was run at 75V for 45 minutes to allow dystrophin to slowly enter the gel and then at 150V for 2 hours and 15 minutes.

2.3.22. Western transfer and Immunoblot

a) Turbo Blot

At completion of the run, the Tris-Acetate 3-8% gel (section 2.3.21) was placed in contact with a polyvinylidene difluoride (PVDF) low-fluorescence transfer membrane (Thermo Fisher Scientific Cat. n° 22860), previously activated in 10 ml methanol and washed in 20X NuPAGE® Transfer buffer (Cat. n° NP0006) diluted to 1X in distilled water. Transfer was performed with the Trans-Blot® Turbo™ Transfer System in 7 minutes. After transfer, the membrane was washed with Tris-buffered saline (TBS) supplemented with 0.5% (v/v) Tween 20 (Sigma) (TBS-T) for 20 minutes, blocked for 60 minutes at room temperature with 4 ml of Odyssey Blocking solution and incubated overnight at 4°C with the following primary antibodies: rabbit anti-human dystrophin (Abcam, 15277, 1:1000) and mouse anti-human β-actin (Abcam, 1:2000). The membrane was washed with TBS-T three times and incubated with the Alexa Fluor 680 and Alexa Fluor 790 secondary antibodies (Thermo Fisher Scientific) (both diluted 1:15000 in Odyssey Blocking buffer). After another three washes with TBS-T, the membrane was developed with Odyssey® Infrared Imaging System.

b) Wet Blot

Protein samples were prepared and run in a Tris-Acetate 3-8% gel (section 2.3.21). At completion of the run, the gel was removed from the thin plastic layers and placed in contact with a PVDF low-fluorescence transfer membrane as detailed in section

2.3.22.a. The electrophoresis apparatus XCell Sure Lock™ Mini-Cell was used for protein transfer, performed at 30V for 3 hours. 300 ml of diluted Transfer buffer (270 ml distilled water, 15 ml 20X NuPAGE® Transfer buffer and 15 ml methanol) were placed in the inner chamber, while the outer chamber was filled with 600 ml of cold water, both replaced after 1.5 hour. The transfer apparatus was placed in ice to avoid overheating.

Following blotting, the membrane was blocked with 10% non-fat milk powder (Cat. n° LP0031, OXOID) diluted in TBS for 1 hour at room temperature. The PVDF membrane was then incubated overnight at 4°C with 10 ml 5% non-fat milk diluted in TBS-T, which contained the primary rabbit anti-human dystrophin (Abcam 15277, 1µg/µl 1:200) and mouse anti-human vinculin (Sigma, V9131, 1:100.000) antibodies. The membrane was then washed three times with TBS-T and incubated with biotinylated secondary antibodies (Abcam) (1:1000) for 1 hour at room temperature. Following another three TBS-T washes, the membrane was incubated 1 hour at room temperature with streptavidin-HRP (1:5000) (Thermo Fisher Scientific) in order to amplify the signal. The membrane was washed with TBS-T and incubated for 1 minute with 2 ml of the Luminata Forte™ Western HRP substrate (Millipore). Signal was detected placing the membrane in a ChemiDoc™ Imager and densitometric analysis of detected bands was performed by using the ImageLab software (BioRad). Each lane (where protein were loaded) was manually detected by using the “Lane and Bands” tool, and the background adjusted prior to band detection. The sensitivity of band detection was set up as high or low for clearly visible or faint bands, respectively. Bands that were missed through the automatic detection were manually added. For each lane, information regarding bands densitometry was finally retrieved through the ImageLab “Report” tool.

2.3.23. Microscopy and image capture

The Olympus IX Inverted microscope (Olympus Life Science) and Leica DMR microscope (Leica Microsystem) were used to acquire images. Generally, a minimum of five images were acquired for each tested condition to ensure the reliability of the results. Any image processing was done by Fiji software (<https://fiji.sc/>).

2.3.24. Statistical analysis

Where specified, replicate experiments were expressed as the mean +/- the standard error of the mean. Statistical analysis aimed to verify the effect of CRISPR/Cas9 treatment versus untreated controls were performed by means of the GraphPad Prism software. Results were analysed by applying the Mann-Whitney test (comparison between two groups) or Kruskal-Wallis (comparison between three groups or more) as, due the small sample size, data were assumed to be not normally distributed. Statistical

significance was set at P-values below 0.05. Data falling outside 2 x standard error of the mean were considered outliers and were therefore excluded from the analysis.

Chapter 3. Design and validation of CRISPR/Cas nucleases targeting duplications in DMD gene

3.1 INTRODUCTION

3.1.1 The origin of genomic duplications

The duplication of genetic material has for a long time been thought to play a fundamental role in species evolution. In his masterpiece “Evolution by gene duplication” published in 1970, the geneticist and evolutionary biologist Susumu Ohno identified gene duplication as the major force which allowed evolution starting from ancestor genes⁴¹⁸. He hypothesized that, through additional mutational events, the duplicated gene could potentially provide a selective advantage by acquiring a new function compared to the conventional one maintained by the original copy, a phenomenon known as neo-functionalization.

Mutational events of the duplication type involve either whole genomes, entire genes or only part of a gene and occur either in coding or non-coding regions, affecting chromatin structure⁴¹⁹. In eukaryotes, the genome-wide duplication rate varies between 10^{-4} and 10^{-9} gene/year⁴²⁰. Most of the mechanisms linked to the duplication of genomic regions involve errors occurring during DNA replication (slipped-strand mispairing) or meiosis (chromatids misalignments).

Slipped-strand mispairing involves the slippage of the DNA polymerase at the replication fork in the presence of short repetitive sequences (a few nucleotides long), so that the read sequence is accidentally copied more than once⁴²¹.

Also, during meiosis, the misalignment of chromatids in regions containing repetitive sequences (such as the 300 bp long tandem Alu elements) might lead to the exchange of an extra-portion of DNA. This generates a duplication at the recombination site and a parallel deletion on the homologous chromosome⁴²².

Nevertheless, nondisjunction phenomenon occurring during meiosis results in an abnormal number of chromosomes (aneuploidy) or genome copies (polyploidy). Usually, aneuploidy is associated with deleterious effects (as for the trisomy of chromosome 21, also known as Down syndrome) and is therefore less likely to have a selective advantage in the population, while polyploidy is most frequently observed in plants, less commonly in some animals like amphibians and fish^{423,424} and only rarely in humans⁴²⁵.

Finally, gene duplication could appear as a result of cell infection by retroviruses which retro-transcribe the transcripts and integrate the derived retrogenes back in the parental DNA. These duplications could then acquire novel functions or be subjected to a different regulation^{426,427}. This specific case might explain the origin of different genes with high degree of homology, such as dystrophin and utrophin, for example¹⁸⁵.

3.1.2 Pathogenic gene duplications and Duchenne Muscular Dystrophy

The final effect of duplications on biological fitness depends on the population, genomic locus and extent to which they occur. As opposed to entire gene duplications, well-known pathogenic duplications involve only a few exons (or short sequences within them), causing an alteration of the mRNA reading frame or of the protein conformation and thus leading to a defective, non-functional and/or unstable product.

The first identification of partial gene duplications in the dystrophin gene in patients with DMD and BMD dates back to late 1980s^{428–430}. Since then, many more dystrophin duplications were reported across the *DMD* gene (section 1.3.2), whose locus specific duplication rate was estimated as 1.0×10^{-8} duplications/gene/generation⁴³¹. To date, *DMD* duplications account for 10-15% of *DMD* mutations¹. Despite most of the *DMD* duplications (50%) tending to cluster close to the exons 2-20 mutational hotspot¹ (section 1.3.2.), duplications can be found heterogeneously within *DMD*, as observed by Tuffery-Giraud *et al.*⁴³². This work, based on a Caucasian cohort of patients, showed that 70% of *DMD* duplications were reported only once, while the most common duplications in DMD and BMD patients were duplications of exon 2 and exons 2-7, respectively⁴³². Interestingly, such distribution is different in Asian patients, where the most common *DMD* duplications involve exons 8 and 9^{433,434}.

3.1.3 Therapeutic approaches to dystrophin duplications

All patients carrying duplications could potentially benefit from mutation-agnostic approaches such as cell-based therapies (section 1.6.2.c) and AAV-based gene therapies aimed to provide a miniaturized but functional dystrophin molecule (section 1.6.2.d), currently in clinical trials.

Exon-skipping could represent a further therapeutic strategy for small DMD duplications encompassing a single exon, but not for larger multi-exon duplications (as for example the duplication spanning exons 52-62)⁴³⁵. Targeted skipping of exon 2 and exons 2-7 was achieved in human myoblasts carrying exon 2 duplication by using a phosphorodiamidate morpholino antisense oligonucleotide chemistry⁴³⁶. The applicability of antisense therapy for skipping a variety of single exon duplications was also recently confirmed in human fibroblasts by Wein *et al.*⁴³⁷. Moreover, this and other

studies showed that the skipping efficiency varies depending on the chosen exons: successful skipping was obtained with antisense oligonucleotides designed to skip exon 17, exon 18, both exons 43 and 44, exon 44 itself and exon 45^{435,438}. However, when designing exon skipping molecules for removing duplications, the level of skipping has to be carefully modulated. In fact, as observed in these studies, too efficient skipping lead to the skipping of both exon copies, resulting in the transcription of a shorter mutated transcript.

While these approaches should in theory be beneficial for limiting the effects of DMD, none of them acts at the genomic level and repair the mutation itself.

The work set out here aims to explore *in vitro* a therapeutic strategy based on the CRISPR/Cas9 editing tool designed for duplications. CRISPR/Cas9 exploits two specific components (sgRNA and Cas9 nuclease) to recognize and cleave chosen genomic regions (section 1.8). Since it can be directed to the mutated gene itself, CRISPR/Cas9 system would allow the permanent removal of the genomic mutation and the restoration of the wild-type genomic sequence. Additionally, the novelty of my study relies on the removal of the extra genomic portion using a single CRISPR/Cas9 nuclease designed to cleave a specific site within *DMD* duplication (**Fig.1.9**, section 1.9).

Once the applicability of CRISPR/Cas9 for correcting duplications has been demonstrated, the ease of customization and the limited cost associated with this tool would allow CRISPR/Cas9 to be targeted to any *DMD* duplication, contributing to the field of personalized medicine⁴³⁹.

3.2 AIMS

The design of nucleases targeted to genomic duplications relies on a series of experimental steps, detailed in the results sections below and summarized in **Fig. 3.1**. These steps not only include the selection of a genomic target within the duplication and of a vector suitable for the nuclease expression, but also require the identification of appropriate CRISPR/Cas9 delivery tools for the cell types in which it will be tested.

The following aims were therefore pursued:

- The selection of a genomic target within a specific duplicated region of the *DMD* gene
- The design of CRISPR/Cas9 nucleases targeted to the chosen *DMD* sequences
- The assessment of the efficiency of designed CRISPR/Cas9 nucleases in the immortalized human embryonic kidney cells (HEK 293T)
- The identification of the best delivery method for the expression of CRISPR/Cas9 nucleases in DMD patient-derived cells

Design of CRISPR/Cas9 nucleases targeting *DMD* duplication

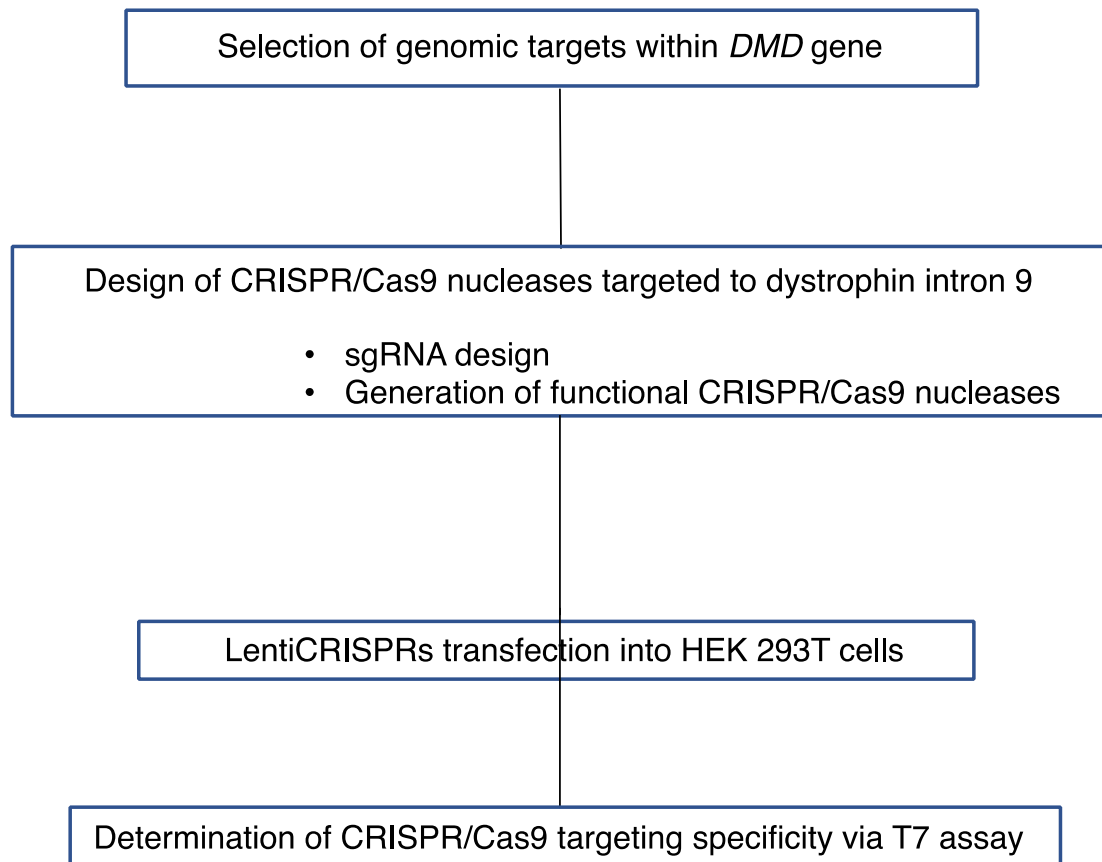


Figure 3.1. Schematic of the experimental steps required for CRISPR/Cas9 design and validation.

3.3 RESULTS

3.3.1. Selection of genomic targets within dystrophin

To demonstrate the ability of a CRISPR/Cas9-based approach to repair dystrophin duplications, I decided to design CRISPR/Cas9 nucleases targeted to DMD intron 9, as this genomic region was included within most of the dystrophin duplications harboured in the cell lines available to me at the time of the experimental set up (**table 3.1**). Moreover, intron 9 was among the introns within the 2-20 hotspot where about 5% of genomic breakpoints were detected in DMD patients carrying duplications⁴⁴⁰.

Patient #	Duplicated exons	Approximate duplication size (Mb)	Cell type available
1	3-25	0.40	Myoblasts
2	5-11	0.17	Myoblasts; Fibroblasts
3	8-11	0.10	Myoblasts
4	8-13	0.12	Myoblasts
5	8-15	0.13	Myoblasts; Fibroblasts
6	8-17	0.18	Fibroblasts
7	8-43	0.50	Fibroblasts

Table 3.1. Range of duplications in DMD patient cells lines stored in the MRC Centre for Neuromuscular Diseases BioBank.

All these cells could be corrected using CRISPR-Cas9 nucleases targeted to DMD intron 9.

First, I retrieved the full sequence of intron 9 of the human *DMD* gene through the Ensemble genome browser (www.ensembl.org/). Then, I selected four regions devoid of known polymorphisms and spanning from 176 to 225 bp throughout the intron. The absence of polymorphisms was a necessary requirement to prevent mismatches between the annotated sgRNA and its target sequence, as polymorphisms have been shown to lead to a reduction or loss in the efficiency of the nuclease⁴⁴¹. The length of the selected sequences was instead chosen in accordance with the specifications provided by the algorithm I used for the design of sgRNAs (section 2.1.5.a), which required inputs of genomic sequences not longer than 500 bp.

Selected genomic sequences are listed below in order from the 3' to the 5' of intron 9,

and were named Targets 1-4. Information about the genomic location is also provided, flanking each sequence:

Target 1

32,646,235.CAGCAGACATGTAGTTTATCCCAAGAAGCCTCTCCTGTCACTTGAGT
GTATGGTTCCTATAGCATGCAAATCTCTGCAAACATATGGCCAGGTTTACAGCTC
ATGTAAACTACCCCGTCTCTATTACTTGGCTTCAGATGGGGAGAAGTTAGAGCCA
CGCAGAGCGCTCCAGCCTCTCCAGAGGAAGCCTTTCTCACCAGATTTTTCTAATC
GA CCTCCTTGCTT.32,646,010

Target 2

32,662,890.AAAATTGCCACAGTATACATTAAATATCTTTGTTTATCCCATTTTAGTT
ATTTTGTGTGATATCTCTAAATGCTGTTCTGTATATGCTGCTGACAAGGAGCAAAT
TTATTTGTTTAGAAATTCTGCAGCCTTTCCATCCTTTGCCTTAACCTTTCTTGTTGAG
GTACTAATGAGCTGTGA.31,656,710

Target 3

32,691,476.ATTTCCCTGATAATTCATAATGTTGAACATATTTTCATACACCTGTTGA
CCATTTGGGTGTCGTCTTTGGAGAAATGTCTGTTCAAGCCTTTAGCCCATTTTTAA
TTGCGTTTTTAGTTTTTTTTTTTTTTTGCCAATGAGTTGTAGGAATTCTTTATATATTTT
AGAAGATAACCAG.32,691,300

Target 4

32,696,670.TGAGCTGCCTTTTCTAGGTTTTCTCCAATGCCTATTCTGGTAGCACAT
TTTCACCTCATCACCACCTTCATTAAGGCATCGCTCAGGTATGACTCCTTAATGAA
GTAATTACAAATGGACTTACATCTTGCTAGGCTCTGCTCATTCTACCTTTGTGTAA
ATATTACCAAATACATAAT.32,696,490

3.3.2. Design of CRISPR/Cas9 nucleases targeted to dystrophin intron 9

a) sgRNA design

Each of the targets mentioned in the above section 3.3.1. was run through the sgRNA-design algorithm developed by Feng Zhang laboratory at MIT (<http://crispr.mit.edu/>). This algorithm returned a list of sgRNAs recognizing sequences in proximity of a 5'-NGG-3' site, which represents the PAM motif required for Cas9-mediated cleavage (section 1.7.4.a). Returned sgRNAs were accompanied by a quality score predictive of potential off-target effects: the higher the score, the lower the presumed off-targeting activity. As such, for any genomic input I selected the sgRNA with the highest quality score. For each sgRNA sequence, the algorithm also provided additional information such as its strand location and quantified the presumed off-target sites, specifying how many of these lie within genes (**Fig.3.2**).

Preferred sgRNAs (named from 1 to 4) all lie in the forward genomic strand, apart for sgRNA2. Their sequence and the PAM motif they recognize are indicated below:

sgRNA1: ACTACCCCGTCTCTATTACT (PAM = TGG)

sgRNA2: GTACCTCAACAAGAAAGTTA (PAM = AGG)

sgRNA3: GACCATTGTTGGGTGTCGTCTT (PAM = TGG)

sgRNA4: CCTTCATTAAGGCATCGCTC (PAM = AGG)

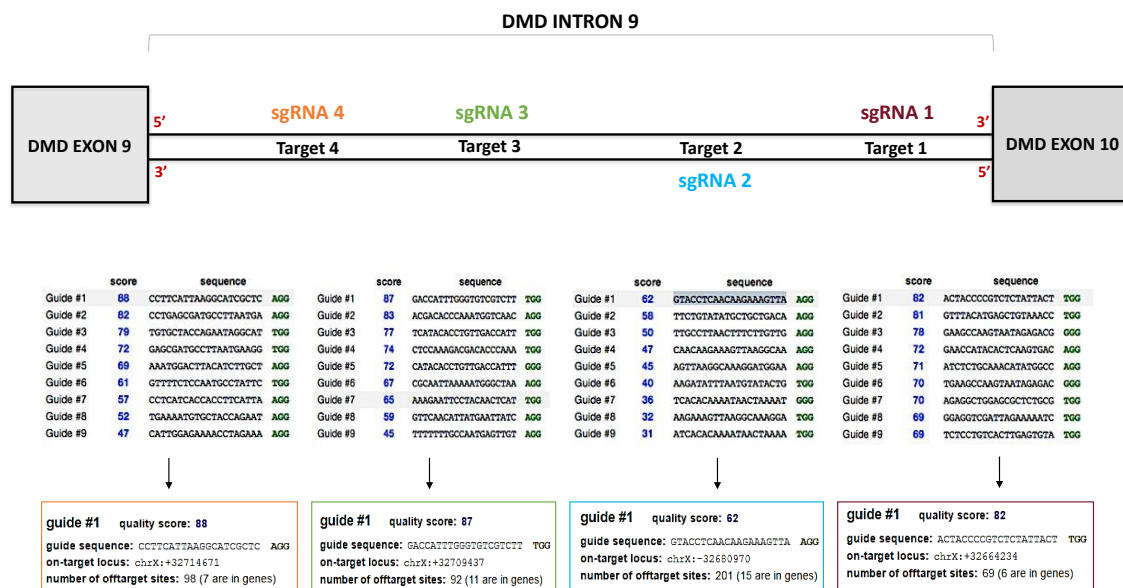


Figure 3.2. Design of sgRNAs targeted to human DMD intron 9.

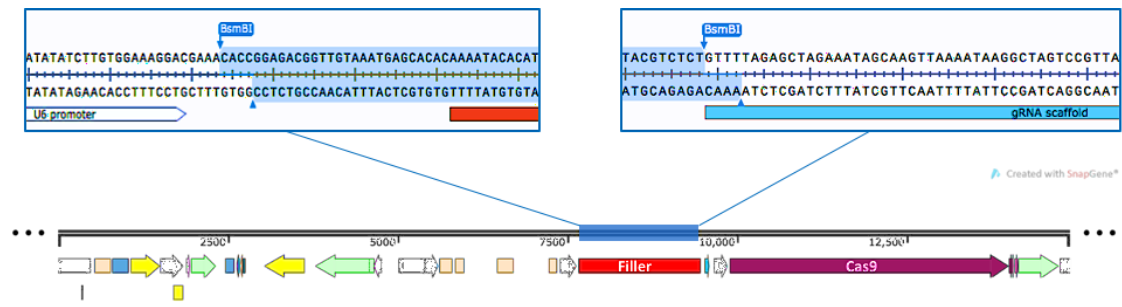
Four sgRNAs were designed to recognize sequences devoid of polymorphisms spanning DMD intron 9 from 3' to 5' (Target 1-4). Input target sequences were run through the algorithm, which returned a number of sgRNAs scored depending on their homology with off-target sites (the highest the score, the lowest the off-target sites detected). Nine sgRNAs (guide#1-9) are shown. For each ranking, the top target sgRNA (guide#1) was selected. As well as the quality score and the nucleotide sequence, information about the on-target site and potential off-target sites were also provided for each sgRNA.

b) Generation of functional CRISPR/Cas9 nucleases

Once the sgRNA design was accomplished, I had to choose a vector capable of expressing the CRISPR/Cas9 nuclease system in human fibroblasts.

From the CRISPR/Cas9 plasmids designed for genomic cutting and deposited in the Addgene online catalogue (<http://www.addgene.org/CRISPR/#cut>), I selected the LentiCRISPR expression vector pXPR-001 (LentiCRISPRv1) (**Fig.2.1**, section 2.1.2.a). The main reason for my choice is that LentiCRISPRv1 has a lentiviral backbone and is therefore suitable to produce lentiviral vectors that I will use to deliver CRISPR/Cas9 in human fibroblasts. I considered viral vectors because transfection of non-viral plasmids by commercial reagents as Lipofectamine 2000, run in our laboratory prior to the beginning of this project, had a poor efficiency in these cells (Francesco Conti, personal communication). Another advantage of LentiCRISPRv1 is that it allows the simultaneous expression of both sgRNA and the Cas9 nuclease. This obviates the need for separate constructs for the expression of the two CRISPR/Cas9 components. LentiCRISPRv1 expresses sgRNA by the constitutive U6 promoter, while the expression of the Cas9 is driven by the constitutive elongation factor 1 α core (EFS) promoter, active in a considerable number of cell types including human fibroblasts⁴⁰¹. This plasmid also expresses two antibiotics (ampicillin and puromycin), each having a specific purpose. Ampicillin is essential for the isolation and amplification of bacterial colonies expressing LentiCRISPRv1, while puromycin aids the isolation of positively transfected or transduced cells only and could therefore be used to monitor the percentage of cells expressing CRISPR/Cas9.

The original stock of bacteria expressing the LentiCRISPRv1 plasmid was propagated as detailed in section 2.1.1.c. and isolated plasmid DNA was quantified by using the Nanodrop spectrophotometer (section 2.3.2). The plasmid DNA with the highest concentration and purity was selected for the cloning of each sgRNA (section 2.1.5.d). This step required the digestion of LentiCRISPR by means of the BsmBI restriction enzyme (section 2.1.5.b), which recognizes two specific plasmid sites. These flank a 2 kb region called filler, which needs to be removed to allow sgRNA cloning into the vector (**Fig.3.3**). As such, BsmBI digestion should result in two fragments of about 11 kb and 2 kb, as I confirmed by running the digestion products in a 1% agarose gel (**Fig. 3.3**).



Digested LentiCRISPRv1

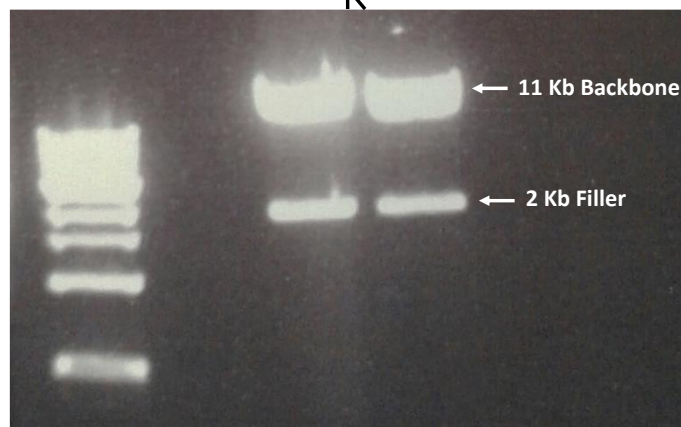


Figure 3.3. LentiCRISPRv1 plasmid digestion.

To allow sgRNA cloning, LentiCRISPRv1 plasmid was digested with BsmBI restriction enzyme. This experimental step allowed the removal of the filler region in between the two BsmBI restriction sites (top panel). The products of BsmBI digestion were run on a 1% agarose gel (bottom panel). As expected, BsmBI digestion resulted in two fragments (indicated by arrows). The bigger 11kb fragment was used for sgRNA cloning, while the smaller 2kb fragment (corresponding to the filler) was discarded.

I gel purified the larger fragment by using the QIAgen Gel extraction kit (section 2.3.6.b) and ligated it to sgRNA oligos designed to have complementary nucleotides to BsmBI cleaved sites, as detailed in section 2.1.5.c. and 2.1.5.d. (**Fig.3.4**).

I used ligation products to transform the recombinase-deficient One Shot® Stbl3 *E. coli* bacteria (section 2.1.1.d), which I then propagated as detailed in section 2.1.1.e. The following day, I extracted bacterial DNA (2.1.1.f.) from three ampicillin-resistant colonies for each cloning reaction and sent it for Sanger sequencing starting from the U6 promoter. This step was needed to authenticate the correct ligation of each sgRNA into the digested LentiCRISPRv1. As I confirmed sgRNA insertion in 11 out of the 12 selected clones, for each sgRNA I selected one successful clone for validating its *in vitro* functionality. From now on, these vectors will be referred to as LentiCRISPR1, LentiCRISPR2, LentiCRISPR3 and LentiCRISPR4 by virtue of their sgRNA targeting dystrophin sequences 1 to 4, respectively.

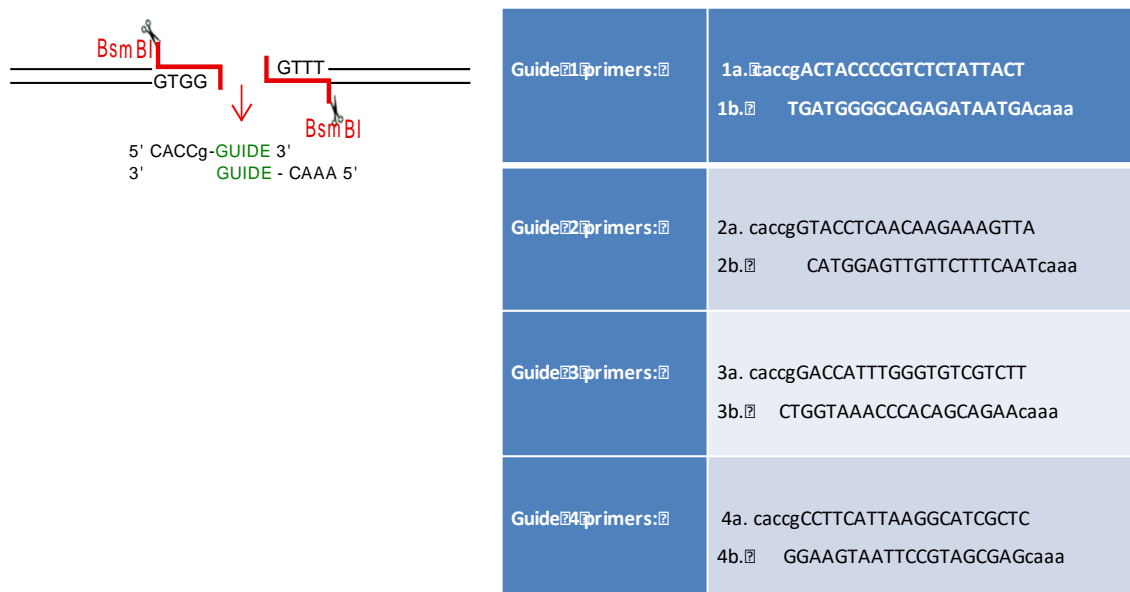


Figure 3.4. Schematic of sgRNA primers design.

sgRNAs primers had to be cloned in the LentCRISPRv1 vector digested with BsmBI restriction enzyme, that performed two staggered cuts (left panel). sgRNA primer pairs were therefore designed to carry at both extremities additional nucleotides complementary to the overhangs generated by BsmBI digestion. The sequence of each sgRNA primers pair is shown in the right panel.

3.3.3 LentiCRISPRs transfection into HEK 293T cells

Before testing the designed LentiCRISPRs in patient-derived cells, I used HEK 293T cells to show that these nucleases were functional and could cleave the genomic regions of interest. This cell line provided the ideal choice for such an experimental step, as it has a low cost of maintenance and can be easily transfected with a variety of reagents and techniques⁴⁴².

I transfected each LentiCRISPR plasmid in HEK 293T cells using the Lipofectamine 2000 (LF2000®) transfection reagent, with the DNA:LF2000® ratio of 1:2 (section 2.2.11.a) previously optimized in our laboratory (Francesco Conti, personal communication). As controls I used non-transfected cells and cells transfected with the original non-digested lentiCRISPRv1 vector which, being devoid of any sgRNA, serves to control for any off-target nuclease cutting in the genome.

In parallel, I also transfected HEK 293T cells with the plasmid pEGFP-C2 (**Fig.2.9**) which, as it expresses enhanced GFP, allowed me to visually monitor transfected cells and thus to evaluate how effective the transfection was (**Fig.3.5**).

The negative and positive controls are summarized in **table 3.2**. The experiment was repeated three times, each being a biological replicate.

Control (+/-)	Transfected Plasmid	Function
-	None	Monitor assay background
-	Undigested LentiCRISPRv1 (w/o sgRNA)	Monitor non-specific genomic cleavage
+	pEGFP-C2	Assess transfection efficiency

Table 3.2. List of controls included for CRISPR/Cas9 validation in HEK 293T cells

48 hours post-transfection, I fixed the cells transfected with pEGFP-C2 by using 4% paraformaldehyde and I counterstained their nuclei with DAPI (section 2.3.15). I calculated the pEGFP-C2 transfection efficiency by dividing the number of GFP-positive cells by the total number of cells (represented by the number of nuclei) and multiplying the resulting value by 100. The percentage of GFP-positive cells appeared to be uniform across the three transfections, being around 40% (39.1 +/- 1.82).

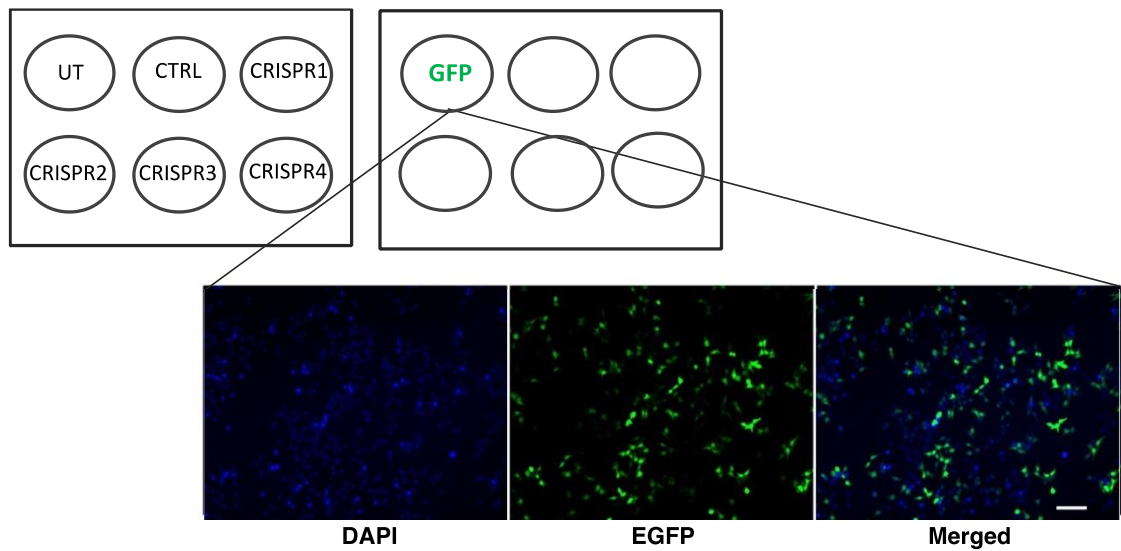


Figure 3.5. HEK 293T cells transfection efficiency.

HEK 293T cells were transfected with either undigested or functional LentiCRISPRs and, in parallel, with the pEGFP-C2 plasmid expressing EGFP. Transfection efficiency was calculated as the percentage of GFP+ cells. The mean transfection efficiency across experimental triplicates was around 40%.

UT=untreated cells; CTRL=cells transfected with undigested LentiCRISPR; CRISPR1=cells transfected with LentiCRISPR1; CRISPR2=cells transfected with LentiCRISPR2; CRISPR3=cells transfected with LentiCRISPR3; CRISPR4=cells transfected with LentiCRISPR4; GFP=cells transfected with pEGFP-C2 plasmid to evaluate transfection efficiency. Images were acquired at a magnification of 10X. Scale bar = 100 μ M

3.3.4. Determining targeting specificity and efficiency of CRISPR/Cas9 into HEK 293T cells

As there was such a high percentage of transfected HEK 293T cells, I did not exploit antibiotic-resistance screening for selecting only LentiCRISPR-expressing cells. 48 hours post-transfection, I extracted genomic DNA (section 2.3.1) from the total pool of both non-treated HEK 293T and HEK 293T transfected with either the undigested or functional LentiCRISPRs to assess the genomic targeting in this cell type. For this step I did not include cells transfected with pEGFP-C2, as these cells would have represented a further negative control for the genomic cleavage in addition to the two mentioned in the table 3.2.

I assessed the reliability of the primer pairs surrounding each CRISPR/Cas9 target site by amplifying genomic DNA from untreated HEK 293T cells via tdPCR (section 2.3.4).

As shown in **Fig.3.6**, the resulting amplicons had the expected molecular weight, indicated below.

LentiCRISPR1 amplicon = 710 bp

LentiCRISPR2 amplicon = 450 bp

LentiCRISPR3 amplicon = 630 bp

LentiCRISPR4 amplicon = 565 bp

I then repeated tdPCR including the remaining samples (i.e. samples transfected with undigested and functional LentiCRISPRs) and I PCR purified resulting amplicons (section 2.3.6.a).

I evaluated the LentiCRISPRs' cutting ability by performing the T7 endonuclease assay (section 2.3.7) on these amplicons and running the T7 assay products on a gel.

For each negative control sample indicated in table 3.2, I expected to visualise only the band corresponding to the full-length amplicon. The molecular weight of each amplicon would depend on the primer pair used to amplify each CRISPR/Cas9 target site (section 2.3.4).



Figure 3.6. TouchdownPCR of LentiCRISPR target sites.

The genomic DNA of untreated HEK 293T cells was amplified by touchdown PCR to assess the specificity of the primers pairs designed to amplify each LentiCRISPR target site. Detected amplicon bands (named CRISPR1-4) confirmed expected sizes of 710bp (CRISPR1), 450bp (CRISPR2), 630bp (CRISPR3) and 565bp (CRISPR4).

For samples transfected with functional LentiCRISPRs, I expected instead to obtain three bands. The upper molecular-weight band would correspond to the full-length amplicon, while the other two bands would represent the cleaved fragments resulting from the LentiCRISPR cut within the amplicon. The expected size of the smaller fragments indicative of each side of the cleaved LentiCRISPR target sequence is shown in **figure 3.7**.

In the T7 assay performed on the set of samples derived from the first and second transfection, genomic DNA from the negative controls was only amplified with the primer pair designed to amplify LentiCRISPR1 target site. The third T7 assay, instead, included controls amplified with all the remaining primer couples.

I loaded the products of the first T7 assay on a 2% agarose gel but, due to the long run time necessary to distinguish between the different molecular sizes of untargeted and cleaved fragments, the results were not conclusive due to poor resolution of this gel (**Fig.3.7**). In subsequent experiments I therefore resolved the T7 assay products on 4-20%TBE gels (section 2.3.8), achieving a better resolution and consequent band detection.

As expected, only the T7 reaction products from samples transfected with functional LentiCRISPRs contained multiple bands corresponding to the uncleaved higher molecular weight fragment and the two smaller cleaved fragments (**Fig.3.7**.) For each tested LentiCRISPR other than LentiCRISPR3, both the full-length and the cleaved fragments appeared to have the predicted molecular weights, confirming that the designed nucleases efficiently and specifically target the selected genomic regions in HEK 293T cells. Despite the presence of background bands in all samples (even in controls), LentiCRISPR3 amplicons exhibited bands other than the expected ones. However, non-specific bands of the same size were also present in control amplicons obtained with the same primer set. This suggested that non-specific bands may have been derived from misalignment of the primers to similar genomic sequences and not from a non-specific nuclease activity. Considering that the other three nucleases were successful in targeting the region of interest, I did not use LentiCRISPR3 in further experiments.

Having confirmed the functionality of the LentiCRISPRs, I next determined their efficiency by quantifying the enzyme-cut versus uncut fragments (section 2.3.9). This was calculated by band densitometry for each LentiCRISPR except for LentiCRISPR3. LentiCRISPR1 showed the best cleavage activity when resolved on agarose gel (31%), but fragment detection on TBE gels highlighted a similarly strong cutting operated by

LentiCRISPR2 and LentiCRISPR4 (35% and 29%, respectively).

By averaging the LentiCRISPR efficiencies derived from each T7 assay, I found that LentiCRISPR1 and LentiCRISPR2 displayed the highest nuclease activity of 31.8% \pm 0.2 and 35.3% \pm 0.56 cleavage, respectively. These two nucleases were therefore chosen to be expressed in lentiviral vectors and tested in human fibroblasts.

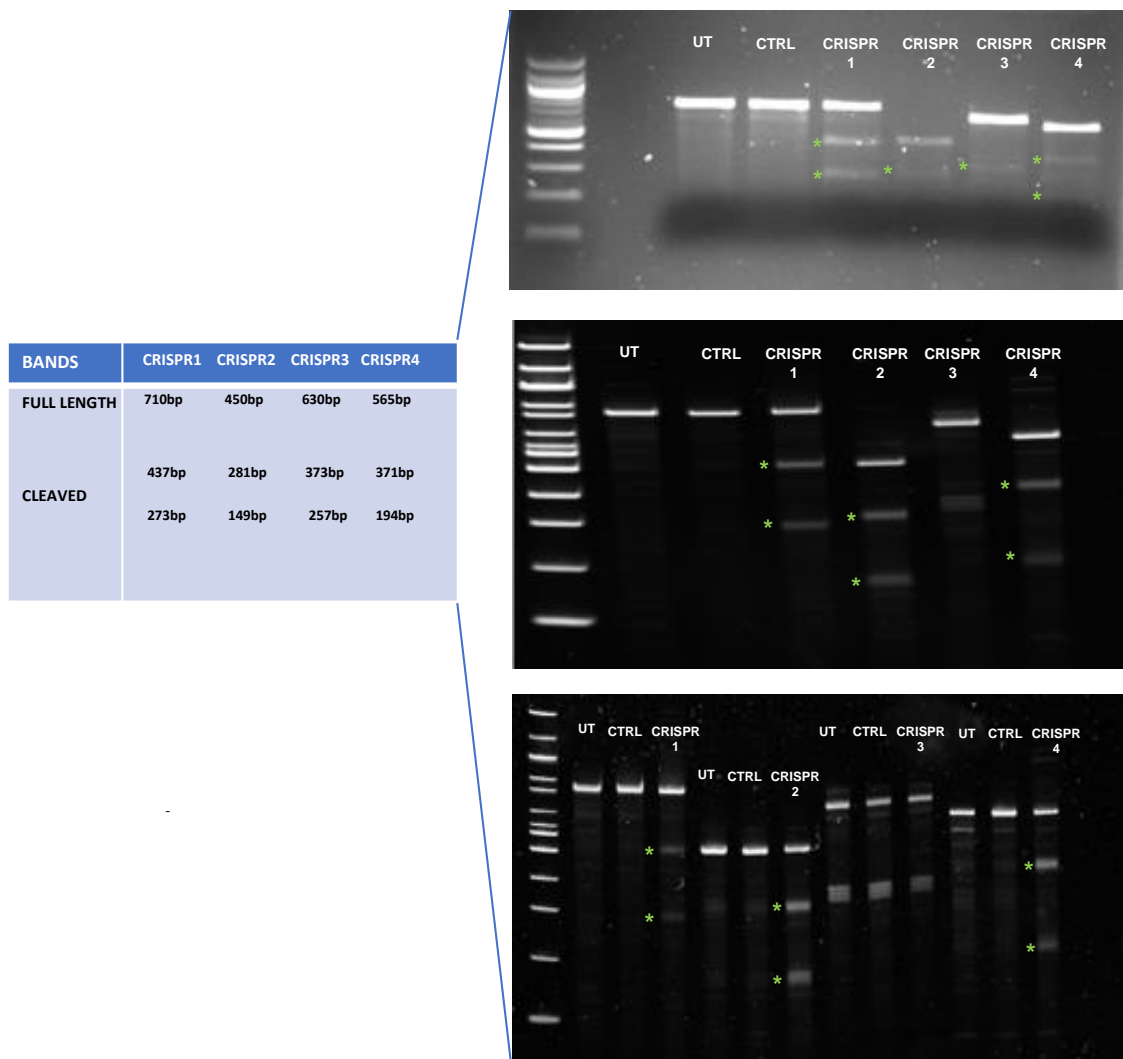


Figure 3.7. Analysis of CRISPR/Cas9 efficiency in HEK 293T cells.

LentiCRISPR efficiency was assessed via T7 assay performed on purified and annealed amplicons. Expected sizes of full length and cleaved fragments derived from T7 assay are shown in the table on the left. T7 assay products were loaded on a 2% agarose gel (top panel) and on 4-20% Novex® TBE gels (middle and bottom panels). All LentiCRISPRs presented with cleavage products, but only LentiCRISPR1, LentiCRISPR2 and LentiCRISPR4 showed products of the predicted size (*). LentiCRISPR3 exhibited multiple, possible non-specific bands.

UT: DNA from non-transfected cells; CTRL: DNA from cells transfected with undigested LentiCRISPR vector; CRISPR1, CRISPR2, CRISPR3, CRISPR4: DNA from cells transfected with LentiCRISPR 1 to 4.

3.4 DISCUSSION

As the purpose of the project is to demonstrate the ability of CRISPR/Cas9 to correct dystrophin duplications causing DMD, I designed CRISPR/Cas9 nucleases to target different regions of the *DMD* gene which were duplicated in patient-derived cell lines available to me.

To identify CRISPR/Cas9 nucleases that can be successfully expressed in DMD patient-derived cells, I had to choose not only a target dystrophin region, but also a pipeline to design the nucleases and an assay suitable to validate their efficiency in a simple cellular model.

As a nuclease target, I chose a non-coding region to avoid misreading of the gene reading-frame that can be introduced by the error-prone NHEJ repair mechanism (the primary mechanism which re-joins cleaved DNA ends in absence of a DNA template)⁴⁴³. Among the introns that were duplicated in my cell lines, I chose to direct the CRISPR/Cas9 nucleases against four different locations across the length of *DMD* intron 9, listed among the introns in which genomic breakpoints are frequently reported⁴⁴⁰.

Moreover, the intron 9 sequences I chose were devoid of polymorphisms, as these could result in mismatches between the sgRNA and target DNA. These can affect Cas9 activity to an extent directly related to the number of mismatches, their distribution within sgRNA and their distance from the trinucleotide 5'-NGG-3' (PAM motif, section 1.7.4.a) that is recognized by the sgRNA⁴⁴⁴. Hsu *et al.* observed, in fact, that Cas9 cleavage is reduced in presence of two mismatches and further decreased when three consecutive bases were changed. Also, no more DNA breaks were observed if the three nucleotide variations appeared in a scattered fashion or if the number of concatenated mismatches increased up to five⁴⁴⁴.

However, Cas9 activity is not affected by certain sgRNA-target DNA mismatches such as those caused by single nucleotide changes, especially when these occur in the PAM-distal region^{445,446}. As such, avoiding polymorphic target regions represented a further precaution to limit the phenomenon of off-target cleavage (section 1.8).

In virtue of these considerations, I looked for a sgRNA design tool able to predict the sgRNA likelihood of resulting in non-specific genomic targeting. More than twenty different on-line algorithms (recently reviewed by Cui *et al.*) have been developed and optimized through the years to improve the performance of sgRNAs and consequently of the CRISPR/Cas system⁴⁴⁷. At the time I had to design sgRNAs to intron 9, the best available algorithm was the one developed by the laboratory of Feng Zhang (Broad

Institute) who, together with George Church, is a well-known scientist in the CRISPR field. In fact, for the first time, their research teams successfully addressed CRISPR/Cas9 editing system to the genome of human cell lines^{380,382}. This algorithm, utilized by Hsu *et al.*⁴⁴⁴, returned a list of sgRNAs correlated by a score indicative of their faithfulness, expressed as a percentage from 0 to 100, with the guides presented at the top of the list having the highest values. SgRNAs scores are inversely correlated with the likelihood of having off-target effects. The advantage of using this tool is that, for each sgRNA, the interface provides the total number of predicted off-target sites and, also, gives the sequence and genomic location of the 20 most-likely off-target sites. I considered this feature essential for monitoring the genome-wide effect of the nuclease I chose, even though the study of the off-target effects was not prioritized and will constitute future work for the project.

Interestingly, sgRNA2 (which appeared to be the most efficient sgRNA in HEK 293T cells) was the sgRNA with the lowest quality score and the highest number of potential off-target sites. I considered that the actual number of off-target sites for this sgRNA might be even higher than what estimated by the algorithm, as Zang's tool refers only to the 5'-NGG-3' PAM sequence without considering the extended cutting ability of Cas9 from *S. pyogenes*. Its enzymatic activity, in fact, is not restricted to 5'-NGG-3' PAM, but is also associated (albeit at lower level) to 5'-NAG-3- or 5'-NGA-3' PAMs⁴⁴⁴.

Through the years following my sgRNA design, several research groups found that specific parameters other than the presence of a PAM sequence close to the target region can influence the sgRNA specificity (and consequently the number of off-target episodes)^{448,449}. I retrospectively analysed the best sgRNA (sgRNA2) in light of these findings, which related the efficiency of sgRNA to the specific PAM motif recognized by the sgRNA, the sgRNA length and its nucleotide composition.

Recent studies also showed that the variable nucleotide (N) in the conventional 5'-NGG-3' PAM affects targeting efficiency⁴⁵⁰, with 5'-CGG-3' PAMs favoured over 5'-TGG-3'. However, when re-analysing the sgRNAs I designed, I observed that none of these recognized the CGG motif.

Furthermore, all the sgRNA I designed were 20 nucleotides long as suggested by Hsu *et al.*⁴⁴⁴, who correlated a sgRNA length of 20-22 nucleotides with a better sgRNA stability. However, Fu *et al.* demonstrated that the use of truncated sgRNA 17-18 nucleotides long (as opposed to the classic sgRNAs length of 20 nucleotides) can reduced undesired off-targeting without affecting the on-target cleavage⁴⁵¹.

In regards to sgRNA nucleotide composition, a low frequency of T or TT is preferred to preserve sgRNA efficiency⁴⁵². Also, purines play crucial roles, A being preferred in the central region while G being favoured in the PAM-distal region⁴⁵³. Interestingly, I

observed both these requirements in the sgRNA2. This might explain why this sgRNA was associated with a high cleavage efficiency when tested in HEK 293T cells. However, work by Moreno-Mateos *et al.* showed that G-enrichment and A-depletion increase sgRNA stability and efficiency⁴⁵⁴, which is in contrast to the above considerations about sgRNA2, as it is predominantly composed by As.

Considering these new pieces of information, should I have to design new sgRNAs targeting dystrophin intron 9, I would use an algorithm able to provide high efficiency sgRNAs taking into account the above mentioned features, but maintaining the prediction of off-target effects, such as CRISPOR or GuideScan^{455,456}.

The efficiency of each sgRNA in directing the Cas9 nuclease to the specified genomic target for its cleavage was assessed via the T7 assay, which I chose among the assays designed to recognize the insertion/deletions (indels) that cluster around the genomic breaking site introduced by the NHEJ repair pathway.

Similarly to the slightly less sensitive Surveyor assay⁴⁵⁷, T7 exploits the nuclease activity of an endonuclease (in this case the phage-encoded T7 enzyme) to specifically cleave heteroduplexes formed consequent to the annealing on wild-type sequences and edited samples carrying indels⁴⁵⁸.

Recently, Sentmanat *et al.* observed that, compared to newer sequencing technologies, the T7 assay underestimates indels frequency and questioned its accuracy as a tool to assess sgRNAs efficiency⁴⁵⁹. Indeed, she showed that the best sgRNA resulting from NGS analysis had a modest activity when screened via T7 assay, while low-scoring sgRNAs in NGS failed to show cleaved bands in the T7 assay. Also, she observed that sgRNAs having similar efficiency in T7 assays clearly differed following NGS analysis. These results reiterated the findings of Liesche *et al.*⁴⁶⁰, i.e. that above a threshold indels level T7 assay is not accurate. Such underestimations have been correlated to biases introduced at different stages of the assay, either because of faulty heteroduplex formation and/or in results analysis. The former is dependent on the indel type as well as their length and abundance. For example, simple nucleotide changes that do not necessarily give rise to hairpins are rarely detected with T7⁴⁶¹. The latter can instead derive from erroneous densitometric analysis possibly coupled with non-specific background banding. Taken together, these factors contribute to limiting the outcome to be a semi-quantitative output.

The use of NGS to discriminate indels or, similarly, deep-sequencing of single DNA molecules cloned into bacterial backbones could improve the quantitative estimate of double strand breaks, but these approaches are time-consuming and their use is generally limited by their prohibitive costs⁴⁶². In virtue of these considerations and the

fact that T7 assay not only proved to be technically simple, easy to interpret and also cost-effective, I considered it as the best option for screening the nuclease activity of the nucleases I designed.

T7 assay allowed me to demonstrate that three out of four CRISPR nucleases effectively introduced double-stranded cleavage in the chosen sites of *DMD* intron 9. LentiCRISPR1, LentiCRISPR2 and LentiCRISPR4 indeed generated cleaved bands of the expected size in each experimental replicate, confirming the specificity of these nucleases.

Densitometric analysis of cleaved bands allowed me to quantitatively determine the higher cleavage efficiency of LentiCRISPR1 and LentiCRISPR2 over LentiCRISPR4, even though that difference did not appear to be dramatic.

However, despite the promising results obtained by the T7 assay, a direct correlation between the efficiency of cut and dystrophin DNA repair can not be necessarily expected. In fact, for the duplication to be excised and repaired, two double-strand breaks need to simultaneously occur in each of the chosen duplicated locations. I hypothesized several scenarios for the proposed study design (**Fig.3.8**).

One possibility is represented by the nuclease targeting only one of the two target sites, leading to a single double-strand break that would be repaired via NHEJ. In this case, the unreliable NHEJ repair mechanism would introduce indels that could disrupt the sgRNA target site, so that it would not be recognized any further (**Fig.3.8.a**). This would prevent any chance of repairing the duplication.

Alternatively, the nuclease could successfully recognize both sites and excise the intervening region. Dystrophin repair in this case would depend on the prevalent repair pathway: as the eliminated region could be perceived as a template DNA, the HDR repair pathway could re-integrate it within the cleaved sites. Such integration could occur either in the same or inverted orientation (**Fig. 3.8.b1-2**).

Only when NHEJ would prevail over HDR, would the DNA breaks deprived of the extra portion be re-joined and the dystrophin sequence corrected (**Fig. 3.8.b3**). According to this, the T7 signal can be thought as a measure of indels distribution and thus of sgRNA activity, but not of the extent of *DMD* correction.

Despite the impossibility to quantify the extent of dystrophin editing at the genome level, the presence of cleaved bands of the expected molecular weight following T7 assay supported the hypothesis that three LentiCRISPRs I generated were able to efficiently target the chosen dystrophin genomic sequence. This outcome set the foundation to the next experimental step intended to test the best nucleases in patient-derived fibroblasts (detailed in Chapter 4).

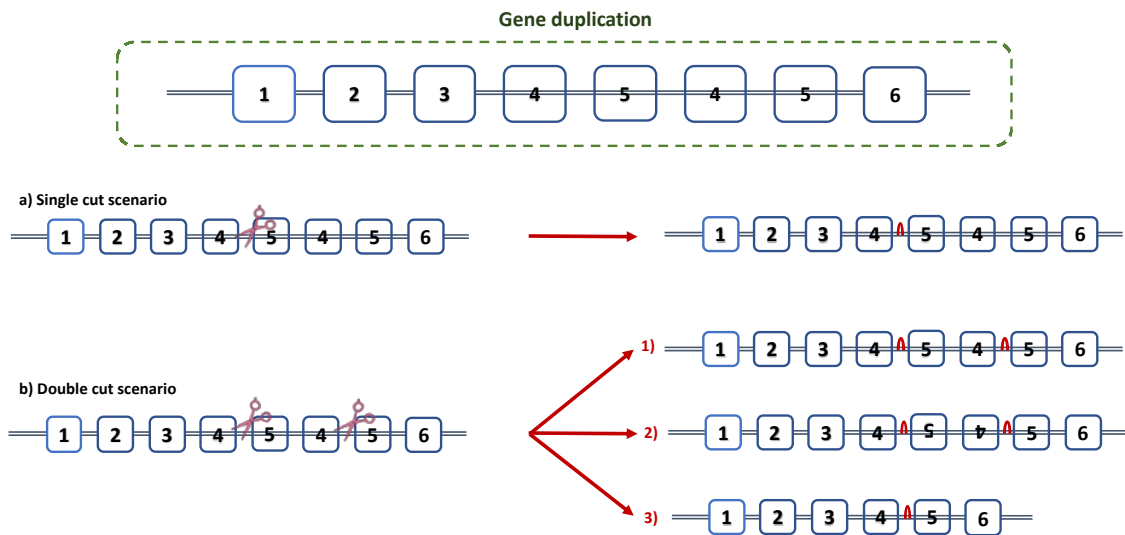


Figure 3.8. Possible scenarios derived from nucleases directed at gene duplications.

Cleaved bands in the T7 assay can be the result of multiple scenarios deriving from a single nuclease approach designed for the removal of genomic duplications; indels (symbolized by the red hairpin) might be the result of either a single (a) or a double (b) nuclease-mediated cleavage. a) Single nuclease cut results in re-ligation and does not allow duplication removal. b) Double cut could either cleaved both duplicated sites and be re-ligated restoring the starting situation (top arrow-1), be re-ligated including the intervening duplication in the inverse orientation (middle arrow-2) or restore the wild-type condition by getting rid of the duplicated fragment in between the two cleaved sites (bottom arrow-3).

Chapter 4. Delivery of CRISPR/Cas9 nucleases to patient-derived fibroblasts carrying dystrophin duplications

4.1 INTRODUCTION

4.1.1. Fibroblasts as a cellular model to prove CRISPR/Cas9 efficiency

Fibroblasts are the most common type of cells within the connective tissue, whose main role is the production and secretion of collagen and other extracellular matrix components such as glycosaminoglycans, reticular and elastic proteins as fibronectin and elastin, and growth factors^{463–466}.

The choice of exploring the nuclease efficiency in fibroblasts was primarily because myoblasts derived from DMD patients carrying dystrophin duplication were not available at the time.

The use of fibroblasts to study neuromuscular disorders has considerable advantages. For example, these cells can be easily isolated from a punch skin biopsy, following which they can be cultured and expanded *in vitro*. Moreover, fibroblasts can be used to generate induced pluripotent stem cells (iPSCs)⁴⁶⁷, from which most lineages of interest to me can be derived, allowing to study the effect of *DMD* mutation in more than one tissue and/or cell type.

However, the analysis of dystrophin transcript and protein in primary fibroblasts is challenging, as only extremely low level of dystrophin transcription has been reported in this cell type⁴⁶⁸. For this reason, observations in regards to dystrophin can only be validated upon conversion of fibroblasts into myoblasts. This is generally achieved through the forced expression of the myogenic regulator MyoD, one of the muscle-specific transcription factors only expressed during embryogenesis and in cells that are committed to myogenic differentiation, but not in adult skeletal muscle, or in any other tissue⁴⁶⁹.

4.1.2. MyoD: forced fibroblast transition to myogenic specification

In late 1980s, Davis *et al.*⁴⁷⁰ isolated the myoblast determination gene number 1 (*MyoD1* or *MyoD*), a regulatory gene essential for initiating myogenic conversion. *MyoD* expression in non-myogenic fibroblasts, in fact, is able to switch on myogenesis and convert them to myoblasts⁴⁷¹. Davis also showed that a highly basic region of 60 amino-

acids within MyoD is responsible for the binding of specific DNA regulatory sequences⁴⁷¹ (enhancers or promoters) of genes controlling myogenesis such as *Myf5*, *myogenin* and *MRF4*. These molecules are involved in the sequential phases of the myogenic process up to the terminal stage of myogenic differentiation, when myoblasts stop proliferating and start to fuse into the multi-nucleated adult skeletal muscle cells known as myofibres⁴⁷².

Halevy *et al.* showed that high levels of MyoD arrest the cell-cycle in the G1 phase of the cell cycle, due to the MyoD-induced inactivation of the cyclin-dependent kinase inhibitor p21⁴⁷³. Blocking the progression into S phase^{474,475}, MyoD expression defines a window of time during which differentiation can occur. Kitzman *et al.* later observed oscillating levels of MyoD in asynchronous myoblasts and showed that *MyoD* and *Myf5* follow a mutually-exclusive expression pattern throughout different phases of the cell cycle, thus suggesting the role of MyoD in the cross-talk between cell proliferation and differentiation⁴⁷⁶.

Other than controlling myogenesis, MyoD also binds to promoter and enhancer sites of other genes whose role may vary from proliferation to migration, fusion and contraction. A recent study by Cao *et al.*⁴⁷⁷ showed that MyoD can affect histone acetylation throughout the whole genome thus altering its epigenetic composition. Interestingly, Shintaku *et al.*⁴⁷⁸ showed that MyoD also binds to metabolic genes involved in the mitochondrial biogenesis, fatty acid oxidation and electron transport chain. This work supported the hypothesis that, by being involved in multiple processes, MyoD contributes to maintain muscle homeostasis.

4.2 AIMS

In this chapter I aim to demonstrate the ability of CRISPR/Cas9 nucleases (previously designed and tested in HEK 293T cells) to remove dystrophin duplications in DMD patient-derived fibroblasts. To do this, I:

- Developed lentiviral vectors to express the CRISPR/Cas9 nucleases in primary fibroblasts
- Investigated alternative non-viral tools to deliver CRISPR/Cas9 nucleases to primary fibroblasts
- Assessed CRISPR/Cas9-mediated cleavage activity at the genomic level
- Attempted forced MyoD-conversion of CRISPR/Cas9-expressing fibroblasts to obtain both dystrophin transcript and protein outputs.

The experimental steps pursued for the accomplishment of these aims (summarized in **Fig. 4.1**) are described in the results sections below.

Dystrophin correction in patient-derived fibroblasts

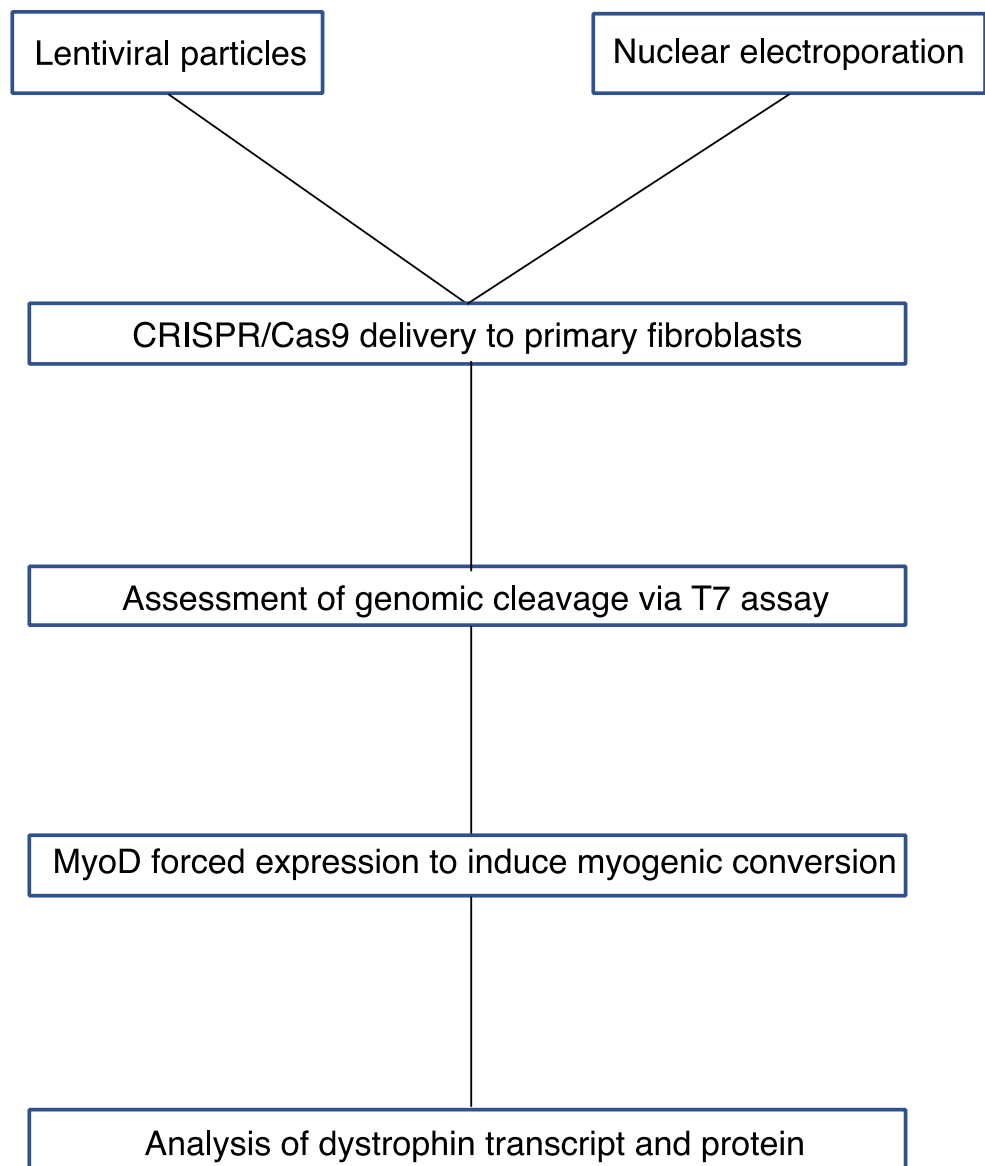


Figure 4.1. Schematic of the experimental steps required for restoring dystrophin in patient-derived fibroblasts.

4.3 RESULTS

4.3.1. Generation of lentiviral particles expressing the best LentiCRISPRs

Since LentiCRISPR1 and LentiCRISPR2 had the highest cleavage efficiency when expressed in HEK 293T cells (section 3.3.4), I decided to test these two nucleases in patient-derived fibroblasts carrying duplications spanning dystrophin exon 9. Specifically, I used DUP1 fibroblasts (section 2.2.2, **table 2.1**), which carry a duplication of *DMD* exons 5-11.

As a vehicle to deliver CRISPR/Cas9 in this cell type I chose the human immunodeficiency virus type I (HIV-1)-based lentiviral particles, as these are able to transduce a broad range of cell lines, including fibroblasts^{479–481}.

For production of lentiviral particles expressing LentiCRISPR1 and LentiCRISPR2 transgenes, I used the second generation packaging system, as this method produces high functional titres⁴⁸².

I also assembled two control lentiviral particles which expressed the LentiCRISPR-EGFPsgRNA5 (Addgene 51764) and the pLJM1-EGFP (Addgene 19319) transgenes (**Fig.2.3 and Fig.2.4**, section 2.1.2.c. and 2.1.2.d).

LentiCRISPR-EGFPsgRNA5 plasmid (from now on referred as LentiCRISPR5) expresses a sgRNA that targets enhanced GFP (EGFP). As EGFP is not expressed in either HEK 293T or primary fibroblasts, the nuclease expressed by this vector is not expected to cause genome cleavage within these cell types. For this reason, I used lentiviral particles expressing LentiCRISPR5 as a negative control for the genome editing.

Conversely, pLJM1-EGFP does not express any Cas9 component but an EGFP protein which can be used to discriminate positively transduced cells and thus quantify the transduction efficiency. For this reason, I considered lentiviral particles expressing pLJM1-EGFP as a positive control for monitoring the transduction success.

These controls are summarized in **table 4.1** (below).

Control (+/-)	Transducing particles	Function
-	None	Monitor assay background
-	LentiCRISPR5	Monitor non-specific genomic cleavage
+	pLJM1-EGFP	Assess transduction efficiency

Table 4.1. List of controls included for CRISPR/Cas9 transduction in primary DMD fibroblasts.

Before generating the viral particles, I verified each of the plasmids requisite for the viral assembly (sections 2.1.2.a and 2.1.2.c-f) by restriction digest with specific enzymes (section 2.1.3) and ran the digestion products on a 2% agarose gel, later visualized in a blue-light transilluminator. The digestion of each plasmid gave rise to fragments of the expected size (section 2.1.3).

To assess the titre of my lentiviral particles I used qPCR, as indicated in section 2.1.7. . The titres, expressed as viral genome (vg)/ml, were:

- LentiCRISPR-EGFPsgRNA5 = 3.80×10^8
- pLMJ1-EGFP = 4.23×10^7
- LentiCRISPR1 = 2.13×10^8
- LentiCRISPR2 = 1.57×10^8

4.3.2. Transduction of HEK 293T cells

I tested the lentiviral particles in HEK 293T cells first. Specifically, I seeded 1×10^5 HEK 293T cells in each well of a six-well plate and transduced them with increasing doses of the lentiviral particles expressing LentiCRISPR1 and LentiCRISPR2, from MOI 1 to 100. I did this experimental step to identify the minimal load of lentiviral particles that allowed me to identify the CRISPR/Cas9-mediated cleavage, and to determine the MOI that combined the highest cleavage efficiency with the lowest cell toxicity.

The lentiviral particles expressing the control transgenes pLMJ1-EGFP (+ control) and LentiCRISPR5 (- control) were used at the intermediate transducing dose of 50 MOI.

At 48 hours post-transduction, I checked the fluorescence of HEK 293T cells transduced with lentiviral particles expressing the pLMJ1-EGFP vector. According to microscopic analysis, I estimated the transduction efficiency to be 60%. As the transduction was successful, 24 hours later I extracted genomic DNA from non-transduced cells and cells transduced with LentiCRISPR1, LentiCRISPR2 and LentiCRISPR5, and amplified it with the primers designed for the T7 assay (sequence detailed in section 2.3.4).

I performed the T7 assay on 200 ng of amplicons and run the reaction products on a 4-20% TBE gel (section 2.3.8.) (**Fig.4.2**). In samples from non-transduced cells and cells transduced with LentiCRISPR5 at MOI 50 (- controls), I only observed the band corresponding to the full-length amplicon. In contrast, in samples transduced with lentiviral particles expressing LentiCRISPR1, I observed the expected cleavage pattern at transducing dosages higher or equal to MOI 50. Transduction with lentiviral particles expressing LentiCRISPR2 (starting from MOI 20) gave cleaved fragments of the

expected size. This experiment demonstrated that my lentiviral particles can efficiently target genomic DNA in HEK 293T cells at doses equal or higher than MOI 20.

Moreover, I observed that the genomic cut efficiency increased proportionally to the increase of lentiviral particle load. Cleaved bands appeared to be more intense in samples expressing LentiCRISPR2 rather than LentiCRISPR1 and densitometric quantification confirmed this. In fact, in cells expressing LentiCRISPR1, the cleavage efficiency was 16.6% (MOI 50) and 26.4% (MOI 100), while in cells expressing LentiCRISPR2 the efficiency of cut was 20.1% at MOI 20, and increased up to 22.9% (MOI 50) and 33.7% (MOI 100).

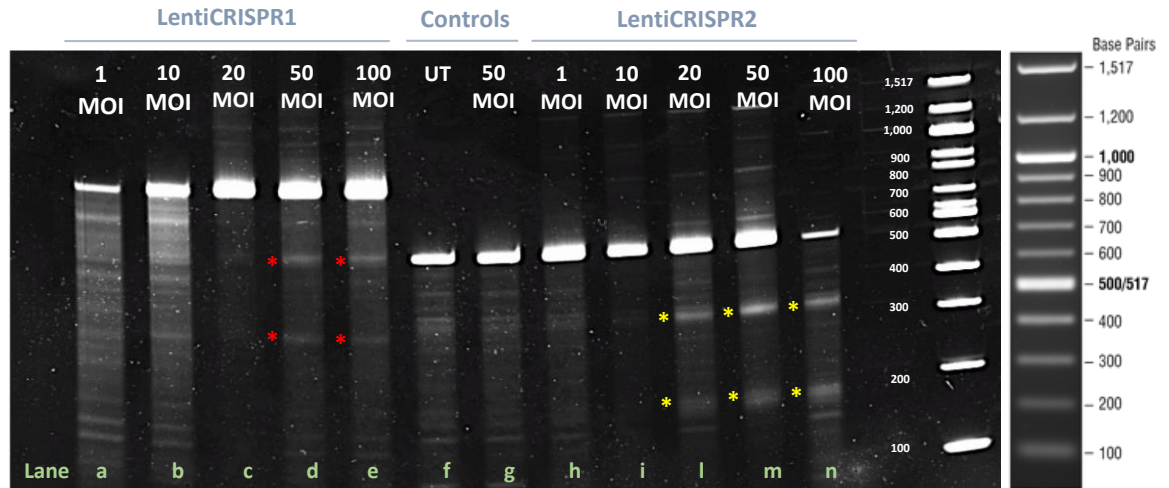


Figure 4.2. T7 assay on transduced HEK 293T cells.

T7 assay was performed on DNA deriving from HEK 293T cells transduced with increasing MOIs of LentiCRISPR1 (lanes a-e) and LentiCRISPR2 (lanes h-n) lentiviral vectors. Negative controls represented by non-transduced cells (UT, lane f) and cells transduced with MOI 50 of the LentiCRISPR5 lentiviral vector (lane g) were at this stage only amplified with the primers designed for the sgRNA2 target site. Cleaved bands of the expected size (red stars for LentiCRISPR1 and yellow stars for LentiCRISPR2 transduced cells) were only detected in HEK 293T cells transduced with a minimum MOI of 50 and 20 for LentiCRISPR1 and LentiCRISPR2, respectively. The size of LentiCRISPR1 cleaved fragments was 437 bp and 237 bp, while the size of cleaved fragments generated by LentiCRISPR2 was 281 bp and 149 bp. The 100 bp ladder from NEB was run alongside the samples.

4.3.3. Transduction of primary fibroblasts

a) Transduction of primary fibroblasts with in-house produced lentiviral vectors

Once I had confirmed that the lentiviral particles expressing LentiCRISPR1 and LentiCRISPR2 were able to target genomic DNA in HEK 293T cells, I screened their efficiency in the control primary fibroblasts DYS+ (**table 2.1**, section 2.2.2), which has a wild-type *DMD* gene. Successful cleavage in DYS+ would have allowed me to test the designed LentiCRISPRs in the patient derived DUP1 cells.

I first verified if cellular stress was already present DYS+ and DUP1 fibroblasts prior to the transduction. I did this expanding both cell lines from passage 5 until passage 10 and staining them with phalloidin. I observed stress fibers in both cell lines. In DUP1 cells at passage 5, 6 and 7 I noticed an higher number of thick stress fibers than in DYS+ (**Fig.4.3**). Stress fibers appeared to be more evident in areas where cells were confluent. I did not observe, instead, a considerable difference between DYS+ and DUP1 cells stained at passage 8, 9 and 10 (**Fig.4.3**). By virtue of this observation, I transduced cells starting from a low passage number and did not consider cells beyond passage 10. This because cells older than passage 10 would have reached senescence too quickly to verify genome editing.

I started by testing a series of increasing MOIs in the DYS+ control fibroblast cell line. I did this to identify the optimal viral load to be used for CRISPR/Cas9 expression in patient-derived fibroblasts. Initially, I screened MOIs previously tested in HEK 293T cells (section 2.2.13.a) to investigate if these doses led to a similar cleavage pattern when transducing primary fibroblasts. I performed each set of transductions twice, each being a biological replicate.

As negative controls for the genomic cleavage, I considered non-transduced cells and cells transduced with the lentiviral vector expressing LentiCRISPR5 at the highest screened MOI (100). I also transduced cells with the lentiviral particles expressing pLJM1-EGFP at the intermediate MOI 50, and estimated the transduction efficiency to be 30 %.

Three days post-transduction, I extracted genomic DNA from 5×10^4 cells transduced with LentiCRISPR vectors and non-transduced cells (section 2.3.1), but not from cells expressing pLJM1-EGFP, as these were only used to monitor the efficiency of transduction.

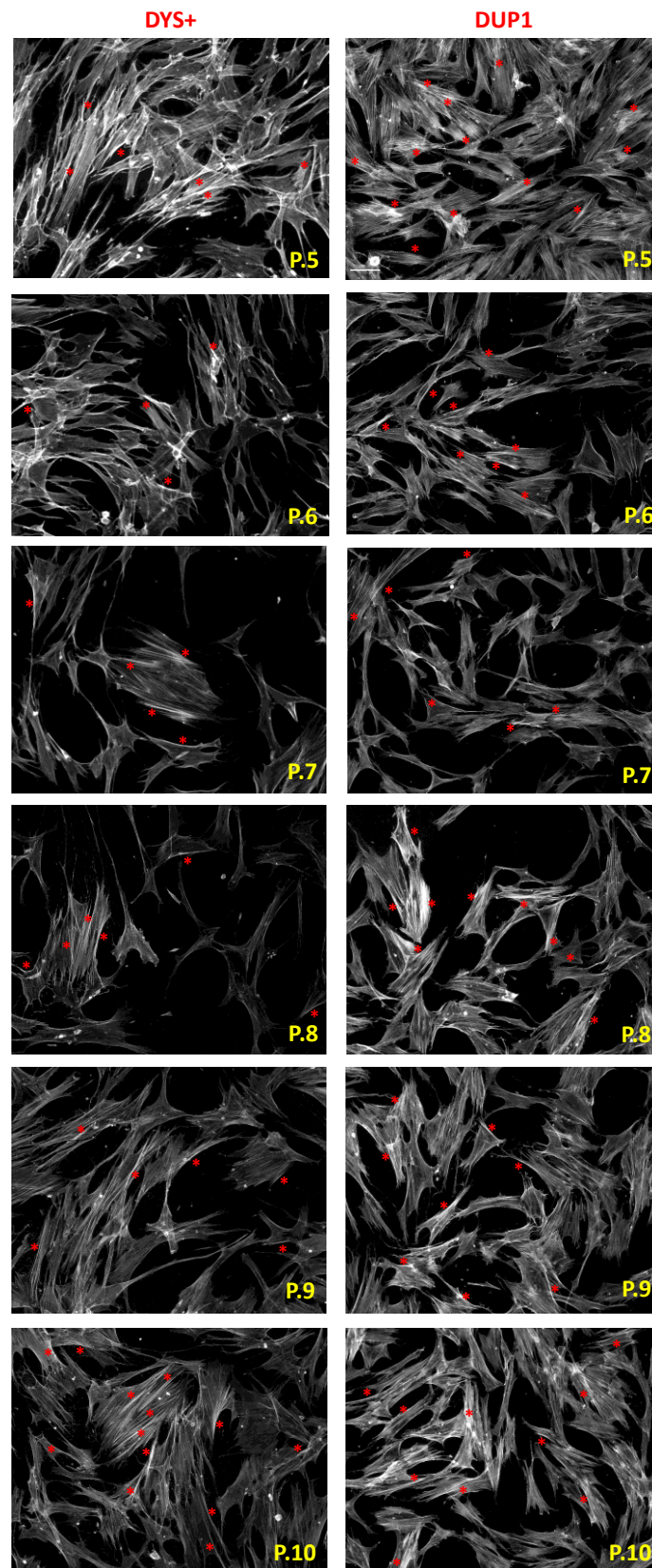


Figure 4.3. Phalloidin staining of primary fibroblasts.

DYS+ and DUP1 fibroblasts from passage 5 until passage 10 were fixed, permeabilized and stained with phalloidin. Red stars indicate the most evident stress fibers detected in both cell lines. Cells were imaged with the Olympus IX71 inverted microscope at a magnification of 20X. Scale bar = 50 μm .

I quantified DNA by NanoDrop spectrophotometer (section 2.3.2) and amplified it by tdPCR (section 2.3.4). After this, I did the T7 assay and ran the reaction products on a 4-20% TBE gel (section 2.3.7 and section 2.3.8).

Unexpectedly, I did not observe any of the cleaved fragments I previously observed in HEK 293T cells transduced with a similar viral load. Furthermore, I detected non-specific bands not only in cells transduced with the nuclease, but also in the negative control samples (i.e. non-transduced cells and cells transduced with 100 MOI of LentiCRISPR5 (**Fig.4.4.A**)).

These results suggested that the MOIs that were effective in HEK 293T cells were not sufficient to mediate genome editing in human control fibroblasts. I therefore repeated the DYS+ transduction with increasing MOIs, up to 150. In this dose range study, I also included the highest MOI used in the previous experiment (MOI 100). Negative controls were represented by non-transduced cells and cells transduced with viral particles expressing LentiCRISPR5 at MOI 150 (the highest). I also transduced the human fibroblasts with the lentiviral vector expressing pLMJ1-EGFP at the same MOI of LentiCRISPR5 (150 MOI). Also, differently from the first fibroblast transduction, I plated cells on Matrigel (section 2.2.3) to aid their survival and enhance their proliferative ability⁴⁸³. This proved to be beneficial, as Matrigel allowed a better survival rate (80% of viable cells 48 hours after transduction) compared to the previous experiment.

Two days after transduction, microscopic analysis of the cells transduced with lentiviral particles expressing pLMJ1-EGFP (MOI 150) showed EGFP expression in approximately 70% of fibroblasts.

I performed DNA extraction, amplification and T7 assays on non-transduced and samples transduced with LentiCRISPR1, LentiCRISPR2 and LentiCRISPR5 (section 2.3.2 and 2.3.7). However, I detected only background nonspecific bands in T7 assay of samples expressing either LentiCRISPR1 or LentiCRISPR2. These background bands were also detected in the negative controls (i.e. untreated cells and cells transduced with LentiCRISPR5 at MOI 150) (**Fig.4.4.B**). This negative result indicated that, for the genomic cleavage to occur, CRISPR/Cas9 nuclease transduction in human fibroblasts needed to be optimized further.

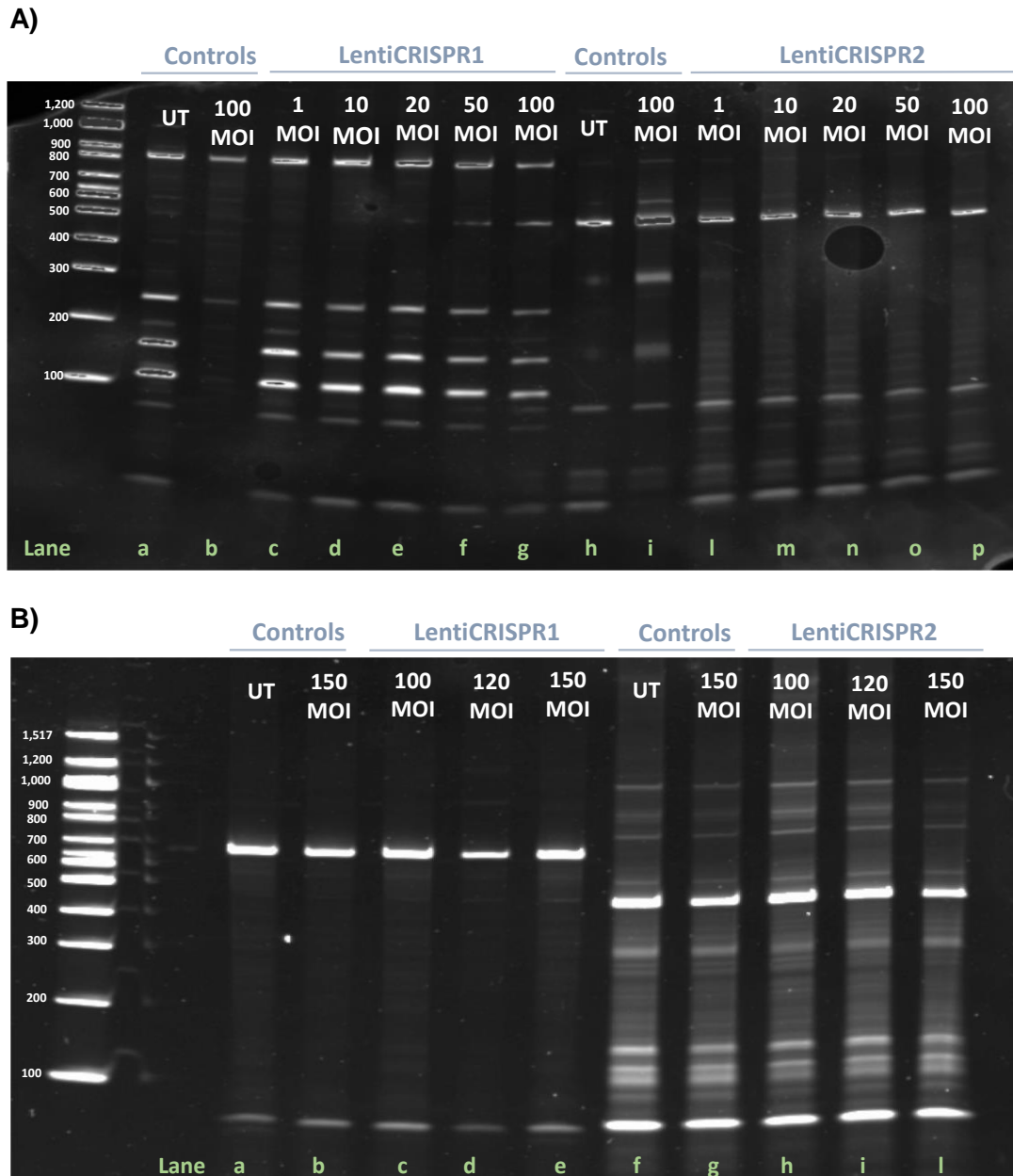


Figure 4.4. *T7 assay on transduced human-derived fibroblasts.*

A) T7 assay was performed on DNA deriving from control fibroblasts cells transduced with the MOIs of LentiCRISPR1 (lanes c-g) and LentiCRISPR2 (lanes l-p) lentiviral vectors previously screened in HEK 293T cells. Negative controls were non-transduced cells (UT, lanes a and h) and cells transduced with MOI 100 of the LentiCRISPR5 lentiviral vector (lanes b and i). None of the expected cleaved bands was detected in this cell type, even at 100 MOIs. B) T7 assay performed on DNA deriving from control fibroblasts transduced with lentiviral vectors expressing LentiCRISPR1 (lanes c-e) and LentiCRISPR2 (lanes h-l). Transducing MOI values were increased up to 150. Negative controls were non-transduced cells (UT, lanes a and f) and cells transduced with MOI 150 of the LentiCRISPR5 lentiviral vector, which targets EGFP (not expressed in human cells) (lanes b and g). Only non-specific bands were detected, indicating the need of either higher viral loads or of having a selection of targeted cells only. The 100 bp ladder from NEB was run alongside the samples.

b) Transduction of primary fibroblasts with commercial lentiviral particles

At this point, I decided to include a selection method to purify the transduced cells. I chose to use a vector that could express both CRISPR/Cas9 and a fluorescence marker, so that I could detect fluorescent cells (expressing the nuclease) at 48 hours post transduction.

For this purpose, I purchased lentiviral particles which co-expressed sgRNA2 (the most efficient sgRNA in HEK 293T cells), Cas9 and EGFP (**Fig.2.7**), to provide a means to select transduced cells by either fluorescence microscopy or fluorescence activated cell sorting (FACS). In the transgene expressed by these lentiviral particles (LentiDMD2), the co-expression of Cas9 and EGFP was achieved via a T2A linker, one of the self-cleaving peptides used to allow bicistronic transgene expression in eukaryotic cells⁴⁸⁴.

As the titre of lentiviral particles provided by the company was higher (10^9) than that obtained for the viral particles I produced (10^7 - 10^8), this allowed me to transduce DYS+ fibroblasts with considerably high dosages (up to 1000 MOI). In particular, I tested three doses: MOI 150, MOI 500 and MOI 1000 (**Fig.4.5**). The only negative control I included in this set of experiment was non-transduced cells. I did not include the positive control previously used to monitor the transduction efficiency (i.e. lentiviral particles expressing pLJM1-EGFP). This was because the new lentiviral particles already express EGFP, and should be therefore indicative of the transduction efficiency. For each MOI, transduction was performed twice (technical replicate). I assessed cell fluorescence by fluorescent microscope analysis at both 48 and 72 hours post-transduction. However, unexpectedly, at neither of these time-points could I detect a fluorescent signal distinguishable from the background that I also observed in non-transduced cells (negative control for transduction) (**Fig.4.5**). I repeated the transduction in culture medium supplemented with 8 µg/mL of the cationic polymer polybrene, as it acts as enhancer of the transduction efficiency by attenuating the charge repulsion occurring among the viral particles and the sialic acid expressed on the cell surface⁴⁸⁵. However, the outcome did not change (**Fig.4.5**).

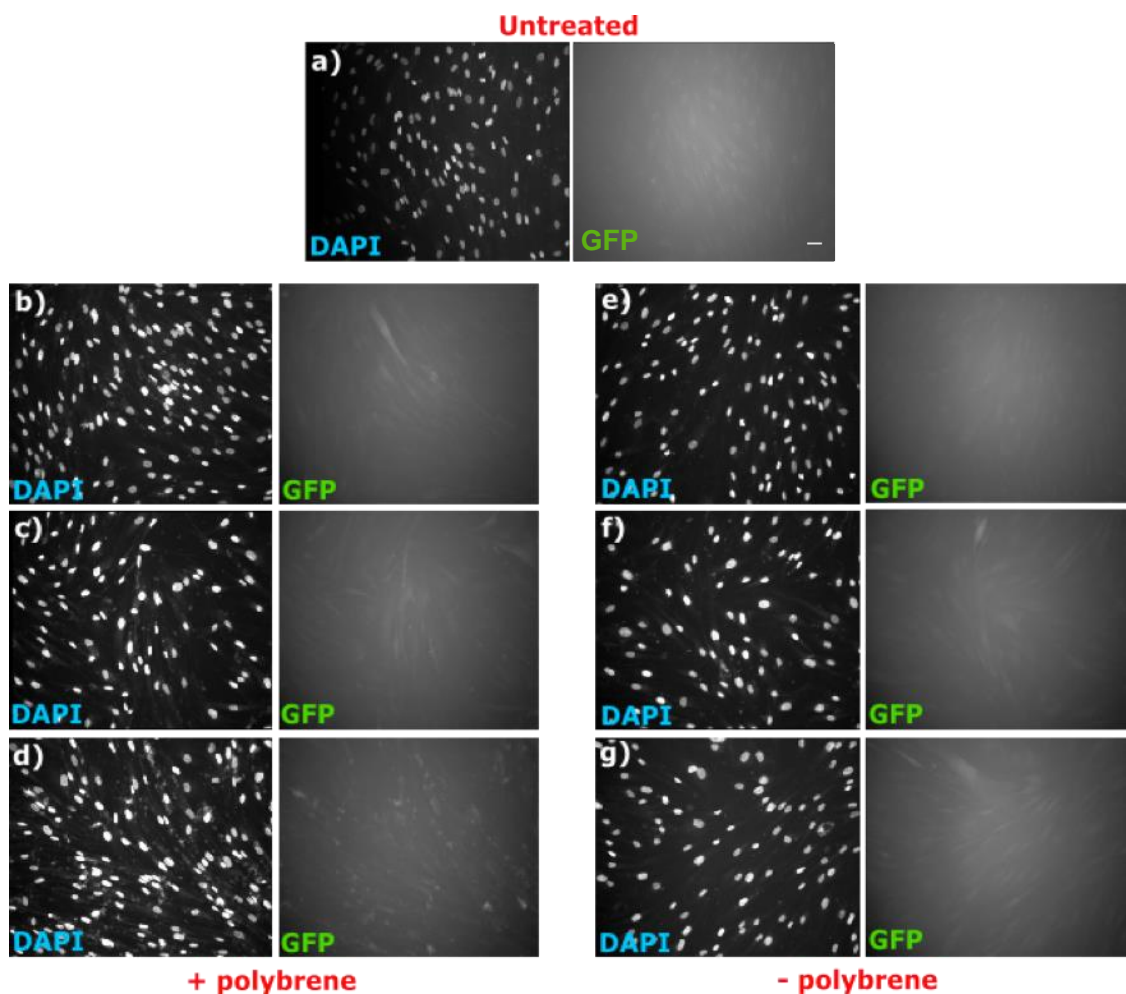


Figure 4.5. Transduction of LentiDMD2 in patient fibroblasts.

DYS+ fibroblasts were transduced with increasing LentiDMD2 doses represented by MOIs 150 (b and e), 500 (c and f) and 1000 (d and g). Transduction was performed both in the presence (b-d) and absence (e-g) of polybrene, a transduction enhancer. As a negative control, untreated cells were used (a). After 48 hours, cell nuclei were stained with DAPI and imaged with the Olympus IX71 inverted microscope. The analysis of GFP fluorescence was inconclusive, as in all samples GFP expression was only background, indistinguishable from the non-transduced cells. Scale bar = 100 μ m.

To investigate the reason behind such poor transduction efficiency, I transduced DYS+ fibroblasts in parallel with either lentiviral particles expressing LentiDMD2 or those expressing pLMJ1-EGFP, produced in-house (section 2.1.6). I used lentiviral particles expressing pLMJ1-EGFP as a positive control for monitoring the transduction efficiency, as these were able to transduce fibroblasts (as I observed from the previous experiments). Non-transduced cells were also included as a negative control. As expected, there were fluorescent cells following transduction with the lentiviral particles expressing pLMJ1-EGFP, even at the lowest viral dose tested (MOI 150) (**Fig.4.6.A**). In contrast, I did not observe fluorescent cells following transduction with lentiviral particles expressing LentiDMD2, even at the highest dose used (MOI 1000) (**Fig.4.6.B**). Taken together, these experiments led me to conclude that LentiDMD2 was ineffective.

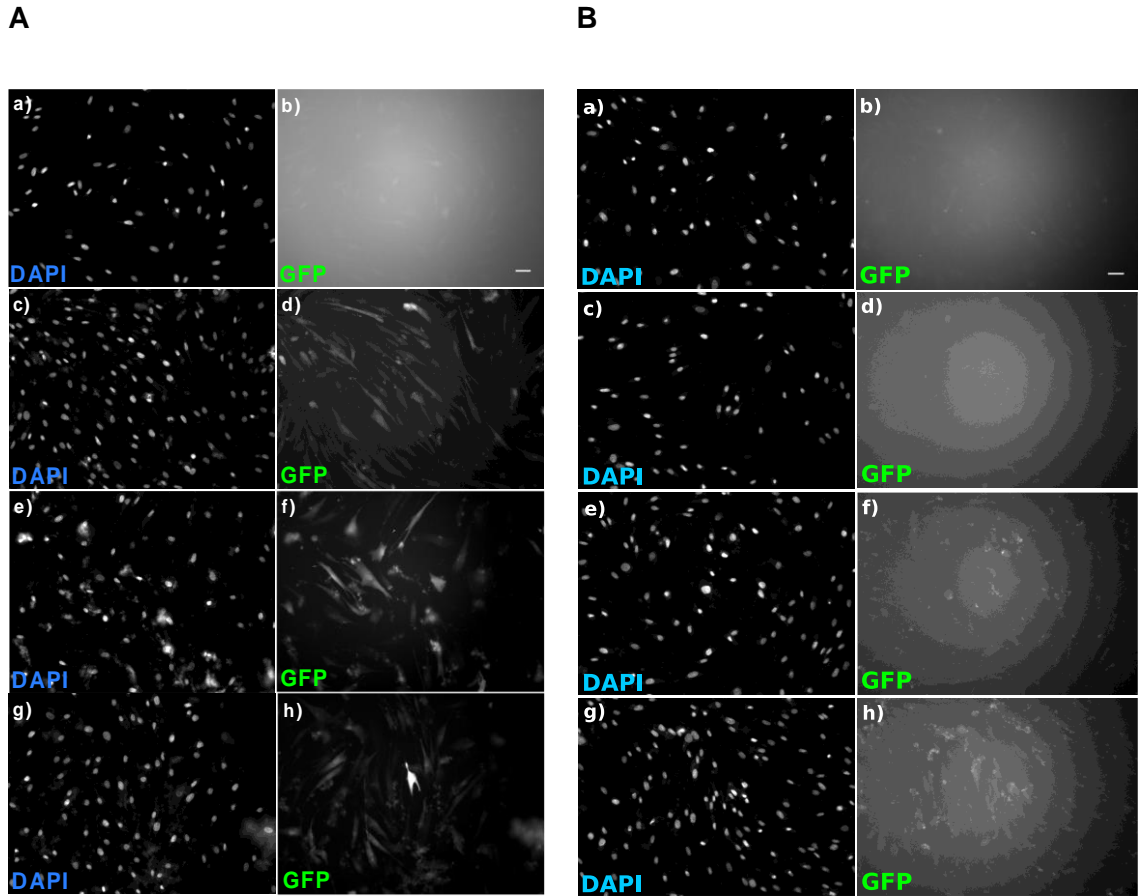


Figure 4.6. Comparison of transduction efficiency between lentiviral particles expressing LentiDMD2 and pLMJ1-EGFP.

Cells were transduced with increasing loads of viral particles expressing either LentiDMD2 (commercial) or pLMJ1-EGFP (produced in-house) transgenes. Untreated cells are represented in panels a) and b) of both A and B panels. A): Transduction with lentiviral particles expressing pLMJ1-EGFP. B): Transduction with lentiviral particles expressing LentiDMD2. c)-d): DAPI and GFP expression in cells transduced with MOI 150; e)-f): DAPI and GFP expression in cells transduced with MOI 500; g)-h): DAPI and GFP expression in cells transduced with MOI 1000. GFP expression was detected in cells transduced with lentiviral particles expressing pLMJ1-EGFP at MOI 150, but not with any of the MOIs tested of lentiviral particles expressing LentiDMD2. Scale bar = 100 μ m.

4.3.4. Nuclear electroporation of primary fibroblasts

In order to overcome the low efficiency observed with the lentiviral vectors, I explored non-viral methods as an alternative approach to deliver CRISPR/Cas9 to primary fibroblasts.

Among the non-viral methods available to express CRISPR/Cas9 in human fibroblasts, I chose nuclear electroporation (nucleofection) through a device developed by the company Amaxa. This was because nucleofection has been previously shown to achieve transfection in cell lines that are resistant to transfection with standard transfection reagents, by using electroporation parameters and reagents standardized for each cell type^{486–488}.

Electroporation efficiency is thought to reduce significantly with increasing DNA sizes⁴⁸⁹. Therefore, as lentiviral plasmids are more than 10 kb in size, I had to consider a different plasmid for CRISPR/Cas9 expression. For this reason, I purchased two non-viral CRISPR/Cas9 plasmids which shared the same backbone but expressed two different sgRNAs. These vectors (CRISPR2 σ and CRISPR0) had a DNA length of around 8.3 kb and could co-express the sgRNA as well as Cas9 and EGFP fused via a 2A sequence, consistent with the LentiDMD2 vector (**Fig.2.8**).

The advantage of these vectors is that they allow the visual identification of cells expressing EGFP (and consequently Cas9) and enable EGFP-selection of cells via FACS. In this way, cells expressing the nuclease can be isolated and expanded for the further assessment of Cas9 cutting efficiency. Similarly to LentiDMD2, sgRNA expression was driven by the U6 promoter but, instead of the tEF1 σ promoter, Cas9-EGFP expression was driven by the constitutive CMV promoter. CRISPR2 σ expresses sgRNA2, while the sgRNA found in CRISPR0 recognizes a sequence not present in the human genome and therefore, together with untreated cells, was used as a negative control.

The controls included in this set of experiments are summarized in **table 4.2** (below).

Control (+/-)	Plasmid to electroporate	Function
-	None	Monitor assay background
-	LentiCRISPR5	Monitor non-specific genomic cleavage
+	pLJM1-EGFP	Assess transfection efficiency

Table 4.2. List of controls included for CRISPR/Cas9 electroporation in primary DMD fibroblasts.

I grew both plasmids and verified them by restriction digest as detailed in sections 2.1.1.c-e and 2.1.3, respectively. The products of the digestion reaction were run on a 2% agarose gel, where I detected fragments of the expected size by means of a blue-light transilluminator (section 2.3.5).

a) Validation of CRISPR2 σ and CRISPR0 in HEK 293T cells via Lipofectamine transfection

I transfected HEK 293T cells with the above plasmids, keeping the previously optimized plasmid DNA:Lipofectamine 2000 ratio constant (section 2.2.11.a.). In parallel, I also transfected HEK 293T cells with the previously tested LentiCRISPR2 and pLMJ1-EGFP vectors. The inclusion of LentiCRISPR2 plasmid allowed me to ensure that CRISPR2 σ plasmid was functional and, when expressed in HEK 293T cells, it could lead to a cleavage pattern similar to that previously obtained with LentiCRISPR2. pLMJ1-EGFP plasmid served as a positive control to monitor the efficiency of transfection. As expected, 48 hours post-transfection, I detected fluorescent cells in cells transfected with pLMJ1-EGFP, CRISPR0 and CRISPR2 σ (**Fig. 4.7**).

The efficiency of transfection (calculated as the ratio between the number of fluorescent cells and the total number of cells in the chosen field) was similar for all the three plasmids (32.0% +/- 1.09). Three days later I extracted the genomic DNA of non-transfected cells and cells transfected with the CRISPR/Cas9-expressing plasmids and I performed DNA amplification. From now on, I used a different primer couple designed to amplify a larger fragment (728 bp) of the sgRNA target region (sections 2.3.1 and 2.3.4).

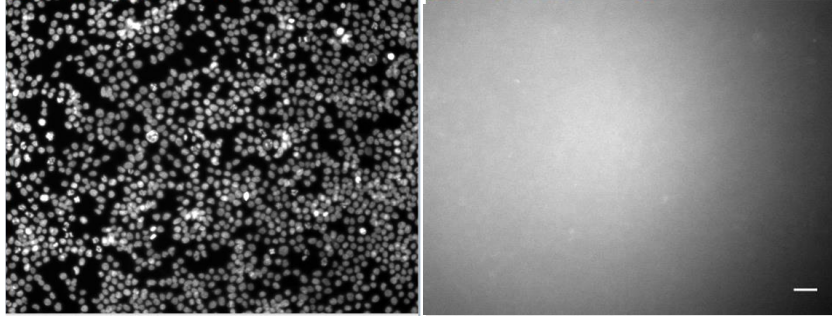
The sequence of these new primers is indicated below:

CRISPR2 New-FW: 5' TGGATGCCAAAACCTACAGT 3'

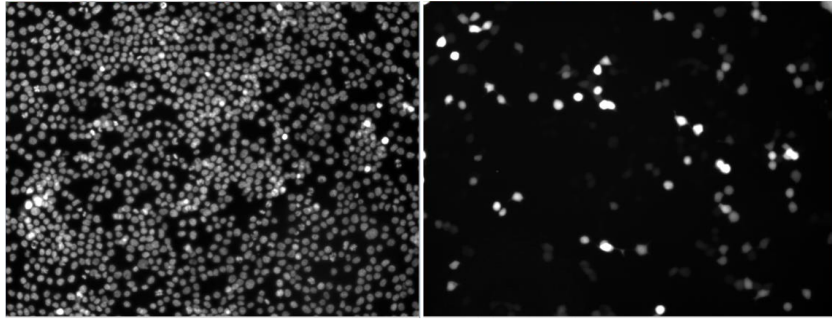
CRISPR2 New-REV: 5' GTACGGCAGCTCTTCAAAAA 3'

I performed the T7 assay on these amplicons and run the products on a 4-20% TBE gel (section 2.3.7 and 2.3.8). This allowed me to confirm that CRISPR0 had no unspecific cleavage activity, while CRISPR2 σ had the same pattern of cutting as LentiCRISPR2. Moreover, LentiCRISPR2 and CRISPR2 σ had a similar efficiency (35.2% for CRISPR2 σ versus 29.9% for LentiCRISPR2) (**Fig.4.8**).

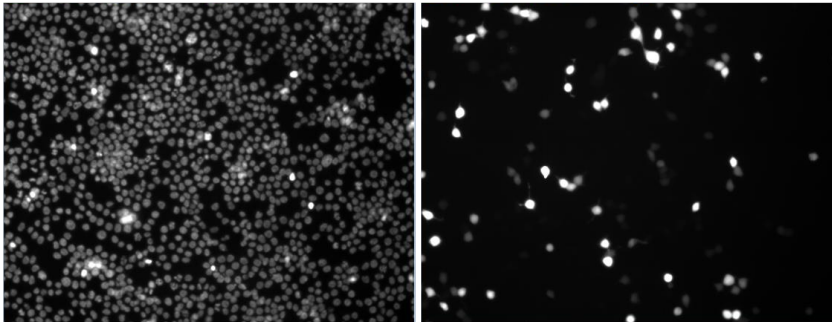
UNTRANSFECTED:



pLMJ1-EGFP vector:



CRISPR0:



CRISPR2 σ :

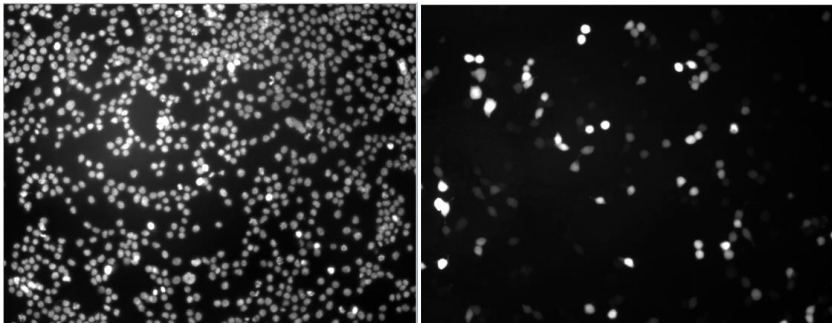


Figure 4.7. GFP fluorescence monitored on transfected HEK 293T cells.

HEK 293T cells were transfected via Lipofectamine 2000 with either pLMJ1-EGFP, CRISPR0 or CRISPR2 σ plasmids. GFP fluorescence was confirmed 48 hours post transfection. Images were captured by the Olympus IX inverted fluorescent microscope at a magnification of 5X. Scale bar = 200 μ m.

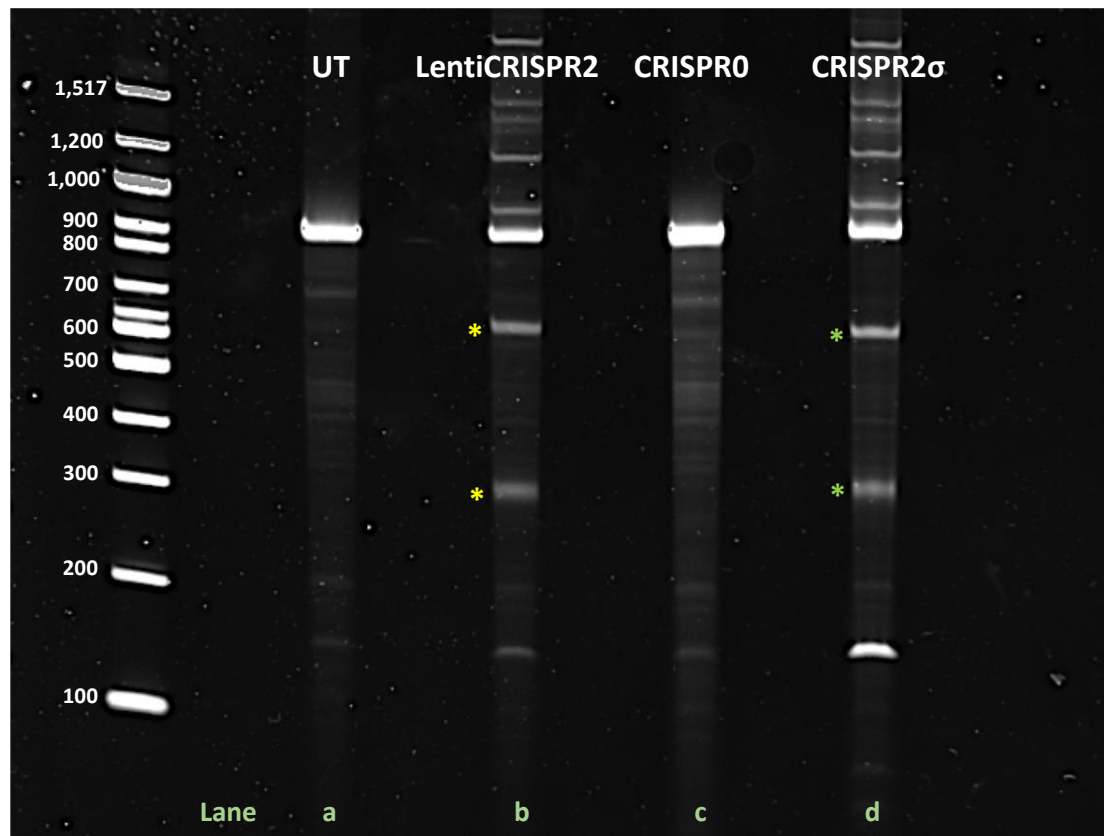


Figure 4.8. T7 assay performed on HEK 293T cells transfected with the pCMV-Cas9-GFP vectors.

CRISPR0 and CRISPR2 σ vectors were transfected into HEK 293T cells. As expected, similar to the untreated cells also used as negative control for genomic cleavage (lane a), CRISPR0 vector did not elicit any specific cleaved bands, thus confirming its inability to cut the target site (lane c). Cleaved fragments of the expected size (511 bp and 217 bp) were detected upon transfection of both CRISPR2 σ (green stars in lane d) and LentiCRISPR2 plasmid (yellow stars in lane b). A 100 bp ladder (NEB) was run alongside the samples.

b) Nucleoporation of CRISPR2 σ and CRISPR0 in primary fibroblasts

The electroporation program recommended by Amaxa for primary fibroblasts was B032. I tested this and two other programs (U024 and P022) which, according to Amaxa, are associated with the highest efficiency of delivery and with the highest viability, respectively. This step allowed me to identify the best electroporation condition for the delivery of CRISPR0 and CRISPR2 σ in primary cells. The DNA amount (2.5 μ g) remained constant across the electroporation conditions. At first, I electroporated DYS+ cells with the control CRISPR0 plasmid only and, 48 hours after, I isolated GFP-positive cells via FACS (sections 2.2.12.b and 2.2.14.a) (**Fig.4.9**). The experiment was repeated twice (technical replicate). Among the three electroporation programs, both B032 and P022 gave an high percentage of EGFP-positive cells (32.7% \pm 0.11 and 23.3% \pm 0.92 respectively), while U024 resulted in only 8% \pm 1.78 of EGFP-positive cells. I chose P022 as, compared to B032, it resulted in a limited cell death (<30%).

Using these optimized nucleofection conditions, I electroporated 2.5 μ g of CRISPR0 and CRISPR2 σ plasmids DNA into 5×10^5 DUP1 patient-derived fibroblasts (**table 2.1**, section 2.2.2). The experiment was performed three times, each being a biological replicate. Non-treated cells were included as a negative control. 48 hours post-electroporation, the percentage of EGFP-positive cells observed by fluorescence microscopy (section 2.3.23) accounted for about 60% of the total cell population. As the majority of cells expressed the nuclease, I collected DNA from the total cell pool (without isolating EGFP positive cells), amplified it and performed the T7 assay (sections 2.3.1 and 2.3.4). However, when I run the products of the T7 assay on the 4-20% TBE gel, I observed only the upper full-length amplicon band. This suggested that the inclusion of the untargeted population might have overcome the edited cells, thus masking the nuclease effect and leading to the cleaved fragments being missed (**Fig.4.10**).

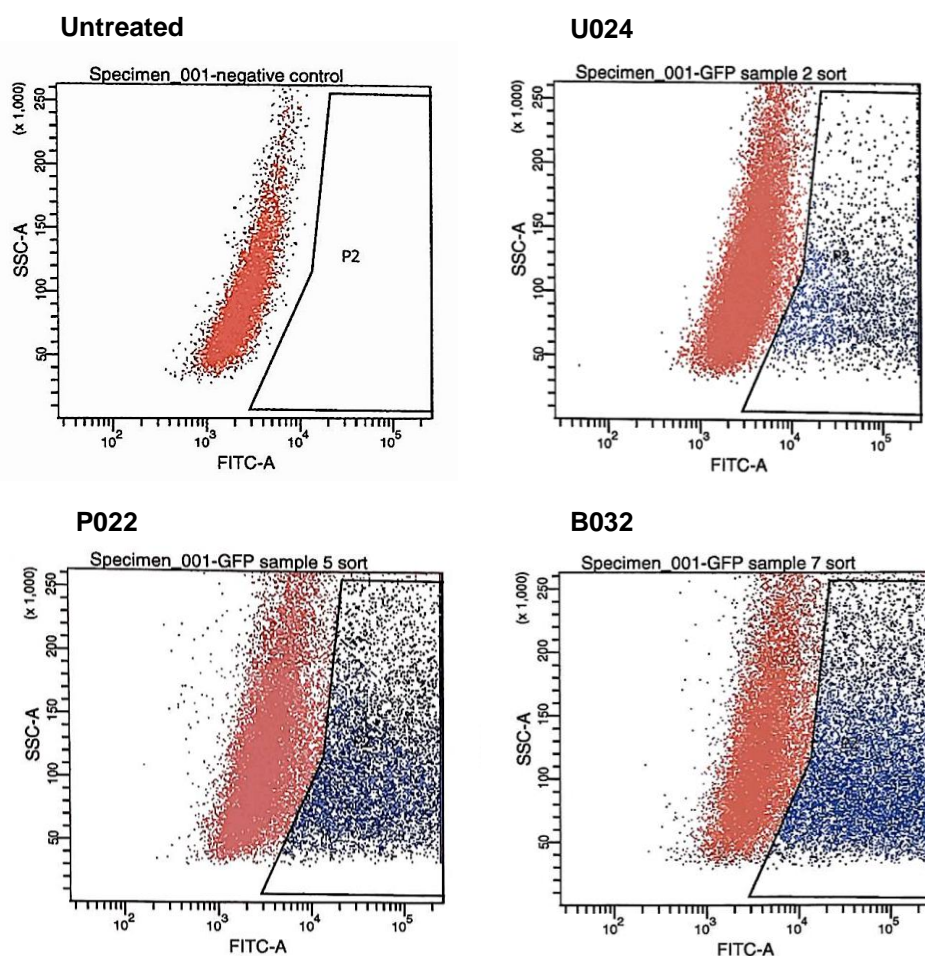


Figure 4.9. Screening of electroporation programs in primary control fibroblasts.

DYS+ fibroblasts were electroporated with the CRISPR0 plasmid by means of electroporation programs U024 (top right), P022 (bottom left) and B032 (bottom right). Since CRISPR0 expressed EGFP, the extent of fluorescence correlated with electroporation efficiency. B032 had the highest efficiency, but P022 resulted in more viable cells.

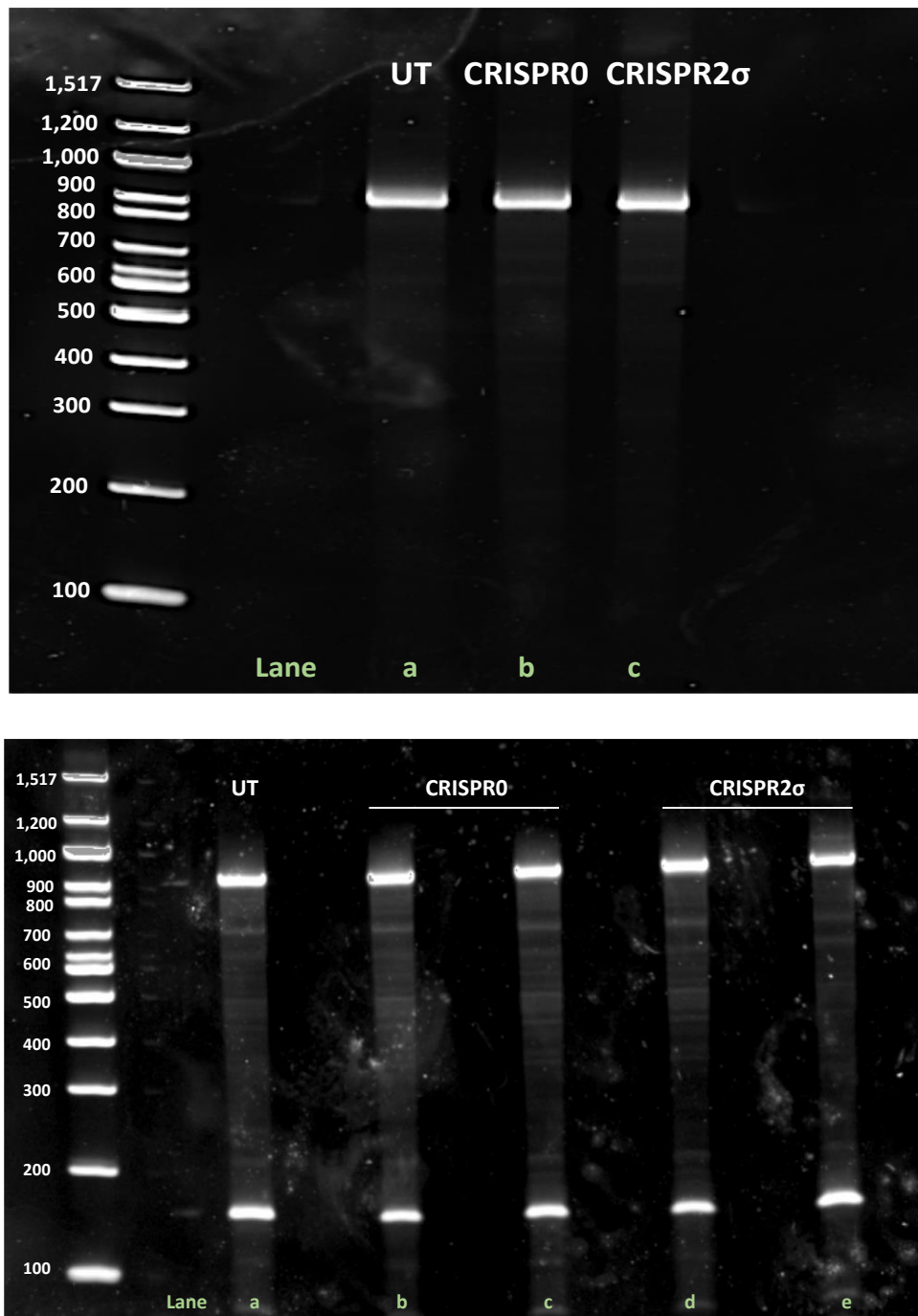


Figure 4.10. T7 assay performed on the total population of patient-derived fibroblasts electroporated with the pCMV-Cas9-GFP vectors.

CRISPR0 and CRISPR2σ vectors were electroporated into patient-derived cells by using the electroporation program P022 of the Amaxa device. T7 was performed on DNA derived from the total cell population, without sorting for GFP-positive fibroblasts. In both samples electroporated with CRISPR/Cas9-expressing plasmids and untreated cells (UT), I only detected the full length-band of 728 bp. A 100 bp ladder (NEB) was run alongside the samples.

To test this hypothesis, DUP1 electroporation was repeated three times (biological replicates). However, instead of collecting the total cell population, at 48 hours post-nucleofection, I isolated EGFP-positive cells (26.02% \pm 0.5) via FACS sorting (section 2.2.14.a). This allowed me to examine only the cells expressing the nuclease while discarding the non-targeted ones. I expanded the sorted cells to about 3×10^5 in one well of a six-well plate and their DNA was extracted and amplified for T7 assay (sections 2.3.1 and 2.3.4-2.3.6.). Even here, the T7 assay outcome was similar to when the whole cell population was considered, as I observed only the molecular band representative of the entire amplicon (**Fig.4.11**).

This led me to assume that Cas9 targeting in human fibroblasts might be dependent on the nuclease load to each cell. To test this, I performed a further nuclear electroporation in DUP1 fibroblasts in triplicate with subsequent sorting of the GFP-positive cells (section 2.2.12.a and 2.2.14.a). However, I collected only the cells with the highest GFP expression (top 20%), which appeared as the brightest. I expanded the sorted cells and extracted their DNA, which was amplified and used for the T7 assay (section 2.3.1 and 2.3.4-2.3.6). This time, running the T7 reaction products on a TBE gel gave cleaved bands of the expected size, whilst in cells electroporated with CRISPR0 and in non-treated cells (i.e. negative controls), there was only the band corresponding to the non-cleaved amplicon (**Fig.4.12**). Quantitative analysis based on band densitometry revealed a mean cleavage efficiency of 28% \pm 1.23 across all triplicates.

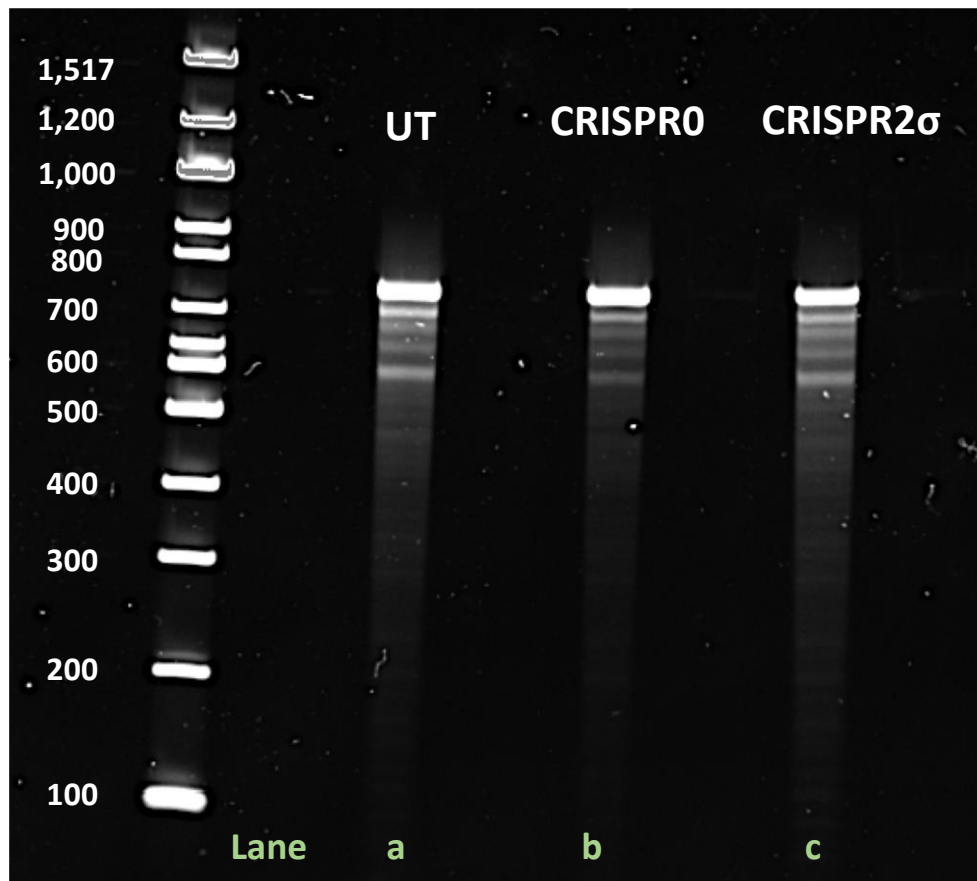


Figure 4.11. T7 assay performed on the GFP-positive patient-derived fibroblasts electroporated with the pCMV-Cas9-GFP vectors.

CRISPR0 and CRISPR2 σ vectors were electroporated into patient-derived cells by using the electroporation program P022 of the Amaxa device. T7 was performed on DNA derived from the total GFP-positive population. Other than the 728 bp amplicon band observed in untreated cells and cells electroporated with CRISPR0 (lanes a and b), only non-specific bands appeared in GFP-positive cells expressing CRISPR2 σ (lane c). A 100 bp ladder (NEB) was run alongside the samples.

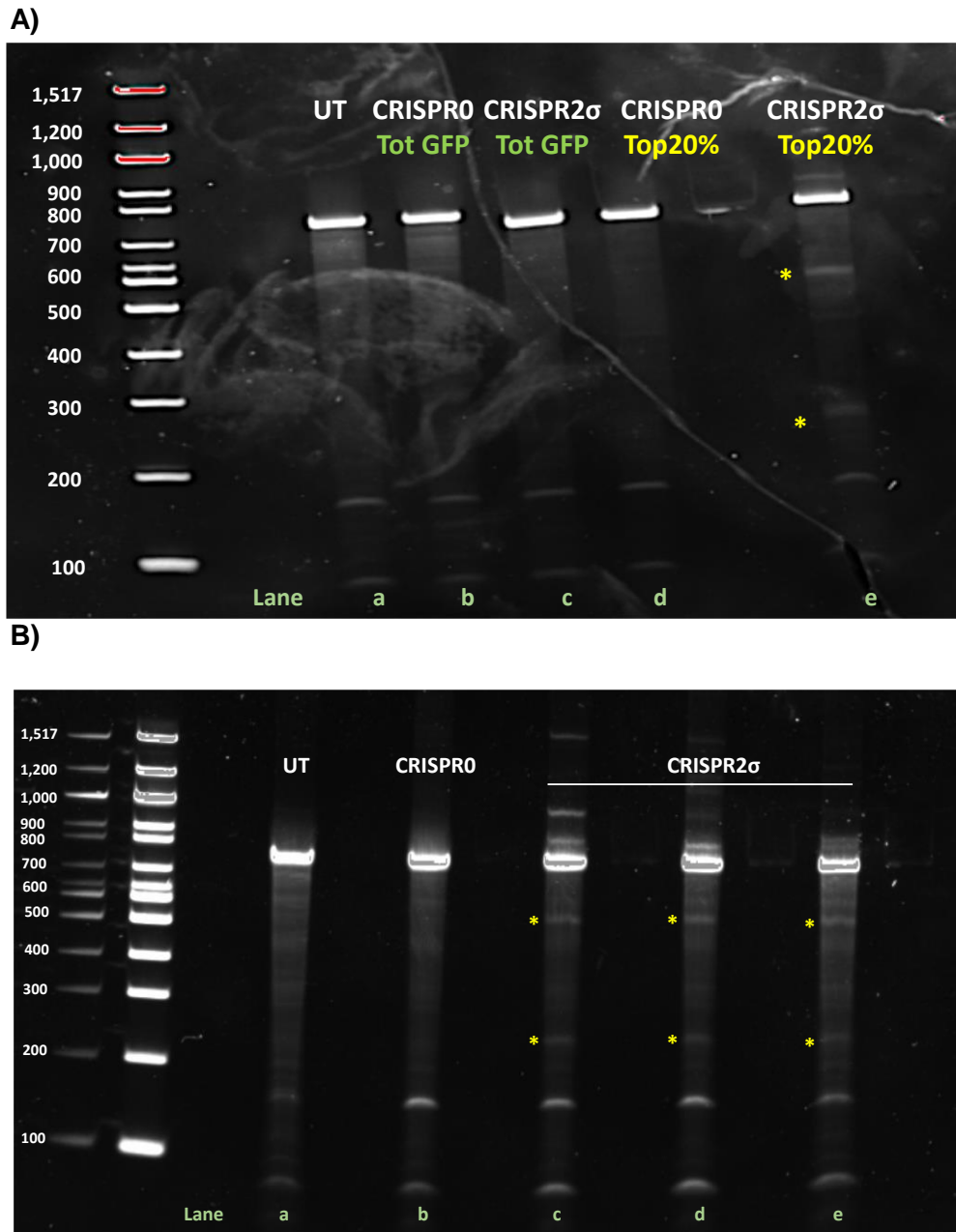


Figure 4.12. T7 assay performed on the Top 20% GFP-positive patient-derived fibroblasts electroporated with the pCMV-Cas9-GFP vectors.

CRISPR0 and CRISPR2 σ vectors were electroporated into DUP1 patient-derived cells by using the electroporation program P022 of the Amaxa device. T7 was performed on DNA derived from the brightest GFP-positive population (Top 20%). A) Cleaved band of the expected size of 511 bp and 217 bp (yellow stars) appeared only in DNA extracted from the brightest cells pool (lane e) but not in the total GFP population expressing CRISPR2 σ (lane c), and was absent in the negative controls (untreated fibroblasts and cells electroporated with CRISPR0) (lanes a, b and d). B) T7 assay repeated on the top 20% GFP-population confirmed the cleavage in CRISPR2 σ expressing cells (lanes c-e) only. A 100 bp ladder (NEB) was run alongside the samples.

4.3.5. MyoD transduction

Since dystrophin is poorly expressed in fibroblasts⁴⁶⁸, I could not detect if the genomic targeting observed in DUP1 cells via the T7 assay led to the correction of dystrophin transcript and protein. Therefore, I had to differentiate DUP1 fibroblasts into myoblasts and eventually fuse them into myotubes, as only myotubes express dystrophin⁴⁹⁰. To induce myogenesis, I transduced DUP1 fibroblasts with an adenoviral vector expressing the MyoD transcription factor⁴⁹¹.

The adenoviral vectors of serotype 5 Ad5.f50.AdApt.MyoD and Ad5-CMV-MyoD (Sirion Biotech) have been previously used to deliver MyoD into human fibroblasts⁴⁹². However, their efficiency considerably varied across the human fibroblasts cell lines in which they were tested and was generally poor. This may have been because non-modified adenoviruses belonging to the serotype type 5 prevalently infect target cells by exploiting the coxsackievirus-adenovirus receptor (CAR), which is expressed at very low levels in human fibroblasts⁴⁹³. Also, these vectors did not carry any selection cassette (either antibiotic or fluorescent marker) so transduced cells could not be selected. To overcome these issues I selected a third adenoviral vector expressing MyoD: Ad(RGD)-GFP-h-MyoD1 (Vector BioLabs). This vector uses its Arginine-Glycine-Asparagine motif (RGD motif) to bind α_v integrins, which are expressed on fibroblast surface⁴⁹⁴. Through this mechanism, they can bypass the entry via CAR receptors. Moreover, the vector concomitantly expresses both MyoD and GFP, allowing visual monitoring of the transduction efficiency by the presence of GFP-expressing cells. As the expression of both MyoD and GFP cDNA is driven by a CMV promoter, I inferred that a positive fluorescence signal in transduced cells would have also been indicative of MyoD expression.

Initially, I tested a range of viral doses (50-100-200 MOIs) to identify the one that gave the highest transduction efficiency (**Fig.4.13**) (section 2.2.13.b.). As recently reported³⁸⁹, around 80% of target cells were expressing GFP 24 hours post-transduction with an MOI of 100, suggesting this was the optimal viral dose. Finally, as myotubes were required to evaluate dystrophin transcript and protein, I cultured transduced cells with DMEM culture medium supplemented with 2% horse serum for 9 days (section 2.2.13.b) to induce terminal differentiation and obtain myotubes.

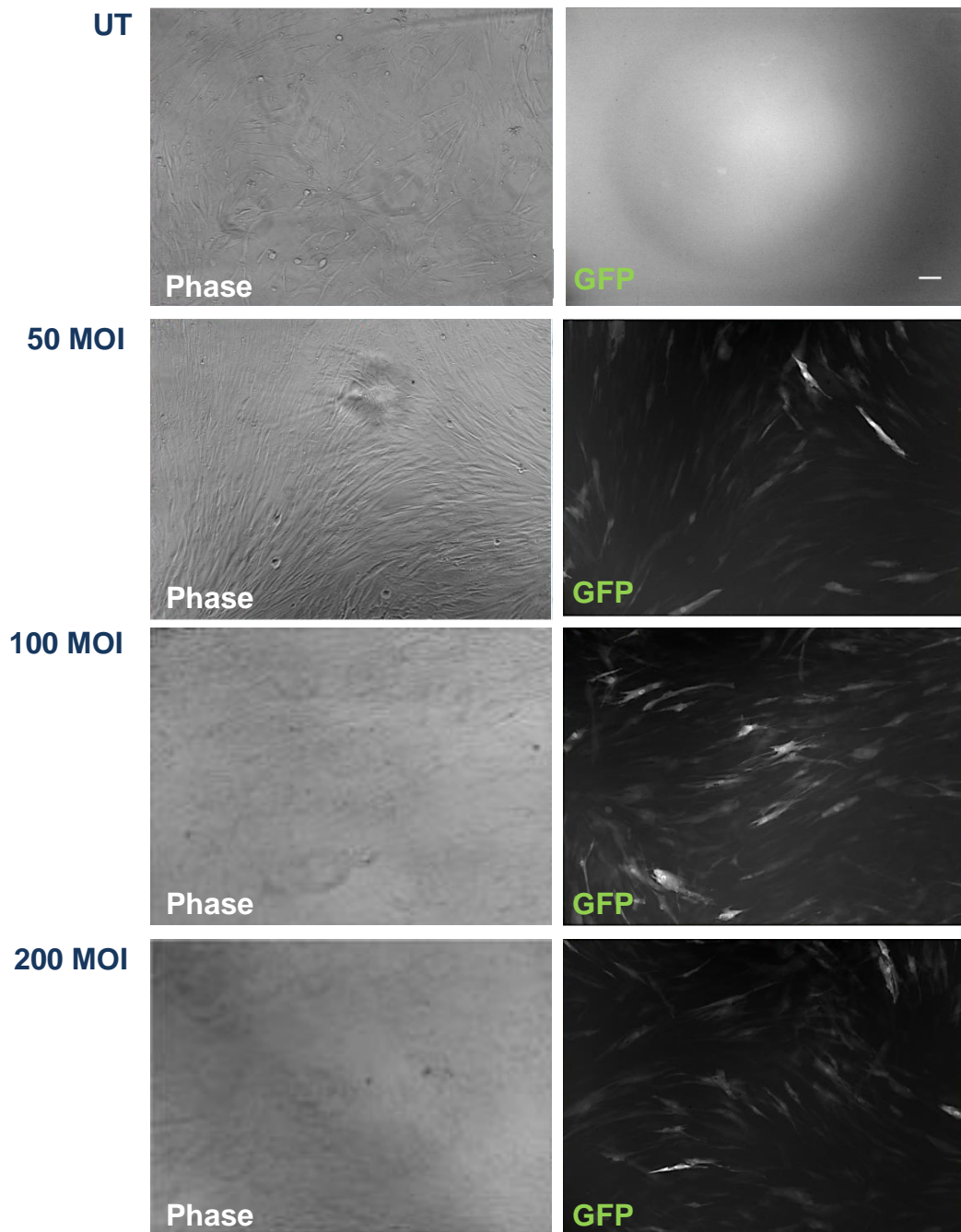


Figure 4.13. Fibroblast transduction with Ad(RGD)-GFP-h-MyoD1 viral vector.

Patient fibroblasts were transduced with increasing doses of the adenoviral vector expressing the MyoD transcription factor. MOIs tested were 50, 100 and 200. Transduction efficiency was calculated by counting the number of GFP+ cells and estimated as 55% \pm 1.23 (MOI 50), 78% \pm 2.30 (MOI 100) and 69% \pm 1.58 (MOI 200). The best transduction efficiency was obtained with MOI 100, used for subsequent experiments. Fluorescence was assessed with the Olympus IX71 Inverted microscope at a magnification of 10X. Untreated cells (UT) were used as a negative control. Scale bar = 100 μ m.

Despite maintaining the viral load suggested by Wojtal *et al.*³⁸⁹, the percentage of nuclei in myotubes or “fusion index” was less than 10%. This was unexpected, giving the high percentage of GFP-positive cells observed at 48 hours post-transduction (around 80%) (**Fig.4.13**). I therefore hypothesized such that poor differentiation efficiency was a consequence of the medium used for the differentiation. To verify this, I repeated the transduction and tested two other differentiation media. One was composed of DMEM supplemented with 1% FBS and 1% GlutaMax. The other was DMEM without serum but supplemented with 1% GlutaMax (section 2.2.13.b) (**Fig.4.14**). The transduction and testing of these differentiation conditions was repeated three times (technical replicate). I fixed and permeabilized cells as detailed in the section 2.3.15. Desmin immunostaining was performed to identify myogenic cells (section 2.3.17.a) but, even using these conditions, I detected only rare myotubes (fusion index less than 15%) in cells cultured with 2% horse serum or 1% FBS (**Fig.4.15**). These results highlighted the need to identifying reproducible conditions for this process. Further optimization is therefore required.

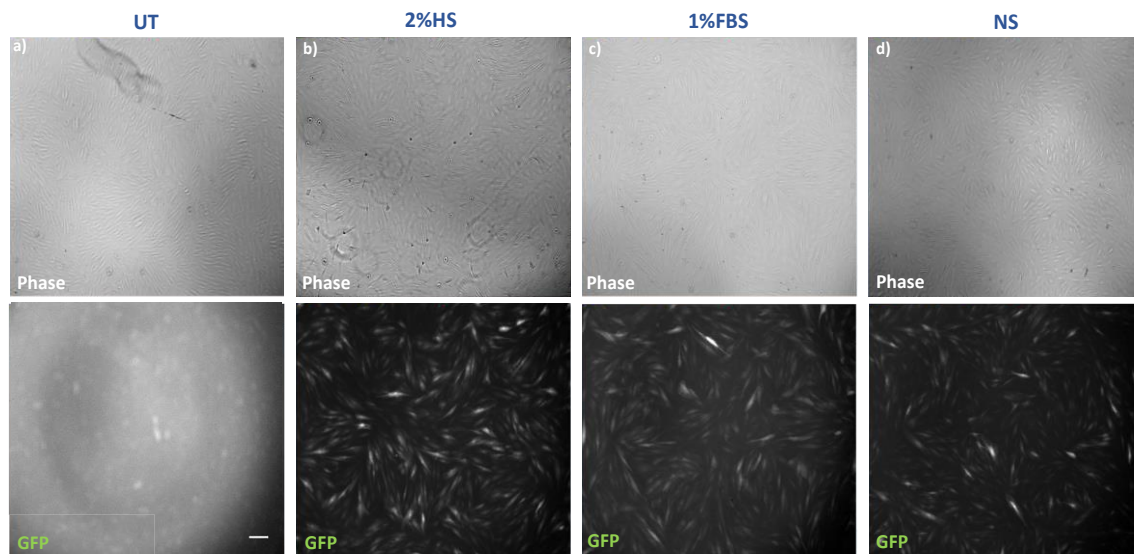


Figure 4.14. Three differentiation conditions tested in transduced fibroblasts (MOI 100).

Successfully transduced cells expressing MyoD were differentiated by using three conditions: 2% horse serum (HS, a), 1% fetal bovine serum (FBS, b), no serum (NS, c). Fluorescence was assessed with the Olympus IX71 Inverted microscope at magnification 5X. Untreated cells were used as a negative control. Scale bar = 500 μ m.

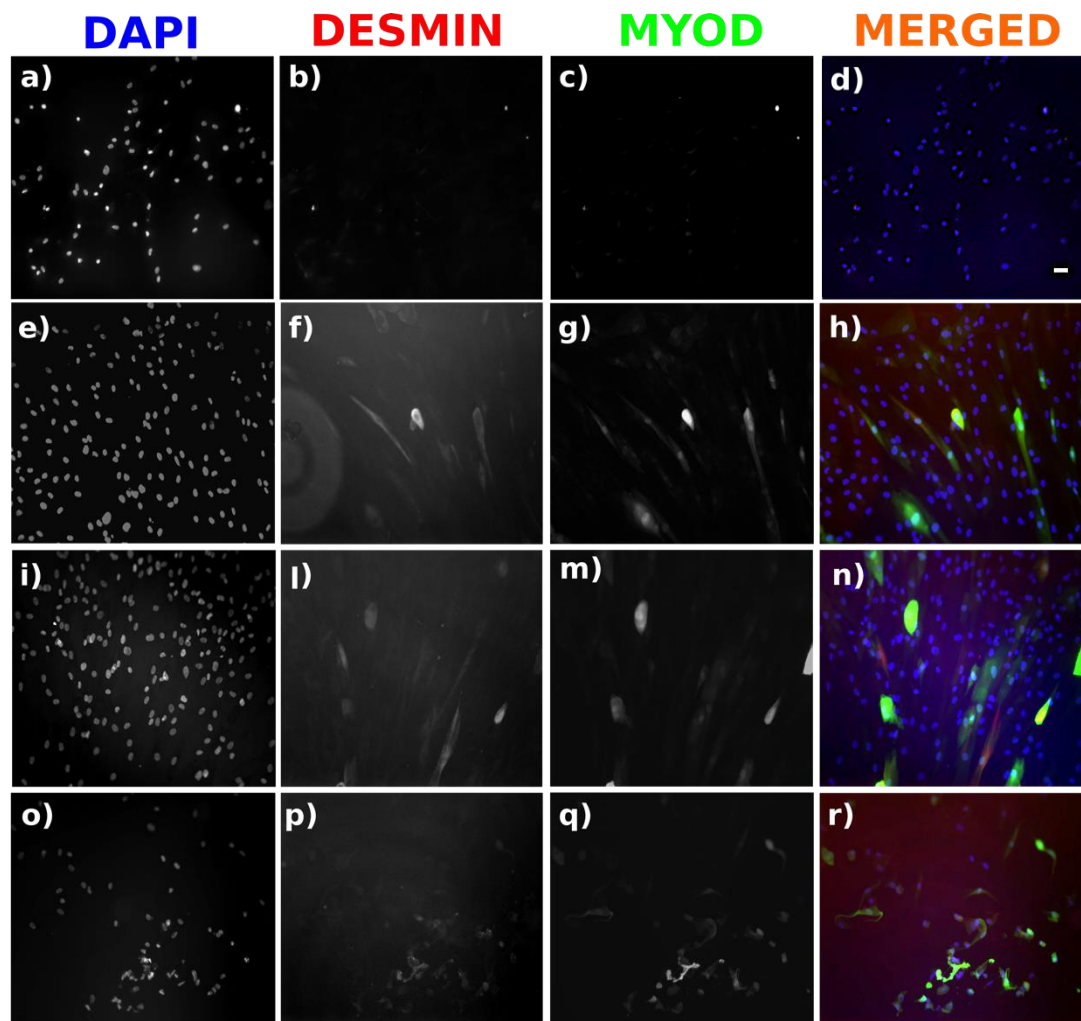


Figure 4.15. MyoD and desmin immunostaining.

Differentiated MyoD-transduced cells were fixed with 4% paraformaldehyde and permeabilized with Triton-X. Then they were incubated with DAPI, the mouse antibody targeting human desmin and the secondary antibody 594, which gives a red fluorescent signal. GFP expressed by successfully transduced cells should be also indicative of MyoD expression. Fluorescence was assessed with the Leica DMR fluorescent microscope at a magnification of 10X. Non-transduced cells (a-d) were used as a negative control. Transduced fibroblasts differentiated with 2% horse serum (e-h), 1% foetal bovine serum (i-n) and no serum (o-r). Scale bar = 100 μ m

4.4 DISCUSSION

In the present chapter, I investigated the ability of CRISPR/Cas9 to target intron 9 of the human DMD gene in DUP1 patient-derived fibroblasts, carrying a DMD duplication spanning exons 5-11.

To deliver CRISPR/Cas9 into the cells, I considered both viral and non-viral methods.

The use of integrating lentiviral particles was my first choice, as transgene transfection has generally a low efficiency in primary fibroblasts, especially with plasmids of considerable size like LentiCRISPR (around 11 kb)⁴⁹⁵. Also, lentiviral vectors provide a number of advantages such as their large packaging capacity, broad cellular tropism and, above all, the ability to permanently integrate into the host genome^{496,497}. This feature in particular was considered advantageous, as the integration of the CRISPR/Cas9 plasmid would allow the nuclease to be continuously expressed. Therefore, the chances of success of targeting are increased.

For this reason, I produced second generation HIV-1-based lentiviral particles to express in fibroblasts the two best LentiCRISPRs that I previously validated in HEK 293T cells.

I decided to first test these viral particles in HEK 293T cells by doing a dose escalation and I identified the minimal effective dose of lentiviral vectors that allowed genomic targeting without compromising cell viability. However, when I used the same transducing doses in fibroblasts, I did not observe genomic cleavage. I justified this outcome considering that the titres of the LentiCRISPRs lentiviral vectors were calculated from HEK 293T cell line, more easily transduced than fibroblasts⁴⁹⁸. Therefore, higher viral doses might have been needed to achieve an equal level of transduction in human fibroblasts. However, cells transduced with lentiviral particles expressing pLJM1-EGFP vector were fluorescent, indicating that the transducing doses I used were high enough to infect fibroblasts. I therefore questioned if a threshold of nuclease expression was required for mediating the cut in this cell type.

I realized the need of having a selection marker to identify the transduced cells and isolate them, so that I could clearly observe how many cells expressed the nuclease.

One possibility of selecting transduced cells was through the antibiotic resistance cassette carried by the LentiCRISPR transgene. However, I chose to avoid antibiotic selection because to obtain a pure CRISPR-expressing fibroblast population, I would have had to culture the cells for about 14 days⁴⁹⁹. Once transduced cells would have been selected, I would have had to expand them to assess their cleavage efficiency. The multiple passaging steps needed to reach this point would have increased the risk

of the cells reaching senescence. This issue would have been further exacerbated by the fact that fibroblasts do not express dystrophin, and for analysing dystrophin transcript and protein I would have had to manipulate them through the forced expression of the myogenic factor MyoD to induce myogenic differentiation. MyoD-expression has to be followed by culturing in a specific medium to induce terminal differentiation, which takes a further 10-14 days.

I therefore decided to use a selection method based on a fluorescent marker which, being detectable at 48 hours post-transduction, would have allowed a more rapid estimate and isolation of the transduced cells, so that I could lower the number of required passaging steps.

I bought lentiviral particles co-expressing both CRISPR/Cas9 components and the EGFP protein and I took advantage of the high titre of the viral particles provided by the company to increase the viral load even more, reaching up to the considerably high dose of 1000 MOI.

Unexpectedly, I observed that this dose did not elicit signs of cells toxicity nor a clearly distinguishable EGFP-positive signal, whereas the same dose of lentiviral particles expressing pLMJ1-EGFP plasmid (produced in house) gave instead both a clearly detectable EGFP expression and toxicity. Taken together, these results suggest that the purchased lentiviral vectors were not infectious and/or defective in their ability to express the transgene they carried. However, different methodologies were used to titrate the commercial lentiviral particles and the ones produced in house (p24 ELISA assay versus qPCR). This might have resulted in a different estimate of the viral genomes/ml, inevitably resulting in different viral doses at the same MOIs. qPCR in fact, estimates the titre by assessing the number of viral copies integrated into the genome, while p24 ELISA quantifies the viral particles expressing the capsid protein p24. Therefore, this last method might overestimate the titre including the empty viral particles that are fully formed (and therefore express p24) but do not express any transgene⁵⁰⁰.

I then investigated non-viral alternatives for delivering CRISPR/Cas9 in primary fibroblasts.

I considered that, if successful, transient transfection would have been safer for future *in vivo* applications, as the plasmid generally remains episomal and is diluted during each cell division. This would limit its risk of creating off-target effects (section 1.8). The tool I chose to achieve transient transfection in the patient-derived fibroblasts was an advanced electroporation method (nucleofection) potentially able to maximize transfection efficiency, as it exploits specific electric pulses to increase the permeability of the nuclear membrane and deliver the vector to the nucleus⁵⁰¹. I considered that,

as it is performed by a device set up with customizable parameters and electrical pulses for a wide range of cells, nucleofection would have abrogated the need of producing high-titres of viral particles and eased the experimental scale-up. I identified a plasmid that co-expressed the same guide as the lentiviral vectors I tested (sgRNA2), Cas9 and EGFP, so that I could visually identify CRISPR-targeted cells. However, in order to detect a genomic cleavage not only I had to isolate CRISPR/Cas9-expressing cells, but I had to consider only cells expressing the highest level of CRISPR/Cas9. This suggested the need of a Cas9 overload for the genomic cleavage to occur in primary fibroblasts. I hypothesized that this overload was needed because in this cell type dystrophin is poorly expressed and therefore the chromatin is potentially in a packaged status that limits the access of the nuclease to the target region. In support of this hypothesis, Horlbeck *et al.*⁵⁰² used chemicals able to alter the chromatin conformation and demonstrated that Cas9 efficiency varies in accordance to the accessibility of the target loci. I considered to investigate this further using chromatin modifier agents (as the histone deacetylase inhibitor butyric acid⁵⁰³), but I did not prioritize these experiments as I reached the first milestone of my project (i.e. identify CRISPR/Cas9-mediated genomic cleavage in patient-derived fibroblasts).

At this point I had to verify if the genomic targeting of CRISPR/Cas9 resulted in the correction of dystrophin transcript and protein. Therefore, I had to force these cells towards the myogenic lineage to obtain myotubes expressing dystrophin. I decided to deliver MyoD into the patient-derived fibroblasts via an adenoviral vector, as these vectors are generally transiently expressed and other investigators have reported on their successful use^{504–506}. I considered this feature essential to mimic the temporary MyoD expression seen in both embryonic development and muscle regeneration, which fluctuates through the cell-cycle⁴⁷⁶. Thinking that the transduction efficiency is the first parameter to be considered for successful MyoD-conversion of fibroblasts, I decided to use the Ad(RGD)-GFP-h-MyoD1 adenoviral vector (Vector BioLabs), as this vector co-expressed MyoD and GFP and allows the visualisation and selection of MyoD-expressing cells. I observed that, among the transducing dosages I tested, the one associated with the best infectivity and lower toxicity (MOI 100) was the same one used by Wojital *et al.*³⁸⁹. However, when I induced terminal differentiation, none of the media conventionally used allowed me to obtain more than 15% of nuclei in myotubes in culture after 9 days.

Even in this case, I wondered if the packed chromatin of my cells could have affected the process. A recent paper by Manandhar *et al.*, in fact, associated the incomplete differentiation of fibroblasts into myoblasts to the chromatin structure of the cells of origin⁵⁰⁷. He observed that, other than inducing the expression of genes important for

myogenesis and its own auto-regulatory loop, MyoD influences the chromatin conformation of the cells in which it is expressed by recruiting remodeling enzymes that change the regulatory landscape of the cells. However, this is achieved only to a certain extent ^{477,508,509}. A particularly packed chromatin status might therefore affect the efficiency with which MyoD triggers myogenic differentiation.

Based on these results, in future I would still rely on the GFP expression cassette for selecting transduced cells expressing MyoD, but I would select a type of vector associated with an easier modality of transduction (e.g. integrating lentiviral vectors) similar to Fuente *et al.*⁵¹⁰. Moreover, as I would have to ensure that the MyoD would not be constitutively expressed, I would exploit the approach with which Chaouch *et al.*⁵¹¹ obtained up to 70% myotubes, based on the use of an inducible promoter. Exploiting both of these factors would allow me to have a better control over MyoD expression in future experiments.

To summarize, the experiments detailed in this chapter show that:

- Electroporation of CRISPR/Cas9 plasmids is efficient in primary fibroblasts.
- The isolation of cells expressing high levels of Cas9 is required to assess genomic cleavage in these cells.
- The myogenic induction of fibroblasts through the forced expression of MyoD needs further optimization, especially if relatively high levels of myotubes are required for the experimental procedure

As I could not verify if the genomic cleavage observed by the T7 assay resulted in the removal of the duplication in dystrophin protein, I moved to the use of alternative patient-derived cells (focus of Chapter 5).

Chapter 5. Delivery of CRISPR/Cas9 nucleases to patient-derived myoblasts and restoration of dystrophin duplications

5.1 INTRODUCTION

5.1.1. Myoblasts: the ancestors of the muscle fiber

Myoblasts are muscle precursor cells of mesenchymal origin that, upon their fusion, give rise to multinucleated myotubes which *in vivo* constitute the muscle fiber⁴⁹⁰.

During development, the formation of mature muscle proceeds through distinct phases involving the sequential and alternate expression of four myogenic regulatory factors named Myf5, MyoD, MRF4 and myogenin⁵¹². The combined expression of Myf5 and MyoD appears indispensable for the early muscle formation, while the expression of MRF4 and myogenin is predominant during late and terminal differentiation, respectively⁵¹³.

In culture, myoblasts proliferate in presence of growth factors such as fibroblast growth factor (FGF) until differentiation is induced: the removal of growth factors from the culture medium prevents myoblast proliferation⁵¹⁴, which ultimately results in myoblast alignment and then their fusion to form the multinucleated myotubes. The possibility to at least partially recreate the muscle structure found *in vivo* makes myotubes a good cell type for the study of muscle protein like dystrophin, not expressed in other more commonly used cell types as fibroblasts.

5.1.2. Patient-derived myoblasts: downsides and benefits

Unlike fibroblasts which can be obtained with a relatively non-invasive method (skin biopsy), the derivation of skeletal muscle myoblasts requires a muscle biopsy, a procedure that might be particularly burdensome for young neuromuscular patients who need a general anesthesia rather than a local anesthesia. A muscle biopsy not only leaves a scar, but the general anesthesia is associated with an increased risk for malignant hyperthermia in susceptible individuals, such as those affected by DMD⁵¹⁵. Muscle biopsy is therefore generally performed only if needed for diagnosis, thus the limited availability of DMD patient-derived myoblasts. Even when a primary myoblast cell line is established, its culture is not devoid of difficulties. In fact, the percentage of satellite cells from which myoblasts are derived is often lower in patients with neuromuscular conditions than healthy muscles, thus providing an obstacle to the

establishment of a proliferative myoblast population. In fact, it has been hypothesised that the loss of dystrophin in satellite cells impairs their ability to undergo asymmetric division from which new myoblasts are generated. Furthermore, the fibrotic/dystrophic environment reduces the efficiency of satellite cells activation¹³⁸. For these reasons these cells cannot efficiently aid muscle regeneration⁵¹⁶.

A pure myoblast population can be difficult to obtain following muscle biopsy, as these cells are often contaminated with muscle fibroblasts⁵¹⁷. Since the myogenicity of each primary myoblast cell line (and consequently the ability of forming myotubes *in vitro*) may vary depending on the percentage of non-myogenic cells present within the preparation, it is important to assess the myogenicity of the cell preparation before conducting experiments. In fact, dystrophin protein can be detected in myotubes but not in myoblasts. Working on a myoblast cell line that has a limited capacity of generating myotubes *in vitro* might therefore lead to the inability to study the protein of interest. The myogenicity of each patient-derived myoblast cell line is assessed via immunostaining for desmin, a cytoskeletal protein belonging to the intermediate filaments, which is expressed in endothelial and muscle cells⁵¹⁸. Although its essential role in myoblast development and differentiation has been debated, desmin has been reported to be expressed in myoblasts prior to MyoD expression⁵¹⁹ and is considered one of the earliest muscle-specific structural protein expressed in myogenesis, whose main role is sustaining muscle strength⁵²⁰.

Although the identification of a pure and myogenic cell line may not be straightforward, the use of patient-derived myoblasts might represent the ideal alternative for *in vitro* proof-of-concept studies, such as the one presented here. Patient-derived myoblasts, in fact, can be used as target cells to test and validate therapeutic tools such as CRISPR/Cas9 so that, once the mutated protein is restored, they could be used for autologous transplantation into patients without the need for immunosuppression.

5.2 AIMS

The present chapter aims to:

- Compare the ability of viral and non-viral approaches for delivering the CRISPR/Cas9 editing tool into myoblasts derived from DMD patients carrying dystrophin duplications
- Demonstrate the ability of a single CRISPR/Cas9 nuclease approach of repairing dystrophin duplications (both at genomic, transcript and protein level) in DMD patient-derived myoblasts

Each experimental step followed to pursue these aims is indicated in **Fig.5.1**.

Dystrophin correction in patient-derived myoblasts

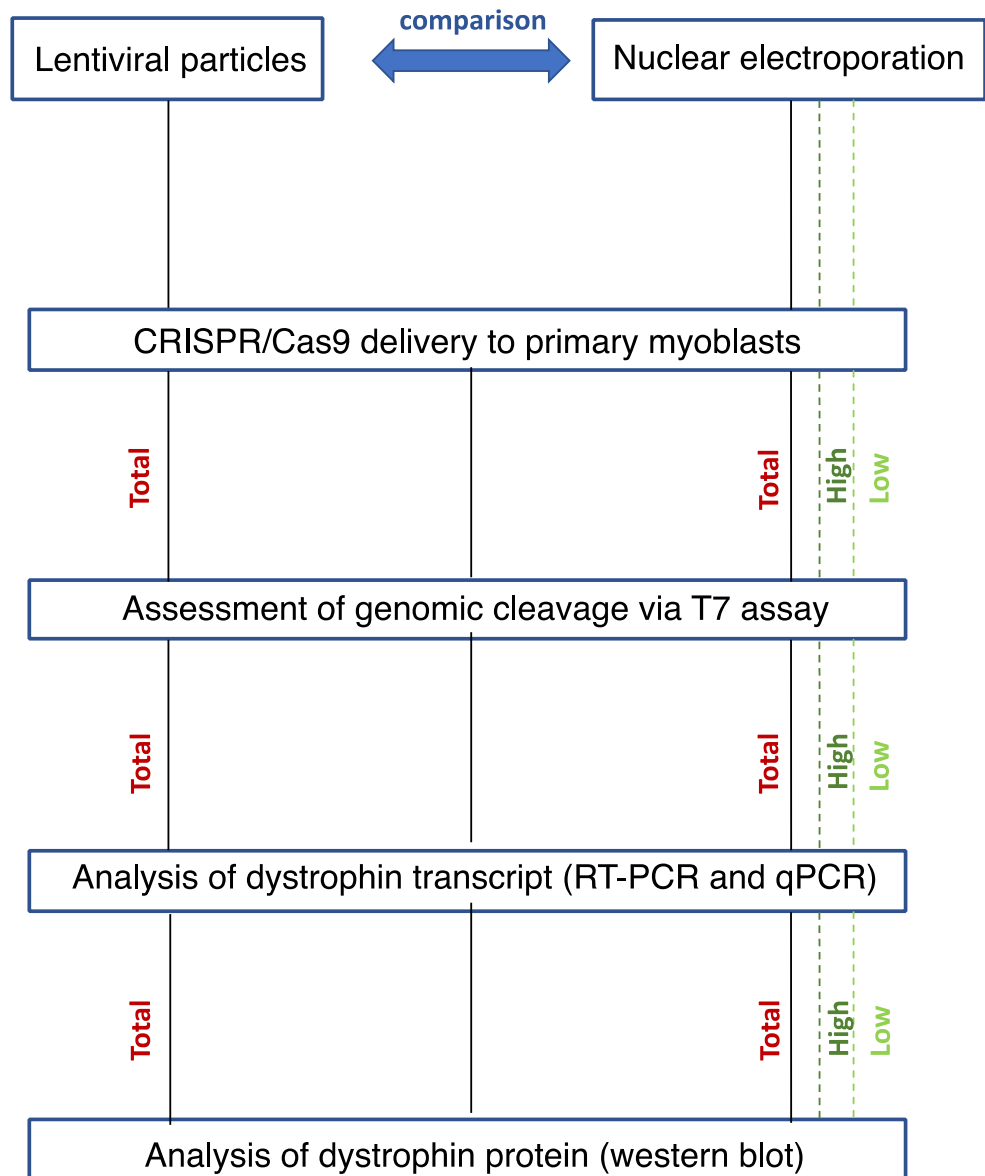


Figure 5.1. Schematic of the experimental steps required for restoring dystrophin in patient-derived myoblasts.

Both lentiviral particles and electroporation will be tested as tool to deliver CRISPR/Cas9 into DMD myoblasts. Total= total cells expressing CRISPR/Cas9; High= cells expressing high levels of CRISPR/Cas9; Low= cells expressing low levels of CRISPR/Cas9.

5.3 RESULTS

5.3.1. Selection of myogenic myoblast cell line

In order to assess if the CRISPR/Cas9 nuclease could remove *DMD* duplications and restore dystrophin protein, I had to identify a myoblast cell line with high myogenicity. This was to ensure the presence of a sufficient number of myotubes expressing dystrophin, so that this could be detected by western blot. Among the DMD patient-derived myoblast cell lines available to me (sections 2.2.1 and 2.2.2) and carrying a dystrophin duplication spanning exon 9, I chose DUP2 (**table 2.1**, section 2.2.2) as it was highly myogenic. DUP2 carries an in-frame duplication of exons 3-16, which results in the production of a mutated dystrophin protein of higher molecular weight (499 kDa) than wild-type dystrophin (427 kDa). The possibility to observe the mutated protein is advantageous as it works as endogenous control for monitoring the extent dystrophin protein repair.

I confirmed the myogenicity of this cell line through desmin staining and, upon induction of terminal differentiation, the quantification of fusion index (**Figures 5.2. and 5.3**). Desmin was detected in both myoblasts and myotubes following incubation with the mouse monoclonal antibody M-0760 against human desmin (section 2.3.17.a). The fusion index was 90%.

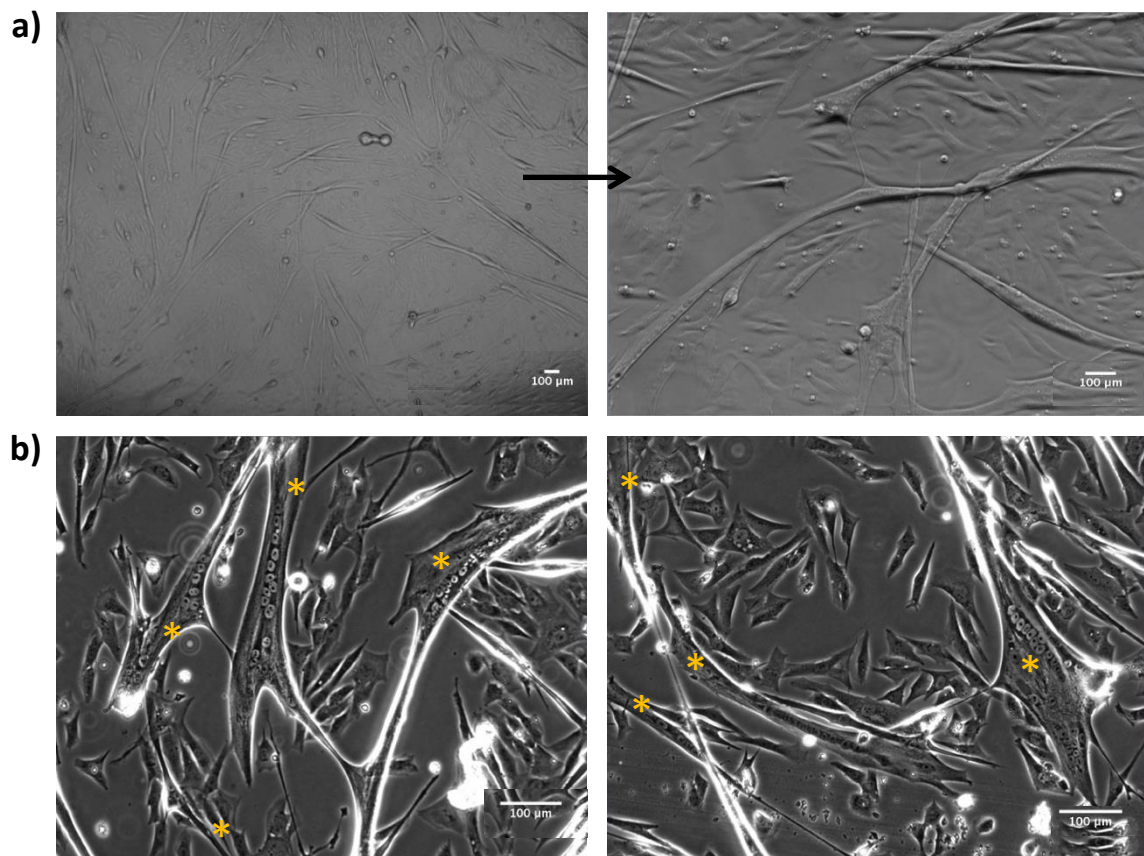


Figure 5.2. Myotubes originated from terminally differentiated DMD1 myoblasts.

Myoblasts DUP2 were seeded in a six-well plate in order to be 80% - 90% confluent the following day. These were cultured with the M2 differentiation medium for 7-9 days and imaged by the Olympus IX inverted microscope to assess the presence of myotubes. Myotubes are easily seen at lower magnification 10X and 20X (a), while with higher magnification 40X their characteristic multinucleated pattern was observed (b). Terminally differentiated multinucleated cells are shown by the yellow asterisks.

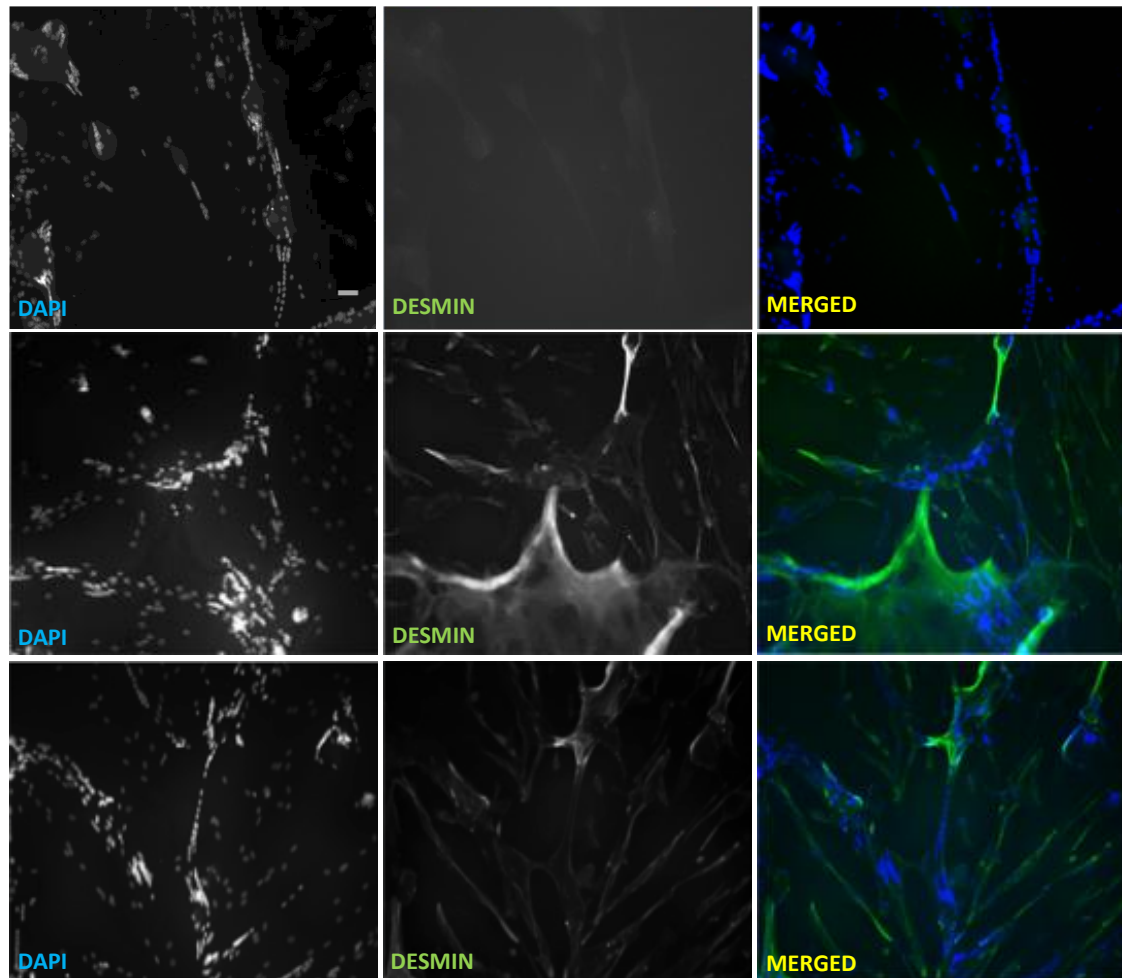


Figure 5.3. Myogenicity of DUP2 myoblasts.

Myoblasts DUP2 seeded in each well of a 8-wells chamber slide were stained with DAPI (1:1000) and the primary antibody AbM0760 (1:100) against human desmin (central and bottom rows). The top row corresponds to cells incubated with DAPI and secondary antibody only (1:1000). The fusion index was 90%. Cells were imaged by the Olympus IX fluorescent inverted microscope at a magnification of 5X. Scale bar = 200 μ m.

5.3.2. CRISPR/Cas9 delivery in patient-derived myoblasts: viral approach

To achieve the highest possible nuclease expression, I decided to deliver CRISPR/Cas9 in DUP2 myoblasts by means of integrating lentiviral particles.

I chose the lentiviral transgene pL-CRISPR.EFS.GFP (section 2.1.2.b), where I cloned the sgRNA2 (section 2.1.5.d). Despite being very similar to LentiCRISPRv1 (tested in human fibroblasts), the plasmid backbone had the puromycin resistance cassette replaced with the EGFP cassette, fused to Cas9 via a P2A element. This would have allowed me to isolate EGFP-positive cells (which also expressed Cas9) via FACS, as I have also performed for the electroporated fibroblasts (section 2.2.14.a). As the transgene itself expresses EGFP, I did not need to include lentiviral particles expressing EGFP transgene as a positive control for monitoring the transduction efficiency. The only control I used for this set of experiments was represented by non-transduced cells (negative control).

I generated lentiviral particles expressing pL-CRISPR.EFS.GFP (section 2.1.6) and tested increasing viral doses (ranging from MOI 0.4 to 25.6) in DUP2 myoblasts. I did this to identify the best dose combining the lowest cell toxicity and the highest percentage of transduced cells. I repeated this experiment another three times (biological replicates).

The highest MOI (25.6) was toxic, as 90% of the cells died. The lowest MOI (0.4) was the safest as, despite a considerably weak fluorescence which was almost undetectable by microscopic analysis after 48 hours (**Fig.5.4**), this dose resulted in a better cell viability (estimated by visually quantifying the live cells as opposed to the rounded floating cells). The mean transduction efficiency estimated by FACS was 4.48% \pm 0.14 (**Fig.5.5**). Since recent evidence showed that the transduced cell populations expressing between 5-30% GFP-positive cells are likely to carry only one viral genome/cell⁵²¹, I considered my transduction efficiency adequate for ensuring reproducibility in the degree of nuclease expression. 72 hours post-transduction, I collected the EGFP-positive cells via FACS (section 2.2.14.a) and expanded them (section 2.2.6).

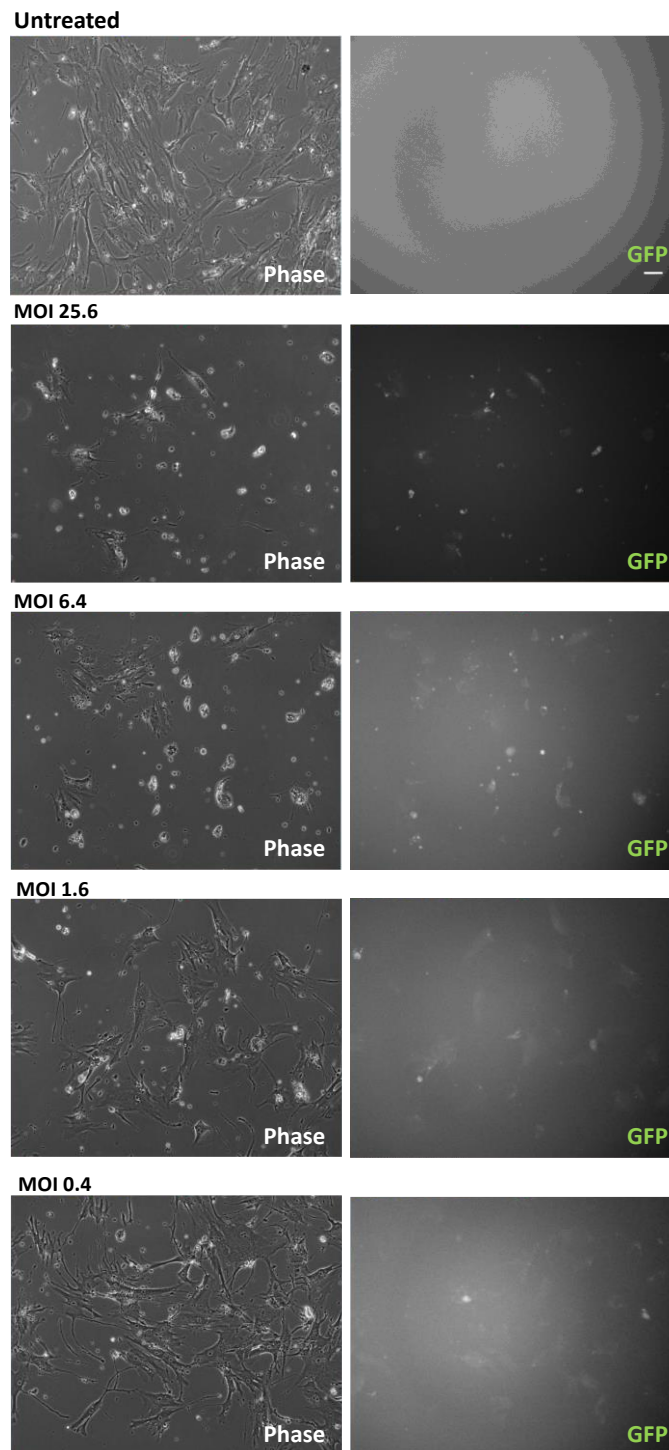


Figure 5.4. Selection of the best transducing MOI.

DUP2 myoblasts were transduced with a range of viral particles doses in order to identify the best dose to give a detectable fluorescence (indicative of the transgene expression) while ensuring the cells' viability. MOI dosages were scaled down from MOI 25.6 to MOI 0.4. Non-transduced cells (top row) were used as a negative control for the transduction, and a positive control to monitor cell wellbeing. Clearly detectable EGFP-positive cells were observed with the highest MOI, although most of these were dead (round floating cells). A faint fluorescence signal was instead detected with MOI 0.4, which represented the less toxic transducing dosage for these cells. Cells were imaged by the Olympus IX fluorescent inverted microscope at a magnification of 10X. Scale bar = 200 μm .

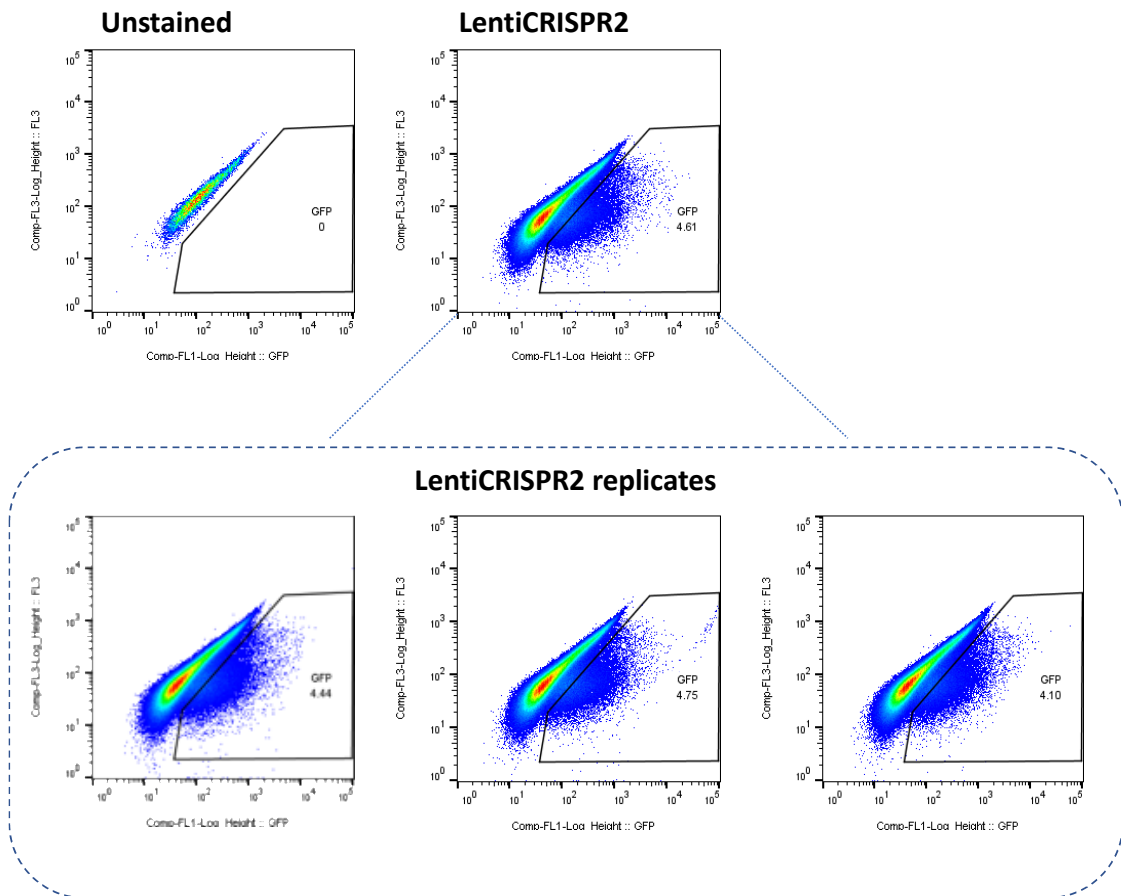


Figure 5.5. FACS sorting of transduced DUP2 myoblasts.

DUP2 myoblasts transduced with 0.4 MOI of the lentiviral particles expressing the pL-CRISPR.EFS.GFP vector were sorted via FACS to precisely estimate the transduction efficiency. The average transduction efficiency across the four replicates was 4.48 % \pm 0.14. Most of the fluorescent cells were close to the gated non-targeted cells, confirming the low level of EGFP expression associated with this MOI and previously observed microscopically.

a) Verification of CRISPR/Cas9-mediated cleavage at genomic level

Once the sorted and expanded cells reached a number of about 3×10^5 , I extracted their DNA and amplified it for performing the T7 assay (sections 2.3.1 and 2.3.4-2.3.6). By running the T7 assay products on a 4-20% TBE gel (section 2.3.8), I observed cleaved bands of the expected size in all the transduced samples expressing the nuclease, but not in the non-transduced samples (negative controls). The editing efficiency appeared to be stable across different CRISPR/Cas9 lentiviral transduction (LVs): LV1 = 28.3%, LV2 = 29.5%, LV3 = 30.03%, LV4 = 28.2% (**Fig.5.6**). The mean editing efficiency was 29.0% +/- 0.45.

b) Verification of dystrophin transcript repair following CRISPR/Cas9 activity

I kept an aliquot of the sorted cells in proliferation medium for further expansion until cells reached a confluence of 70-90%. Then I cultured these cells in the M2 medium (Sigma Aldrich) to induce terminal differentiation (section 2.2.6). Terminally differentiated myotubes (50% fusion index) were collected after 7 days (section 2.2.8) (**Fig.5.7**), so that I could extract RNA and protein (section 2.3.10 and 2.3.19).

I investigated if the genomic cleavage observed in cells expressing CRISPR/Cas9 resulted into the repair of dystrophin transcript. I generated cDNA from the RNA I extracted by non-transduced and transduced myotubes (section 2.3.11.a), and I used it to do both RT-PCR and qPCR (sections 2.3.12 and 2.3.13.b).

I ran two parallel RT-PCRs with parameters set to selectively amplify either the restored transcript only, or both (section 2.3.12). In this experiment, I included cDNA (provided by Dr. John Counsell) that was previously derived from differentiated wild-type human CD133+ cells. As these cells expressed wild-type dystrophin, I consider them as a positive control. The controls used for this experiment are summarized in the below **table 5.1**. The RT-PCR was repeated three times (biological replicates).

Control (+/-)	cDNA	Function
-	Non-transduced cells	Amplify mutated transcript
+	CD133+ cells	Amplify wild-type transcript

Table 5.1. List of controls included in the RT-PCR to assess repair of dystrophin transcript.

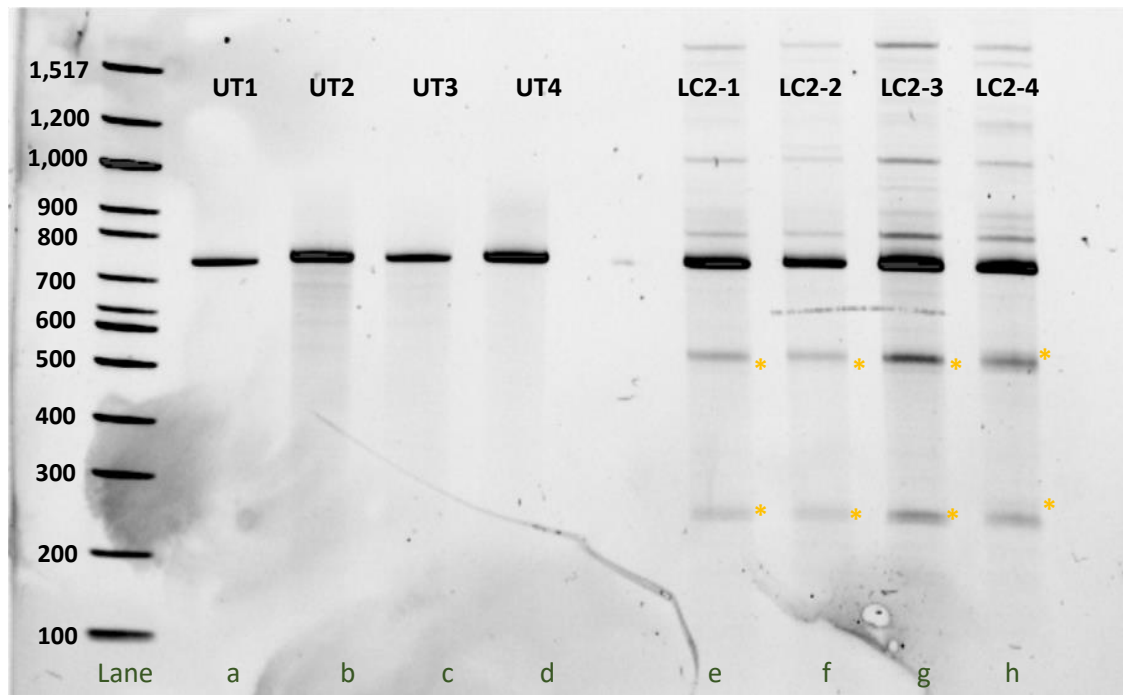


Figure 5.6. T7 assay on transduced DUP2 patient-derived myoblasts.

T7 assay was performed on DNA amplicons derived from non-transduced and transduced DUP2 myoblasts expressing CRISPR/Cas9 (MOI 0.4). Lanes a-d indicate the full-length amplicon derived from non-transduced cells, while lanes e-h indicate transduced cells expressing the nuclease. Two cleaved bands of the expected molecular weight of 511 bp and 217 bp (*) are present in the transduced samples, indicating successful genomic cleavage (mean efficiency: 29.0% \pm 0.45). A 100 bp ladder (NEB) was run alongside the samples to monitor the size of each detected fragment.

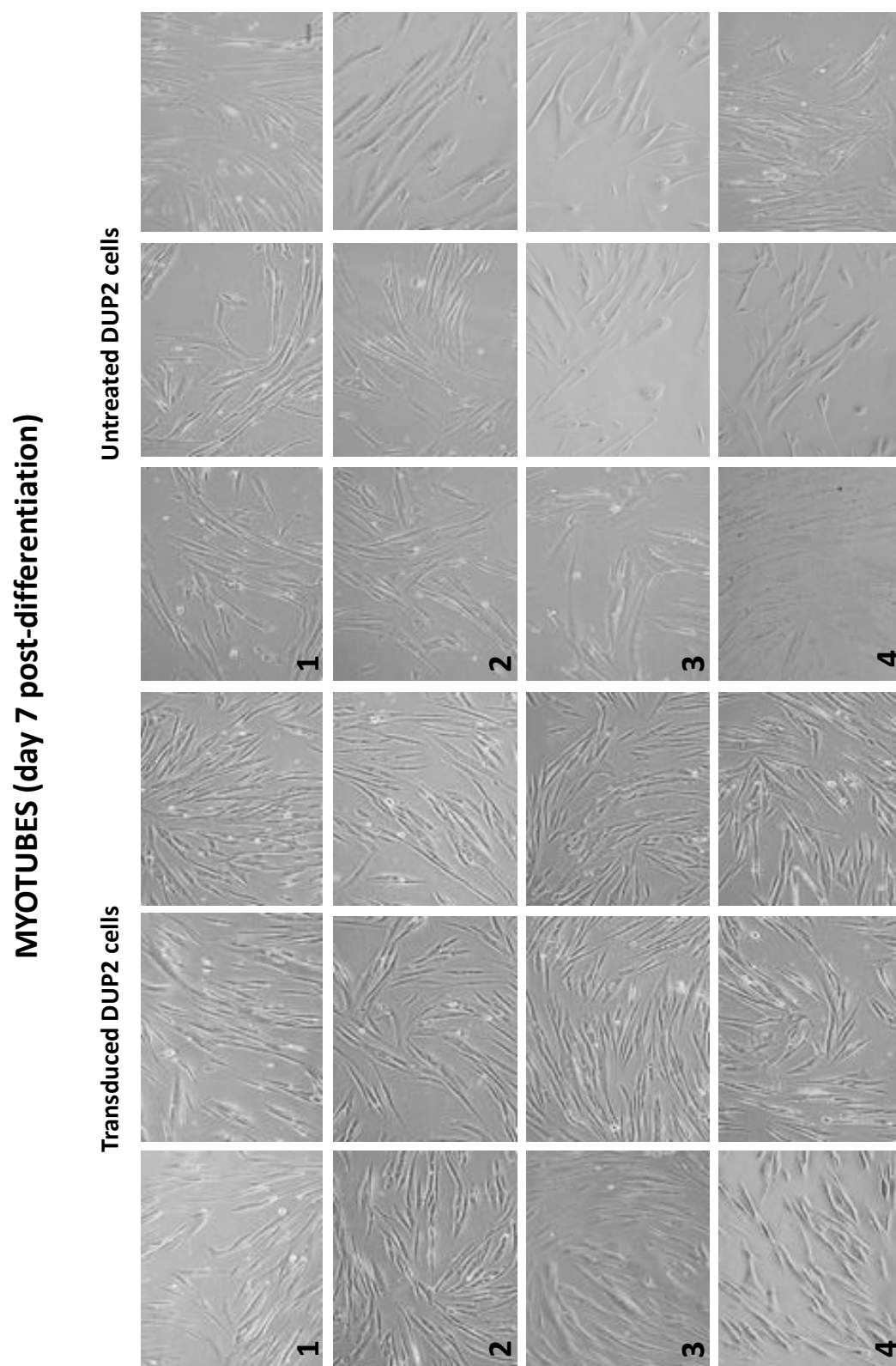


Figure 5.7. Terminal differentiation of DUP2 cells.

Untreated DUP2 cells and DUP2 cells transduced with lentiviral particles expressing CRISPR/Cas9 were induced towards terminal differentiation and imaged after 7 days. Both fusing cells and mature myotubes were detected at this stage and appeared to be more prominent in the untreated cells (right panel). Cells were imaged by light microscope at a magnification of 10X. Scale bar = 500 μ m.

I observed that the intensity of the band corresponding to the restored transcript was high in both amplification conditions. However the band corresponding to the mutated transcript was very weak in the untreated sample and, moreover, also appeared in lanes containing the wild-type sample (**Fig.5.8**). I observed this result also in different experimental RT-PCR replicates, indicating the need to assess the accuracy of the experimental design and primers used to assess the repair of dystrophin transcript.

The RT-PCR experiment was not pursued further, as the parallel approach based on qPCR (section 2.3.13.b) allowed me to assess more accurately the restoration of dystrophin transcript.

For the qPCR I used a different set of primers (2.3.13.b) designed to amplify regions of about 100 bp. I decided to amplify regions spanning dystrophin exon 8 and 9 (within the duplication) and the constitutively expressed GAPDH gene, that I chose as a normalizer gene.

To confirm the specificity of the chosen primer pairs, I performed a qPCR based on serial dilutions of cDNA derived from non-treated DUP2 myotubes and I analysed their amplification profile. These primers did not show any non-specific activity. I therefore repeated the qPCR. I included cDNAs from both non-treated (negative control) and transduced myotubes. The qPCR plate was set-up including a triplicate of each sample (technical replicate). I analysed the qPCR data by the $\Delta\Delta C_t$ method (section 2.3.14), following which I derived fold changes values.

I observed that the expression of dystrophin exons 8-9 normalized to GAPDH was diminished in transduced samples as opposed to the non-transduced ones. This suggested a reduction of the duplicated dystrophin transcript caused by the treatment with the nuclease (**Fig.5.9.a**). Such reduction was statistically significant when the Mann-Whitney test was applied (p value= 0.0036).

However, I hypothesized that GAPDH might not have been the best normalizer to use. This consideration came from the notion that GAPDH levels vary depending on the level of terminal differentiation of the cells⁵²² and the observation that, visually, the myogenicity of the cells was reduced following the nuclease treatment (**Fig.5.10**). Therefore, I introduced a different normalizer to quantify the extent of dystrophin transcript correction. This was represented by a dystrophin region located outside the duplication (anchor region) (section 2.3.13.b). To avoid the amplification of multiple dystrophin transcript isoforms, I chose the anchor region before any of the dystrophin promoters that are located throughout *DMD* gene (section 1.3.2), specifically within *DMD* exon 20. Similarly to what I detected in data normalized against GAPDH, upon

data analysis performed with $\Delta\Delta C_t$ method (section 2.3.14), I observed a reduction of the mutated dystrophin transcript. However, differently than what I observed with GAPDH, such reduction was not statistically different when I applied the Mann-Whitney test (p-value = 0.28)(**Fig.5.9.b**).

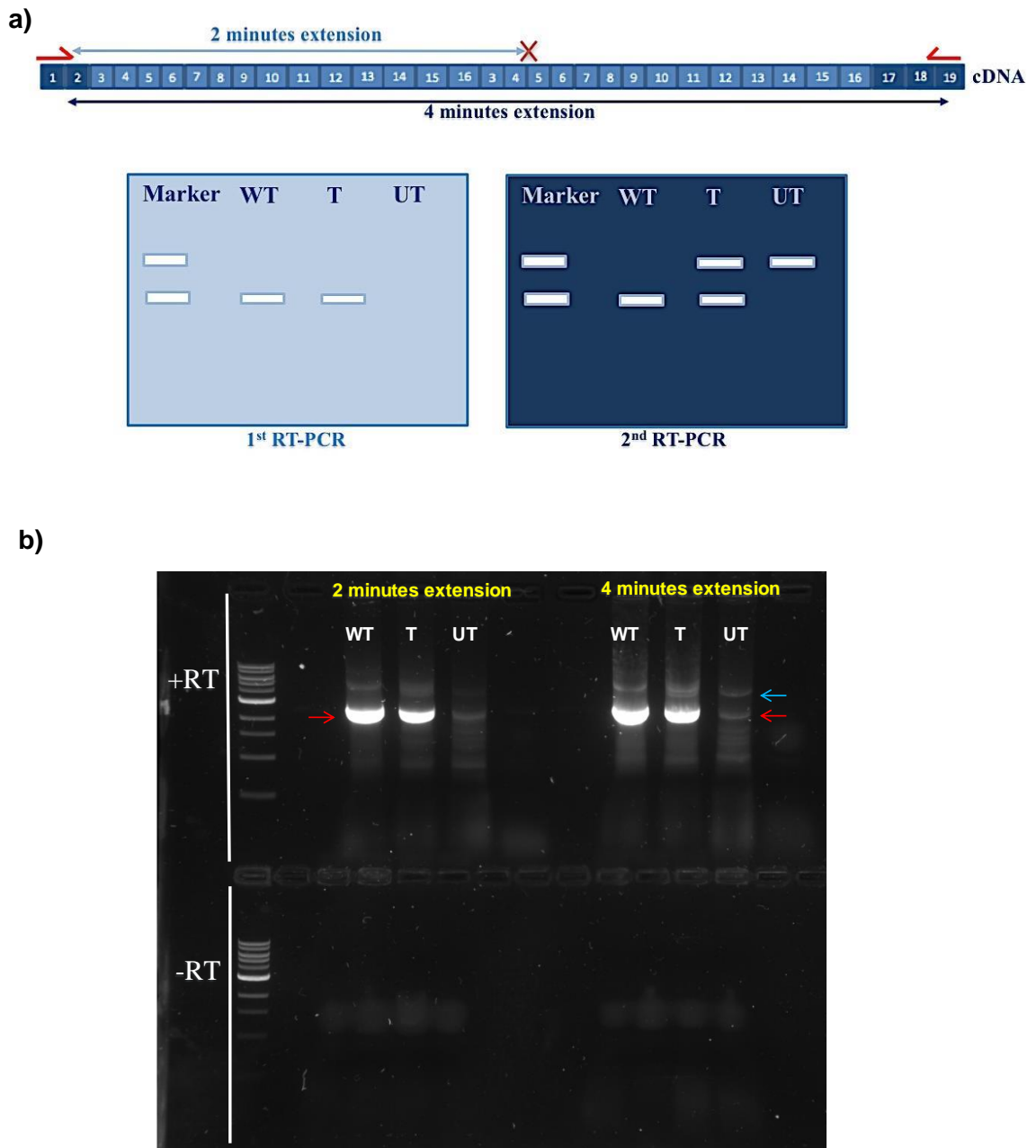


Figure 5.8. Identification of corrected dystrophin transcript via RT-PCR.

a) RT-PCR strategy design. An extension time of 2 minutes would allow the amplification of wild-type (WT) and corrected transcripts (T) only, whose size is expected to be 2,348 kb. Conversely, mutated transcripts derived from untreated (UT) cells (4,337 kb) could only be amplified with a 4 minute extension time. b) Arrows indicate the size of expected amplicons. Results obtained by RT-PCR matched those predicted, as the corrected transcript was amplified (red arrows) with both extension times, as expected. Amplicons whose size corresponded to untreated transcript were detected only with 4 minute extension times (blue arrow), even though non-specific amplicons appeared also in the WT sample. +RT and -RT indicate reactions set up with and without the reverse transcriptase enzyme, respectively.

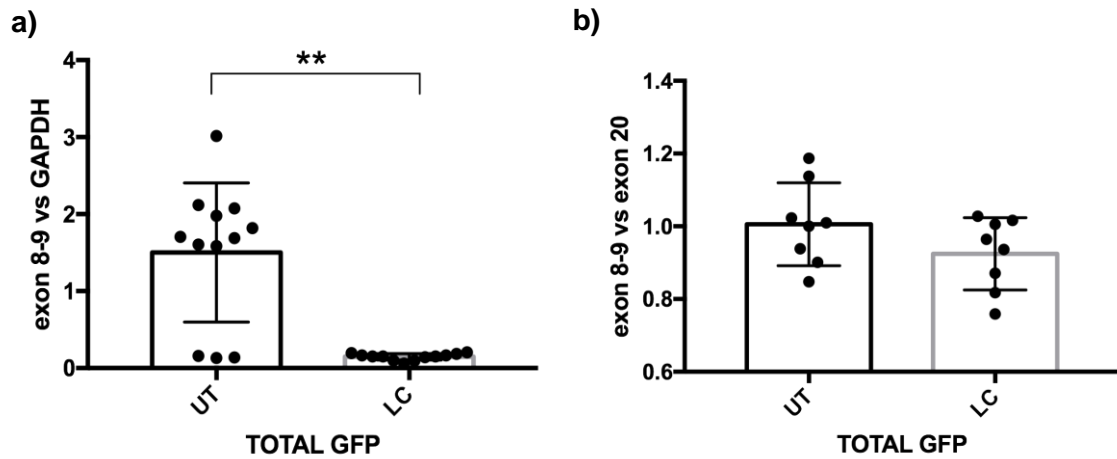


Figure 5.9. Quantification of dystrophin transcript correction in transduced DUP2 cells.

qPCR was performed to assess if CRISPR/Cas9 treatment resulted in a reduced expression of mutated transcript (represented by dystrophin exon 8-9 junction, within the duplication). qPCR data showed a significant reduction of mutated transcript in transduced cells expressing CRISPR/Cas9 when data were normalized to GAPDH ($p=0.0036$) (a) but not when normalization was based on the anchor region dystrophin exon 20 ($p=0.28$) (b). Error bars represent the standard deviation.

Untreated DUP2



Transduced DUP2

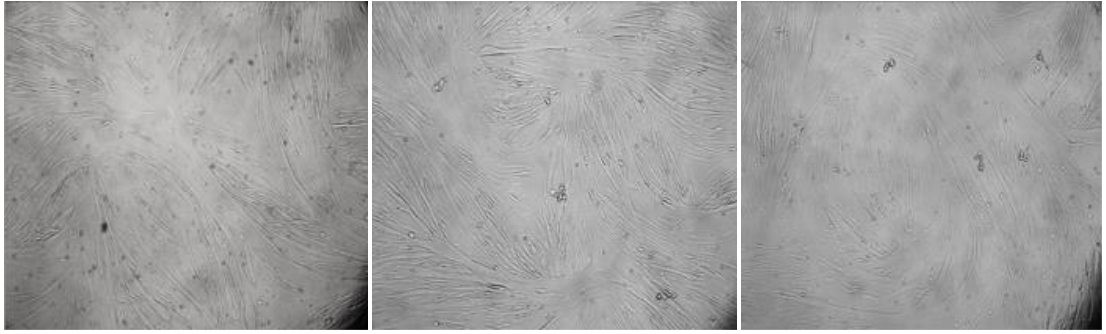


Figure 5.10. Effect of CRISPR/Cas9 transduction on myoblast myogenicity.

DUP2 cells were induced towards terminal differentiation and imaged after 7 days by the Olympus IX fluorescent inverted microscope at a magnification of 10X. Visually, cells transduction (bottom row) appeared to result in the formation of fewer myotubes than in untreated cells (top row). Scale bar = 200 μ m

c) Verification of dystrophin protein repair following CRISPR/Cas9 activity in DUP2 cell line

I assessed dystrophin correction at the protein level via western blot (section 2.3.22). As negative control, I included protein extracted from non-transduced cells, which should only express the mutated dystrophin isoform. As none of the control myoblasts cell line was myogenic, I could not obtain myotubes from which extract protein. As such, the positive control was represented by protein extracted from the immortalized H2K 2B4 cells (provided by Dr. Silvia Torelli), which express wild-type dystrophin. These controls are indicated in **table 5.2**.

Control (+/-)	Cells from which protein were extracted	Function
-	Non-transduced cells	Visualize mutated protein
+	H2K 2B4 cells	Visualize wild-type protein

Table 5.2. List of controls included in the western blot to assess repair of dystrophin protein.

Initially, I tested TurboBlot (section 2.3.22.a) as a method to transfer dystrophin to the PVDF membrane. I incubated the PVDF membrane with the dystrophin polyclonal antibody Ab15227 (recognizing dystrophin C-terminus) and β -actin (section 2.3.22.a), which I chose to monitor the success of protein transfer. I did not detect dystrophin when the PDVF membrane was analysed with the Odyssey Infrared Imager, whereas I could observe β -actin (**Fig.5.11.a**). I hypothesized that the lack of detectable dystrophin was due to the short transfer time imposed by the Trans-Blot apparatus, which might not have been sufficient to entirely transfer a large molecular weight protein as dystrophin. In fact, with this protocol I could not detect even the mutated dystrophin of higher molecular weight resulting from the in-frame mutation carried by DUP2 cells (**table 2.1**, section 2.2.2).

Therefore I tested a different protocol (section 2.3.22.b) based on a longer transfer time (Wet Blot) and signal amplification of the biotinylated secondary dystrophin antibody by means of streptavidin-HRP (section 2.3.22.b). I incubated the PDVF membrane with the same dystrophin polyclonal antibody as above (Ab15227) and, instead of β -actin, with the V9131 monoclonal vinculin antibody (section 2.3.22.b). I calculated the amount of restored protein by the ratio between the amount of duplicated and wild-type dystrophin within the same lane. In addition, I also included vinculin in the western blot to ensure that the same amount of starting protein was being analysed.

I chose this vinculin antibody as it also recognizes the vinculin isoform known as meta-vinculin, which can be distinguished from vinculin because of its slightly higher

molecular weight. Meta-vinculin is only expressed in cardiac and skeletal muscle tissue but not in non-muscle cells, and increases on skeletal muscle differentiation⁵²³. Because of that, meta-vinculin can be indicative of how well differentiated cells were at the moment of protein extraction and, consequently, of how much of dystrophin to expect in each sample.

I observed that, despite the little dystrophin transcript restoration measured by qPCR, protein editing was extremely efficient: the wild-type protein observed in each replicate was quantified as approximately 50% of the total dystrophin, with a mean value of 59.9% +/- 6.74. (**Fig.5.11.b and c**). The values of dystrophin restoration quantified in each of the four replicates were 49.7%, 70.1%, 72.9% and 46.9%. The outcome of the western blot allowed me to confirm that lentiviral particles could successfully deliver and express CRISPR/Cas9 in patient-derived myoblasts. This result indicated that the constitutive CRISPR/Cas9 expression can correct dystrophin duplication in these cells.

Once demonstrated the efficiency of my approach, however, I considered that the continuous expression of the nuclease in the cells might lead to the cleavage of non-target genomic regions, thus limiting the applicability of such approach for therapeutic purposes.

I therefore wondered if a non-viral delivery methods, that would allow a transient CRISPR/Cas9 expression, could be used as an alternative to ensure a comparable level of correction to that observed by means of integrating lentiviral vectors.

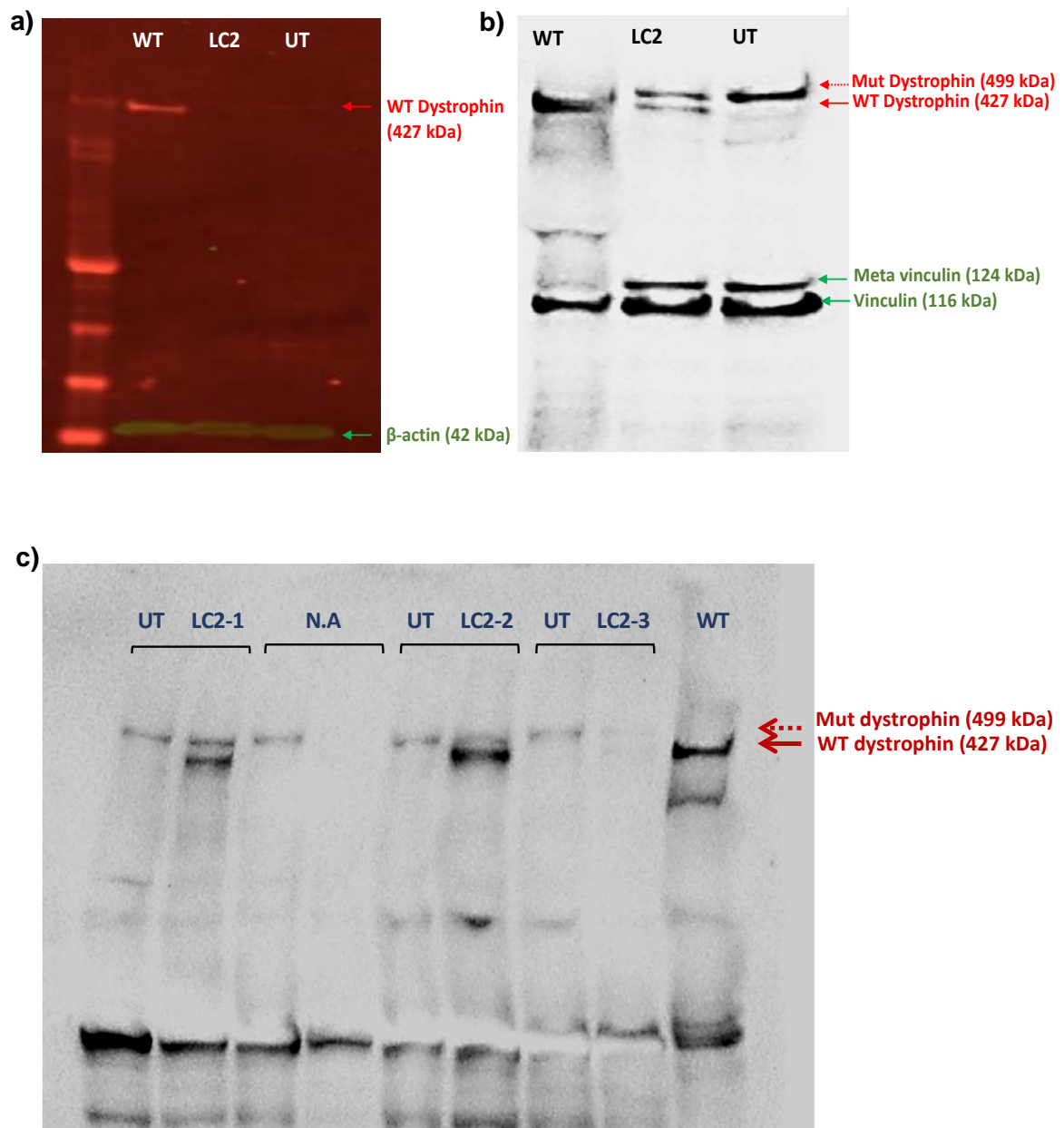


Figure 5.11. Validation of dystrophin repair in transduced DUP2i myoblasts.

Western blot was performed to verify dystrophin protein correction in transduced DUP2 cells expressing CRISPR/Cas9. UT = protein derived from untreated cells (negative control); LC2 = protein derived from transduced cells; WT= protein derived from the murine H2K 2B4 cells, expressing wild-type dystrophin (positive control). a) Turbo Blot only allowed the transfer of the wild-type (but not of the mutated) dystrophin isoform to the PVDF membrane. Conversely, Wet blot followed by signal amplification (b) allowed the transfer and visualization of both isoforms. c) Experimental replicates of independent DUP2 transduction processed as (b). Wild-type (WT) and mutated dystrophin (Mut) of 427 kDa and 499 kDa, respectively, are indicated by the red arrows. The mean efficiency of dystrophin repair was estimated as 59.9 ± 6.74 .

5.3.3. CRISPR/Cas9 delivery in patient-derived myoblasts: non-viral approaches

a) Transfection by using commercially available transfection reagents

To identify a suitable non-viral method to deliver CRISPR/Cas9 to human primary myoblasts I screened three commercially available transfection reagents (Lipofectamine2000, TurboFect and GeneJuice), each having different mechanisms to allow DNA entry into the cell.

Similarly to Lipofectamine 2000, which forms cationic liposomes able to entrap the DNA, TurboFect is based on cationic polymers able to complex with the negatively charged DNA, while GeneJuice is composed of a small amount of polyamine and non-toxic cellular protein^{524–526}. I performed each transfection three times (biological replicates) using the EGFP-expressing plasmid pCMV-GFP plasmid (**Fig.2.10**), whose size was 4,479 kb. I chose this plasmid as I wanted first to identify the best transfection condition that I would have later used for the transfection of CRISPR/Cas9 plasmids. As a negative control, I used non-transfected cells. For each reagent, I transfected 2×10^5 cells seeded in one well of a six-well plate (sections 2.2.11.b-d). I tested four lipofectamine conditions based on different DNA:Lipofectamine ratios, to identify the best one among those suggested by the manufacturer's protocol. The amount of DNA (2.5 μ g) was kept constant across the Lipofectamine transfections, while I increased the dose of Lipofectamine 2000 from 6 μ l to 15 μ l. Initially, I assessed the efficiency of transfection by fluorescence microscopy (section 2.3.23). I quantified the ratio between the number of cells in the plate and the GFP-positive cells in five fields randomly acquired for each of the transfected wells (**Fig.5.12**). According to microscopic analysis, none of the tested conditions had a transfection efficiency greater than 5% (**Table 5.3**). To avoid under- or overestimation of the transfection efficiency which may relate to the stochastic choice of the fields I imaged, I performed FACS analysis. In this way, I could more precisely determine the transfection efficiency. I observed that Lipofectamine 2000 transfection (section 2.2.11.b) had a very low efficiency in the tested conditions, apart from when 9 μ l of reagents were used, which gave intensely bright GFP-positive cells with a mean transfection efficiency of 6.84% \pm 0.64 (**Fig.5.13**). However, I observed that this reagent led to a considerable toxicity which resulted in substantial cell death (visually estimated as 60-80% depending on Lipofectamine 2000 dose). I therefore concluded that this method was not suitable for my study.

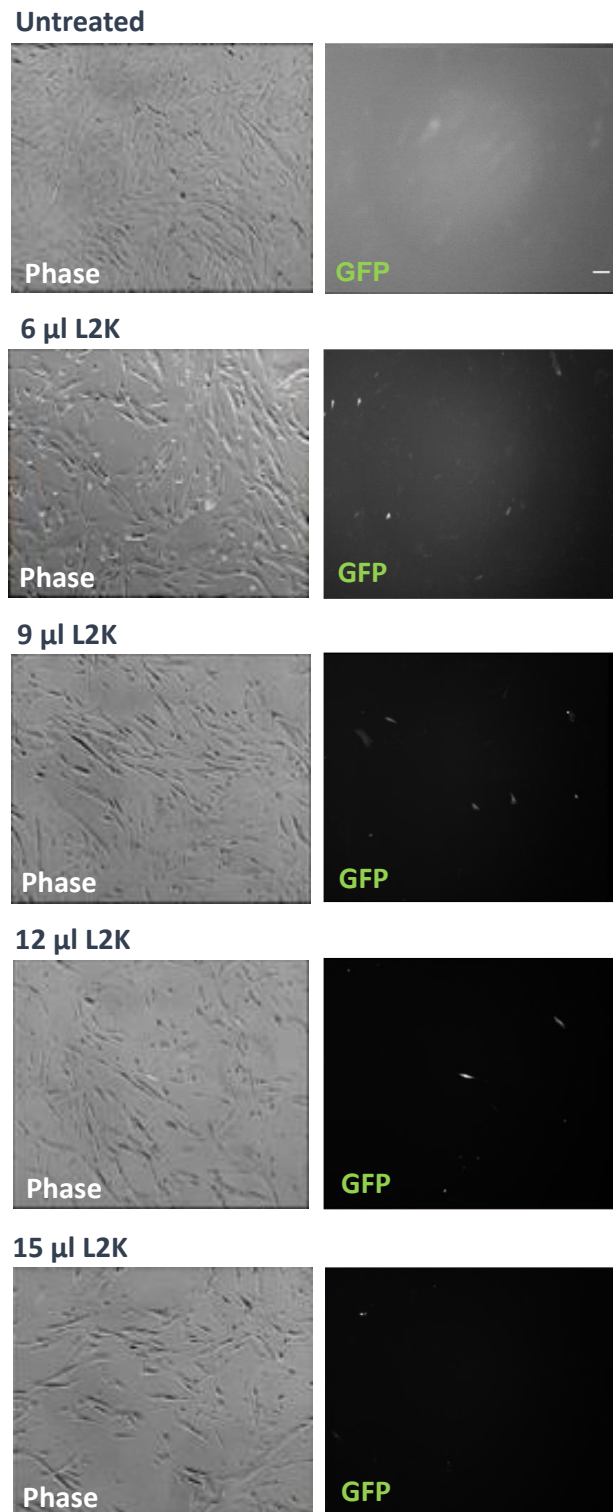


Figure 5.12. Lipofectamine transfection of primary myoblasts. Identification of (Lipofectamine 2000).

Increasing doses of Lipofectamine 2000 (6 μ l, 9 μ l, 12 μ l and 15 μ l) were used to deliver 2.5 μ g of the pCMV-GFP plasmid DNA amount to patient myoblasts. Transfection efficiency was low in all the tested conditions, with an average efficiency of 3.5 ± 1.62 . Among those tested, the highest percentage of GFP-positive cells was obtained with a Lipofectamine 2000 volume of 9 μ l (4.84 ± 0.21). Images were acquired at a magnification of 10X. Scale bar = 200 μ m

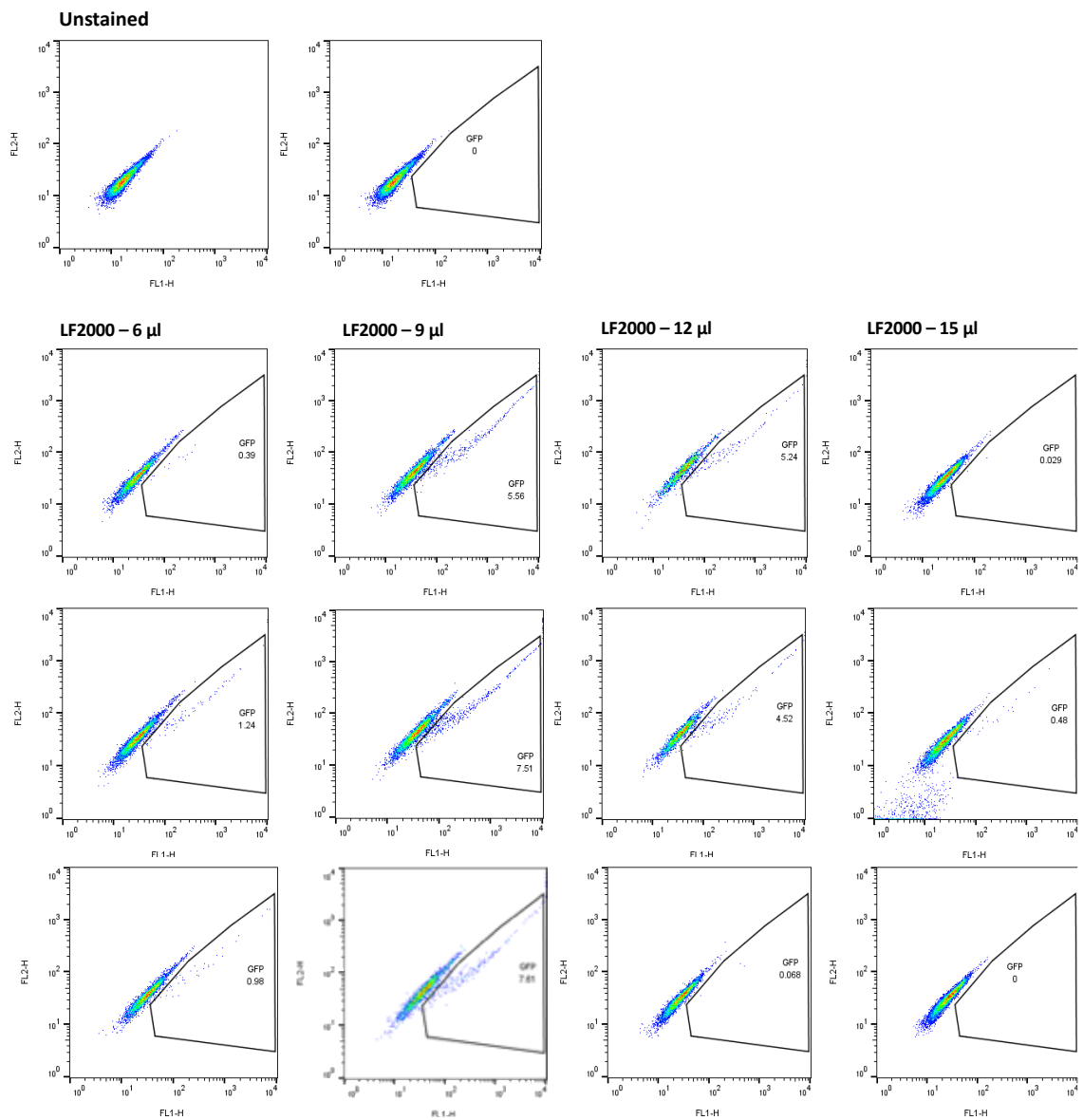


Figure 5.13. FACS analysis of patient-derived myoblasts transfected via Lipofectamine 2000 (LF2000).

DUP2 patient-derived myoblasts were transfected with four different doses of Lipofectamine 2000 ranging from 6 µl to 15 µl. The percentage of GFP-positive cells varied across conditions and was reported to be below 1% when 6 µl (0.86% +/- 0.26) and 15 µl (0.18% +/- 0.15) were used. The best Lipofectamine 2000 yield appeared to be 9 µl, which resulted in an average efficiency of 6.84% +/- 0.64, followed by 12 µl (3.27% +/- 1.62).

I also transfected cells with TurboFect, for which I used 2 µg of plasmid DNA and increasing dosages of TurboFect from 4 µl to 8 µl (section 2.2.11.c). By microscopic analysis this reagent seemed to result in a greater transfection efficiency as, on average, 11.14% +/- 0.74 of the cells expressed GFP (**Fig.5.14**) (**Table 5.3**). However, FACS analysis gave a different outcome, as it showed a similar efficiency to cells transfected via Lipofectamine 2000: only one condition (8 µl TurboFect) resulted in a slightly lower (5.3% +/-0.83) transfection efficiency compared to Lipofectamine 2000 (**Fig.5.15**), although with a better toxicity profile (visually cell death was estimated as 20-30%). Finally, not even GeneJuice transfection (section 2.2.11.d) produced a better result. Its efficiency was estimated around 2% both by microscopic (2.51% +/- 0.43) (**Fig.5.14**) and FACS analysis (1.89% +/- 0.18) (**Fig.5.15**) (**Table 5.3**).

Transfection reagent	Lipofectamine2000	TurboFect	GeneJuice
4 µl	N.A.	12.33 +/-0.63	N.A.
6 µl	2.47 +/-0.08	12.42 +/-3.06	2.51% +/- 0.43
8 µl	N.A.	10.15 +/- 0.36	N.A.
9 µl	4.85 +/- 0.21	N.A.	N.A.
12 µl	4.99 +/- 0.35	N.A.	N.A.
15 µl	1.89 +/- 0.42	N.A.	N.A.

Table 5.3. Myoblasts transfection efficiency assessed by microscopic analysis.

Considering that I could not rely on any of the transfection reagents I tested, I considered nuclear electroporation, as this technique worked in patient-derived fibroblasts (section 2.2.12.a).

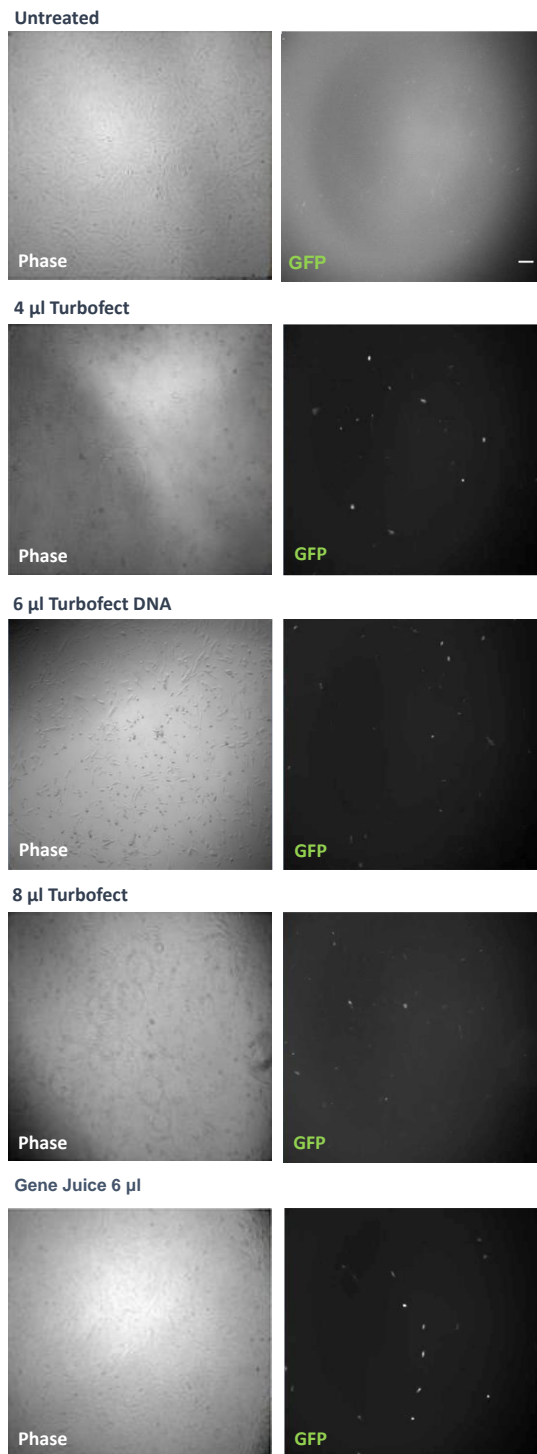


Figure 5.14. Identification of the best transfection condition (TurboFect and GeneJuice).

TurboFect transfection: Increasing doses of TurboFect (4 µl, 6 µl and 8 µl) were tested, while the DNA amount (2 µg) was kept constant. Transfection efficiency resulted to be low in all the tested conditions, with an average efficiency of 11.14% \pm 0.74. Among those tested, the highest percentage of GFP-positive cells was obtained with a TurboFect volume equal to 6 µl (12.42 \pm 3.06). GeneJuice was tested in parallel, by using a volume of reagent equal to 6 µl. The average efficiency of transfection meant as the number of GFP-positive cells over the total cell pool was 2.51% \pm 0.43. Images were acquired at a magnification of 10X. Scale bar = 200 µm

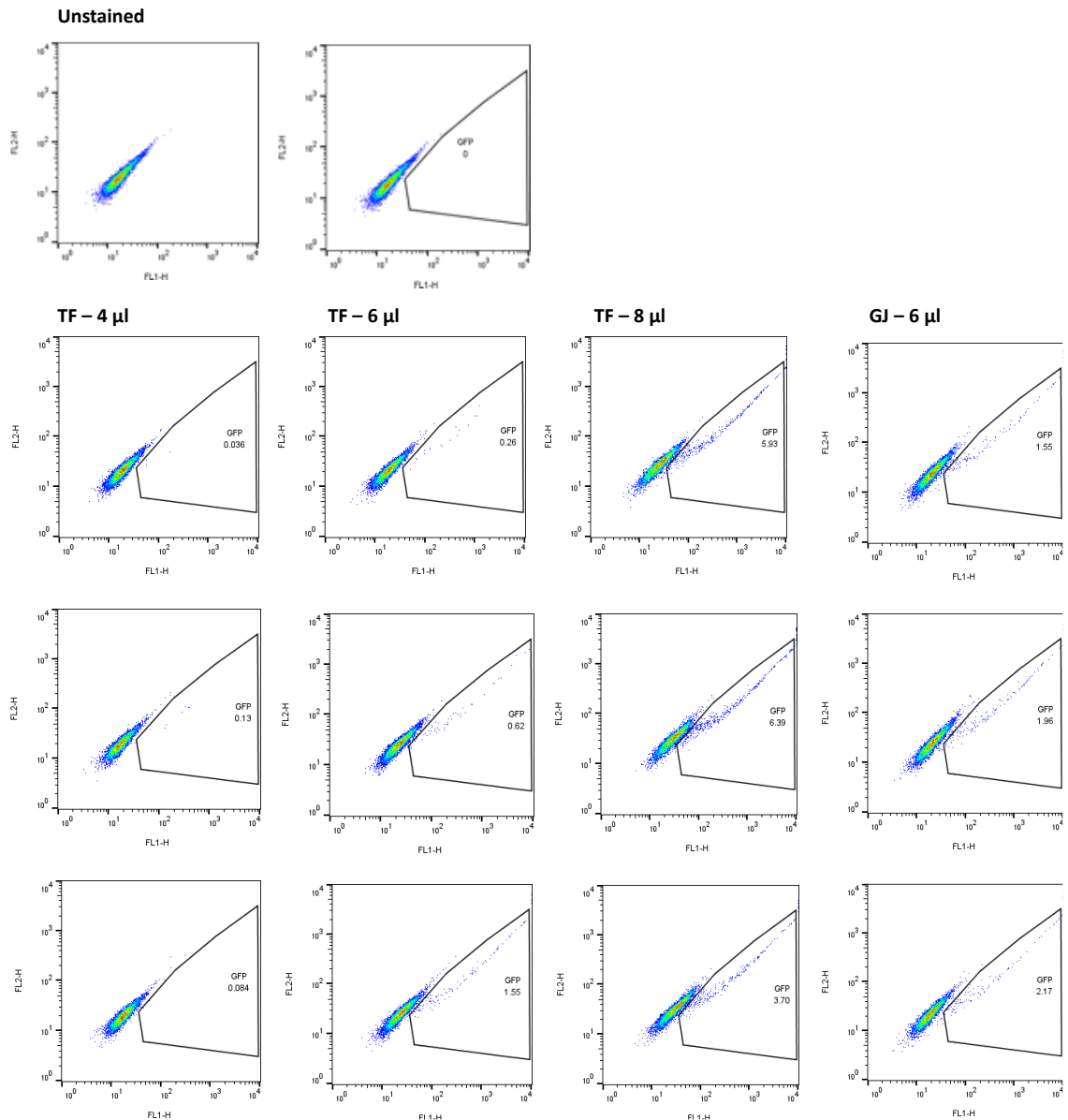


Figure 5.15. FACS analysis of patient-derived myoblasts transfected via TurboFect (TF) and GeneJuice (GJ).

DUP2 patient-derived myoblasts were transfected with three different doses of TurboFect ranging from 4 μ l to 8 μ l and, in parallel, with 6 μ l of the GeneJuice transfection reagent. All the TurboFect conditions resulted in a GFP expression below 1% (4 μ l TurboFect = 0.083% \pm 0.02, 6 μ l TurboFect = 0.71% \pm 0.29) with the exception of the transfection performed with 8 μ l of reagent (5.3% \pm 0.83). Also GeneJuice transfection resulted in an efficiency of transfection of less than 2% (1.89% \pm 0.18).

b) Myoblast electroporation: Amaxa versus NEON

To deliver CRISPR/Cas9 to patient-derived myoblasts, I tested the Amaxa electroporation condition which I used in primary fibroblasts (P022) (section 2.2.12.a), as this was shown to work also for human myoblasts⁵²⁷. I tested four DNA amounts (1 µg, 2 µg, 3 µg and 4 µg) of pCMV-GFP plasmid I used to test Lipofectamine 2000, TurboFect and GeneJuice transfection. For each DNA amount, I performed the electroporation two times (biological replicate). However, differently from fibroblasts, the electroporation of patient-derived myoblasts with the Amaxa device was ineffective, as by microscope analysis almost no fluorescent cells was detected (0.67 % +/- 0.12) with 4 µg of DNA, while only a background fluorescence similar to that observed in untreated cells was seen with the other experimental conditions (**Fig.5.16.**). Since this method was even less efficient than transfection, electroporated cells were discarded and not analyzed by FACS.

I therefore decided to electroporate DUP2 myoblasts using a different device (NEON) provided by our collaborators from the Translational Myology section of our Institute. Unlike Amaxa, NEON allows the customization of electroporation parameters, including the number and duration of electric pulses as well as the voltage. This should result in a better control of the electroporation process⁵²⁸.

I tested the parameters routinely used by our collaborators (i.e. 1 pulse of 1400 volts of a duration of 20 milliseconds, Dumonceaux and Mariot, personal communication). As these conditions were known to work in primary myoblasts, I electroporated 5×10^5 DUP2 myoblasts not only with pCMV-GFP plasmid, but also with the previously used CRISPR/Cas9 plasmids (CRISPR0 and CRISPR2σ) (section 2.2.12.b). Negative controls were represented by the untreated cells and cells electroporated with CRISPR0. I did this experiment three times (biological replicates). The outcome I obtained with NEON was considerably different than that I got from the previous methods. In fact, 48 hours post-electroporation, the mean number of fluorescent cells identified by microscopic analysis accounted for 33 % +/- 2.06 (CRISPR0), 23,10 % +/- 0.57 (CRISPR2σ) and 39.11% +/- 0.38 (GFP) of the total population of cells attached to the plate after the electric shock (**Fig.5.17.**).

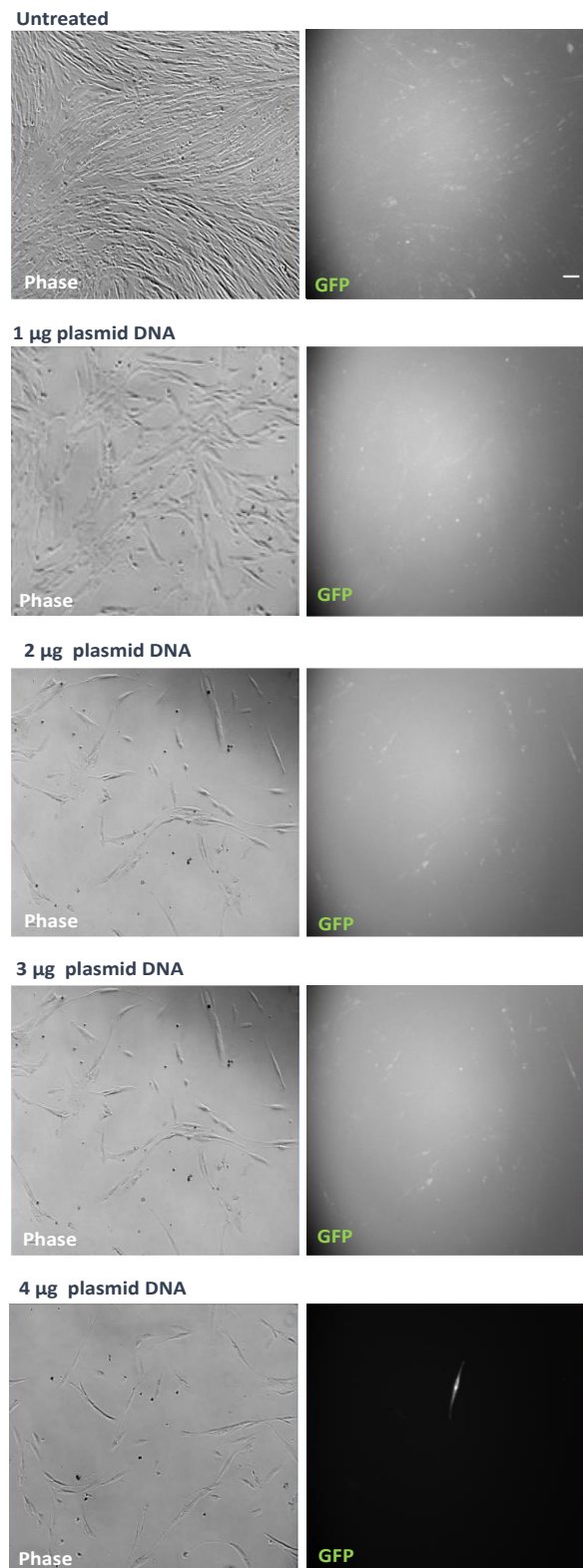


Figure 5.16. Electroporation of CRISPR2 σ in DUP2 patient-derived myoblasts (Amaxa device).

The P022 electroporation program of the Amaxa device was used in combination to increasing amount (from 1 μ g to 4 μ g) of the pCMV-GFP plasmid. Only the use of 4 μ g resulted in detectable GFP fluorescence after 48 hours, even though at a very low efficiency (0.6 % \pm 0.12). Images were acquired at a magnification of 10X. Scale bar = 200 μ m

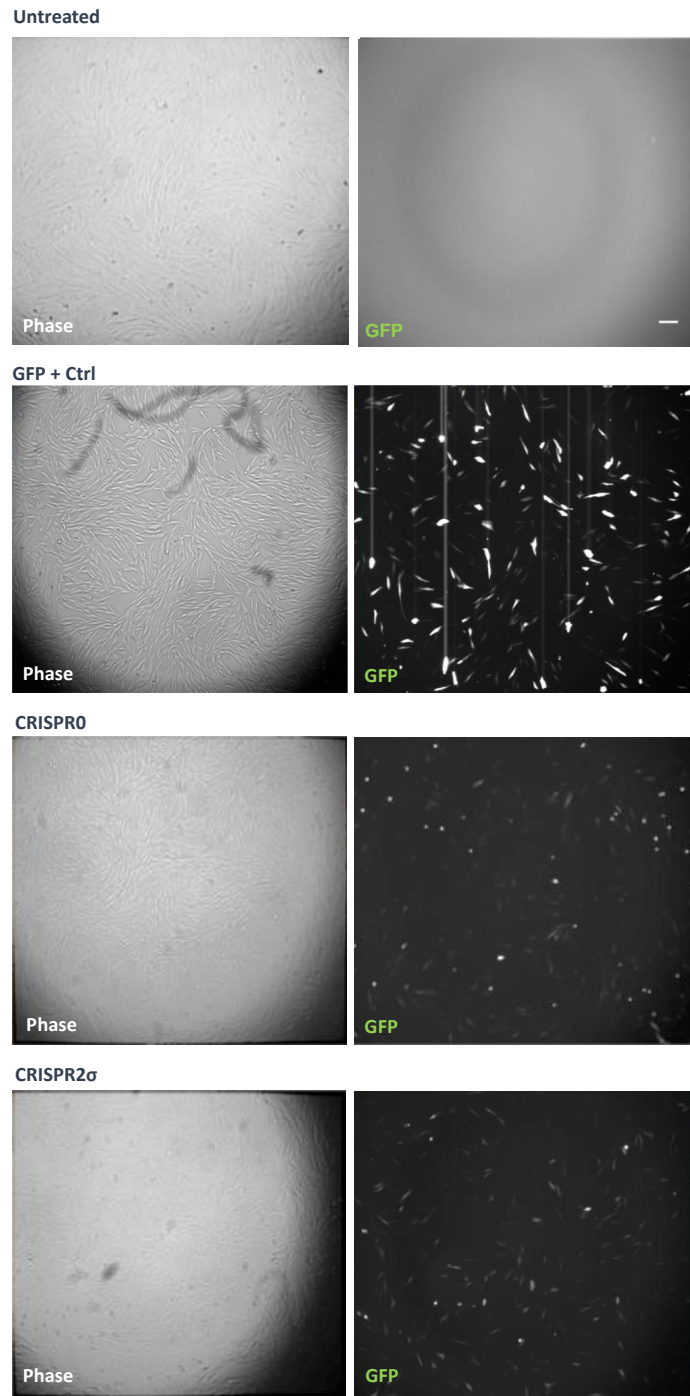


Figure 5.17. Electroporation of CRISPR2 σ in DUP2 patient-derived myoblasts (NEON device).

Custom electroporation parameters (1400 V, 1 pulse, 20 milliseconds) of the NEON device were used to electroporate 1 μ g of the pCMV-GFP plasmid, CRISPR0 and CRISPR2 σ plasmids. NEON electroporation proved to be the most efficient non-viral methods among those tested as, on average, more than 25% of cells appeared to express GFP 48 hours post electroporation 33% \pm 2.06 (CRISPR0), 23,10% \pm 0.57 (CRISPR2 σ) and 39.11% \pm 0.38 (GFP+ Ctrl). Images were acquired at a magnification of 10X. Scale bar = 200 μ m

b.i) Verification of CRISPR/Cas9-mediated cleavage at genomic level

Considering the high number of EGFP-positive DUP2 cells resulting from NEON electroporation, I expanded these cells (section 2.2.6) and used a part of them for DNA extraction (section 2.3.1), without isolating cells expressing CRISPR/Cas9. I amplified the CRISPR/Cas9 target region and performed T7 assay to assess the extent of genomic cleavage (sections 2.3.4 and 2.3.7). However, I could not detect cleaved bands of the expected size (511 bp and 217 bp) in the samples expressing CRISPR2 σ . I observed that the banding pattern observed in these sample resembled the one detected in negative controls (i.e. non-treated cells and cells electroporated with CRISPR0)(**Fig.5.18**). This outcome suggested the need of isolating EGFP- (and therefore only Cas9)-expressing cells to detect the genomic cleavage.

I repeated NEON electroporation with the aim of isolating the total population of cells expressing Cas9. I confirmed the presence of EGFP-positive cells by microscopic analysis and isolated them by FACS. FACS analysis showed a mean percentage of EGFP-positive cells of 27.57 +/- 0.75 (the highest I observed in DUP2 myoblasts via FACS) (**Fig.5.19**). However, I could not sufficiently amplify sorted DUP2 cells up to the point where I should have induced them towards terminal differentiation, as they became senescent due to the excessive passaging. I did this experiments two more times (biological replicates) but I encountered the same issue in each experimental replicate. Therefore, I decided to repeat this experiment using DUP2i cells, i.e. DUP2 myoblasts (carrying the duplication of *DMD* exons 3-16) that were immortalized by Dr. Vincent Mouly⁵²⁹.

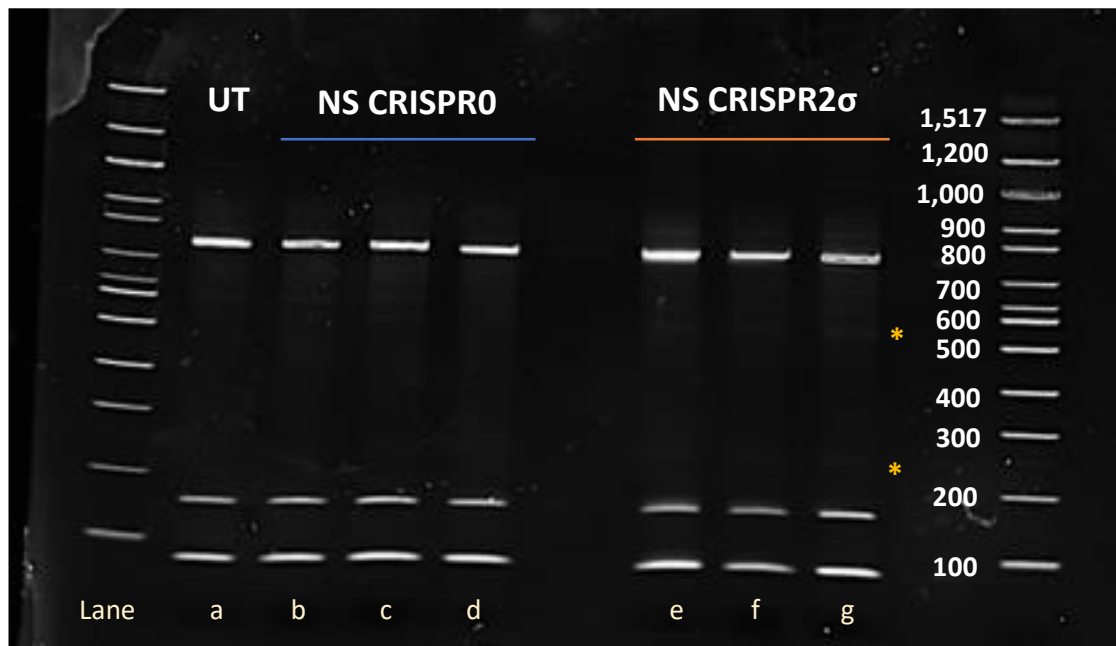


Figure 5.18. T7 assay on electroporated DUP2 patient-derived myoblasts (NEON).

T7 assay was performed on DNA derived from the total pool of DUP2 patient-derived myoblasts electroporated with CRISPR0 and CRISPR2 σ . Lanes a-d represent the amplicons derived from non-treated cells (a) and cells electroporated with CRISPR0 (b-d), respectively, where the full-length band of 728 bp was detected. Lanes e-g are instead representative of samples electroporated with CRISPR2 σ , in which cleaved bands having a molecular weight of 511 bp and 217 bp (*) were expected. However, these bands were only slightly visible in lane g (*). Genomic cleavage was therefore considered as absent. The 100 bp ladder from NEB was run alongside the samples.

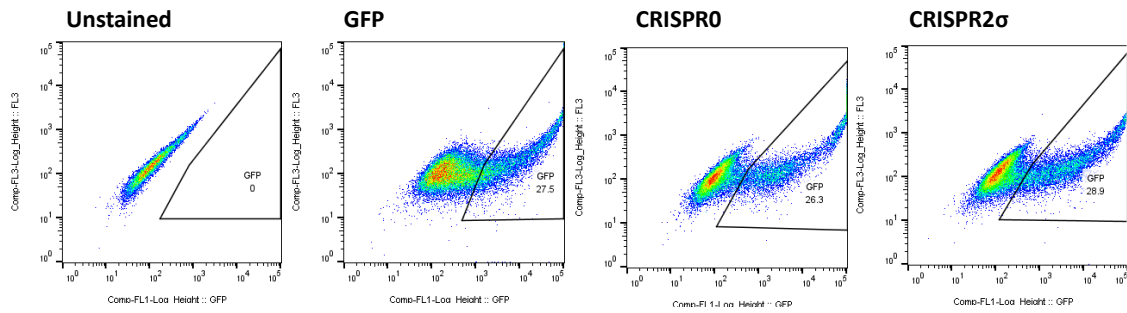


Figure 5.19. FACS analysis of patient-derived DUP2 myoblasts electroporated by NEON.

DUP2 patient-derived myoblasts were electroporated with 1 μ g of pCMV-GFP plasmid (GFP) as well as CRISPR0 and CRISPR2 σ plasmids, which co-express CRISPR/Cas9 components and EGFP. The total EGFP-positive cell population was isolated for each sample.

I repeated NEON electroporation on DUP2 cells immortalized by Vincent Mouly (named DUP2i cells). This experiment was done three times (biological replicates). Controls were represented by untreated cells and cells electroporated with CRISPR0 (negative controls). I verified EGFP expression by microscopic analysis after 48 hours and isolated fluorescent cells via FACS. I noticed that, despite the percentage of EGFP-positive DUP2i cells being lower than in electroporated DUP2 cells (**Fig.5.20**), NEON was still the best of all methods tested for CRISPR/Cas9 delivery. The mean efficiency for each electroporated plasmids was 11.5 % \pm 0.6% (CRISPR0) and 13.27% \pm 1.6% (CRISPR2 σ) (**Fig.5.20**). These values were at least twofold higher than what I achieved in all the viral and non-viral delivery methods previously tested (**Fig.5.5**, **Fig.5.13** and **Fig.5.15**).

I expanded sorted cells and collected their DNA, that was amplified and analyzed via T7 assay (section 2.3.7). This time, I observed both cleaved bands of the expected size of 511 bp and 217 bp in all samples expressing CRISPR2 σ , but not in amplicons derived from untreated cells and cells expressing CRISPR0 (negative controls) (**Fig.5.21**). The cleavage efficiencies quantified for each replicate were 20.57 %, 21.46 % and 24.64%, which resulted in a mean value of 22.2% \pm 1.23.

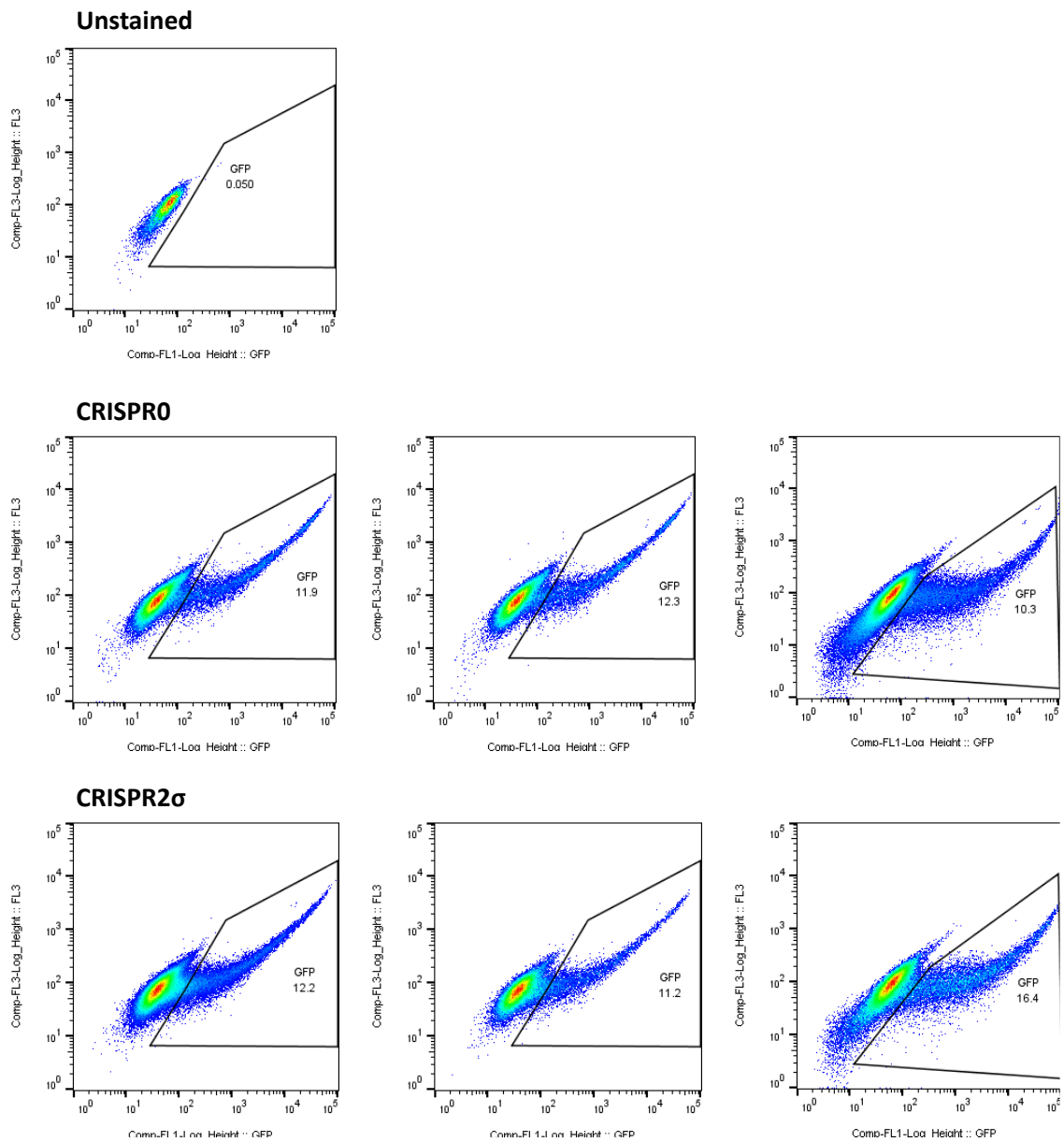


Figure 5.20. FACS analysis of immortalized DUP2i myoblasts electroporated by NEON.

DUP2i myoblasts were electroporated with 1 μ g CRISPR0 and CRISPR2 σ plasmids, which co-express CRISPR/Cas9 components and EGFP. The mean electroporation efficiency, assessed by FACS as the percentage of GFP-expressing cells, was 11.5% \pm 0.6 (CRISPR0) and 13.27% \pm 1.6 (CRISPR2 σ).

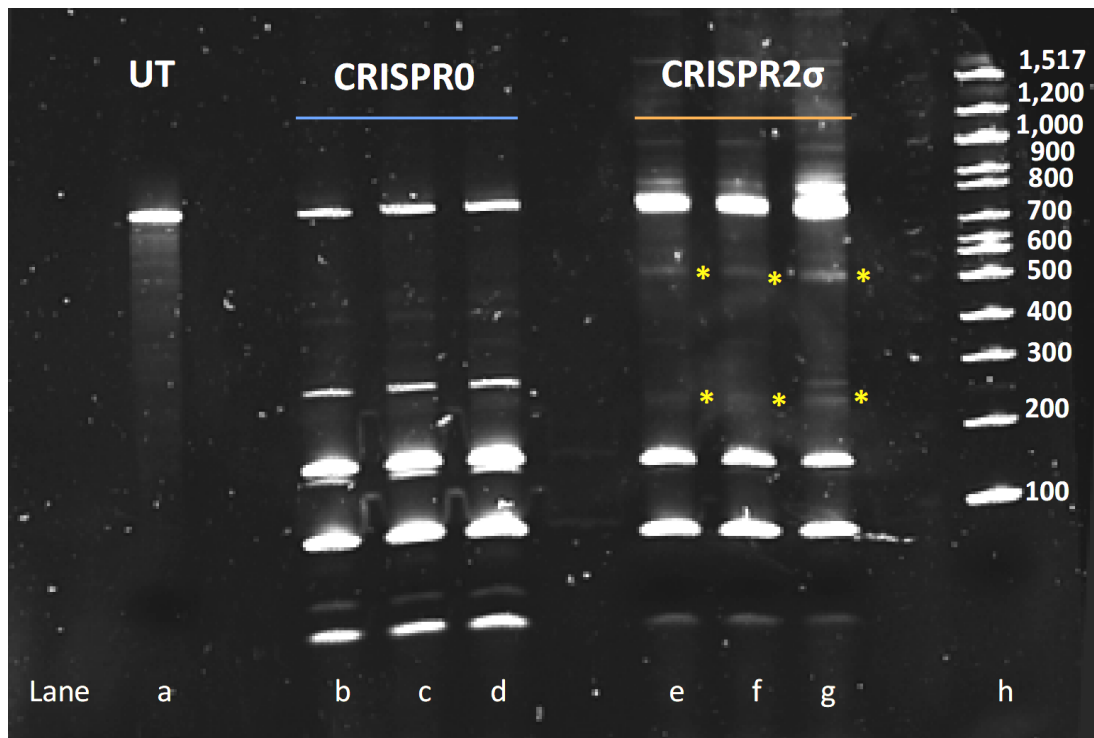


Figure 5.21.T7 assay on electroporated DUP2i myoblasts (NEON) expressing CRISPR/Cas9.

T7 assay was performed on DNA derived from the pool of GFP-positive DUP2i myoblasts electroporated with CRISPR0 and CRISPR2 σ . Lanes a-d represent the amplicons derived from non-treated cells (a) and cells expressing CRISPR0 (b-d), where the full-length band of 728 bp was detected. Lanes e-g represent instead cells expressing CRISPR2 σ . Cleaved bands of the expected molecular weight (511 bp and 217 bp, *) were visible in these samples (yellow stars). The mean editing efficiency in cells treated with CRISPR2 σ was 22.2% \pm 1.23. The 100 bp ladder from NEB was run alongside the samples (lane h).

b.ii) Verification of dystrophin transcript repair following CRISPR/Cas9 activity in DUP2i cells

Considering the positive outcome of the T7 assay, I expanded further the sorted cells kept in culture and seeded them in one well of a six-well plate and one 10 cm² dish. These were cultured with the M2 medium (section 2.2.6) for one week to be induced towards terminal differentiation. RNA and protein were then extracted from resulting myotubes (sections 2.3.10 and 2.3.19).

I assessed the restoration of dystrophin transcript by doing a qPCR (section 2.3.13.b) on cDNA obtained as in section 2.3.11.b. I studied a triplicate of each cDNA sample derived from untreated cells as well as cells derived from each electroporation with CRISPR0 and CRISPR2 σ . Upon qPCR data analysis based on the $\Delta\Delta C_t$ method (section 2.3.14), I saw that the expression of dystrophin region spanning the duplicated exons 8 and 9 normalized to the anchor point (dystrophin exon 20) was diminished in cells expressing CRISPR2 σ , but not in cells expressing CRISPR0 and in untreated cells (**Fig.5.26.a**). This reduction did not reach statistical significance when Kruskal-Wallis test was applied (p. value = 0.055). However, I considered that a non-statistically significant reduction of the mutated transcript could still result in a detectable amount of wild-type dystrophin, as I previously saw in transduced cells (section 5.3.2.c). Therefore, I could not use this outcome to infer the missed dystrophin restoration at protein level.

b.iii) Verification of dystrophin protein repair following CRISPR/Cas9 activity

I ran wet blot as detailed in section 2.3.22.b to assess dystrophin restoration in cells expressing CRISPR2 σ . As a positive control I used protein derived by the differentiated H2K 2B4 cells⁵³⁰, while as a negative control I included protein from non-treated myotubes obtained from DUP2i cells. I confirmed that CRISPR/Cas9 treatment resulted in dystrophin correction, as I saw two bands on the PVDF membrane. These represented the higher molecular weight dystrophin carrying the in frame duplication and the smaller wild-type dystrophin. Densitometric analysis, which I performed by measuring the ratio of intensity between the two dystrophin bands, revealed a mean dystrophin correction of 54% \pm 7.11 (almost the same as that achieved in transduced cells) (section 5.3.2.c). This result showed that even the episomal expression of CRISPR/Cas9 delivered by a non-viral method such as electroporation was able to efficiently target dystrophin and remove its duplicated region, with a success rate comparable to that achieved with lentiviral vectors (**Fig.5.22**).

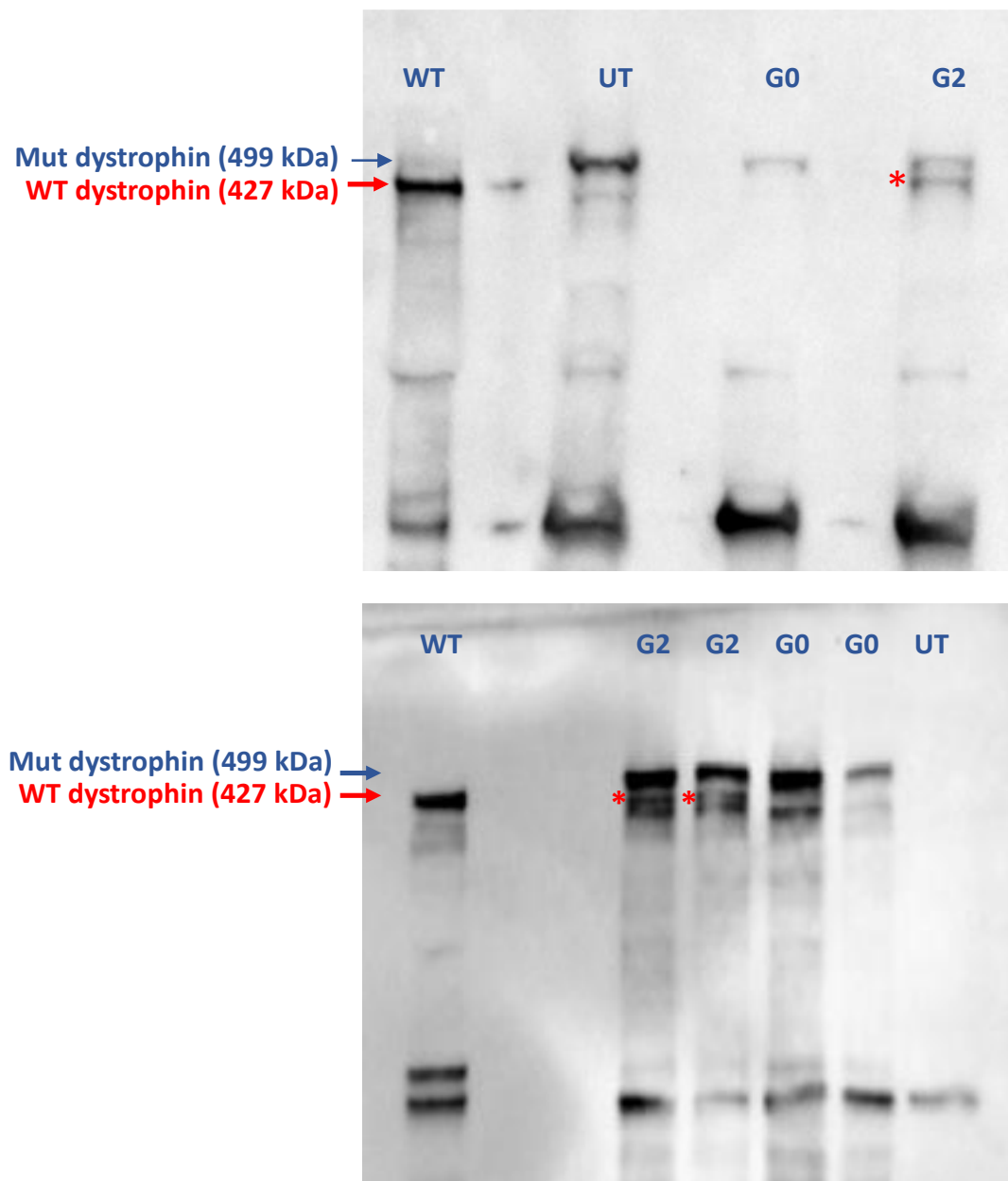


Figure 5.22. Validation of dystrophin repair in electroporated *DUP2i* myoblasts.

Western blot was performed to verify dystrophin protein correction in transduced *DUP2i* cells expressing CRISPR/Cas9. UT = protein derived from untreated cells, negative control; G0 = protein derived from cells electroporated with CRISPR0 plasmid, negative control; G2 = protein derived from cells electroporated with CRISPR2 σ plasmid; WT = protein derived from the murine H2K 2B4 cells, expressing wild-type dystrophin, positive control. Wet blot followed by signal amplification allowed to observe wild-type dystrophin isoform in cells electroporated with CRISPR2 σ plasmid but not in the negative controls, where only the higher molecular weight dystrophin isoform was detected. The percentage of corrected dystrophin for each experimental replicate was 68.4%, 48.8% and 45.7%. Red stars and arrow indicate wild-type dystrophin (427 kDa), while blue arrow indicate the mutated isoform (499 kDa).

c) Contribution of Cas9 expression levels to dystrophin editing

Once I demonstrated that the transient CRISPR/Cas9 expression can restore dystrophin duplications, I next studied the contribution of cell populations expressing different levels of Cas9 in triggering the repair. I did this to understand if a threshold in the level of CRISPR/Cas9 expression was needed to edit dystrophin. I electroporated DUP2i cells with CRISPR/Cas9 plasmids three times (biological replicates) (section 2.2.12.b). However, unlike previously, I gated GFP-positive cells according to their fluorescence intensity, which is representative of Cas9 expression levels, and sorted them accordingly via FACS (section 2.2.14.a). Specifically, gated populations were either expressing minimal or considerable EGFP levels and were therefore named as Low and High in relation to their fluorescence (**Fig.5.23**). By FACS analysis (**Fig.5.23**), I derived the mean percentage of fluorescent cells for each plasmids. These were 8.37 \pm 0.49 (CRISPR0 Low), 3.92 \pm 0.28 (CRISPR0 High), 10.41 \pm 0.8 (CRISPR2 σ Low) and 3.60 \pm 0.39 (CRISPR2 σ High).

c.i) Assessing CRISPR/Cas9-mediated cleavage efficiency at genomic level

I expanded each sorted pool representative of the Low and High fluorescent cells and used part of the cells for DNA extraction (section 2.3.1). I assessed genomic targeting via T7 assay (section 2.3.7). I observed cleaved bands of the expected size (511 bp and 217 bp) in both Low and High DUP2i cell populations, expressing low and high levels of CRISPR2 σ (**Fig.5.24**) but not in untreated cells and cells expressing CRISPR0 (**Fig.5.25**). However, as expected, the percentage of editing was reduced in the Low (19.8 \pm 1.12) compared to the High population (28.0 \pm 3.01). This allowed me to hypothesize that higher level of Cas9 expressing cells contribute to genome editing more than cells with a barely detectable Cas9 expression.

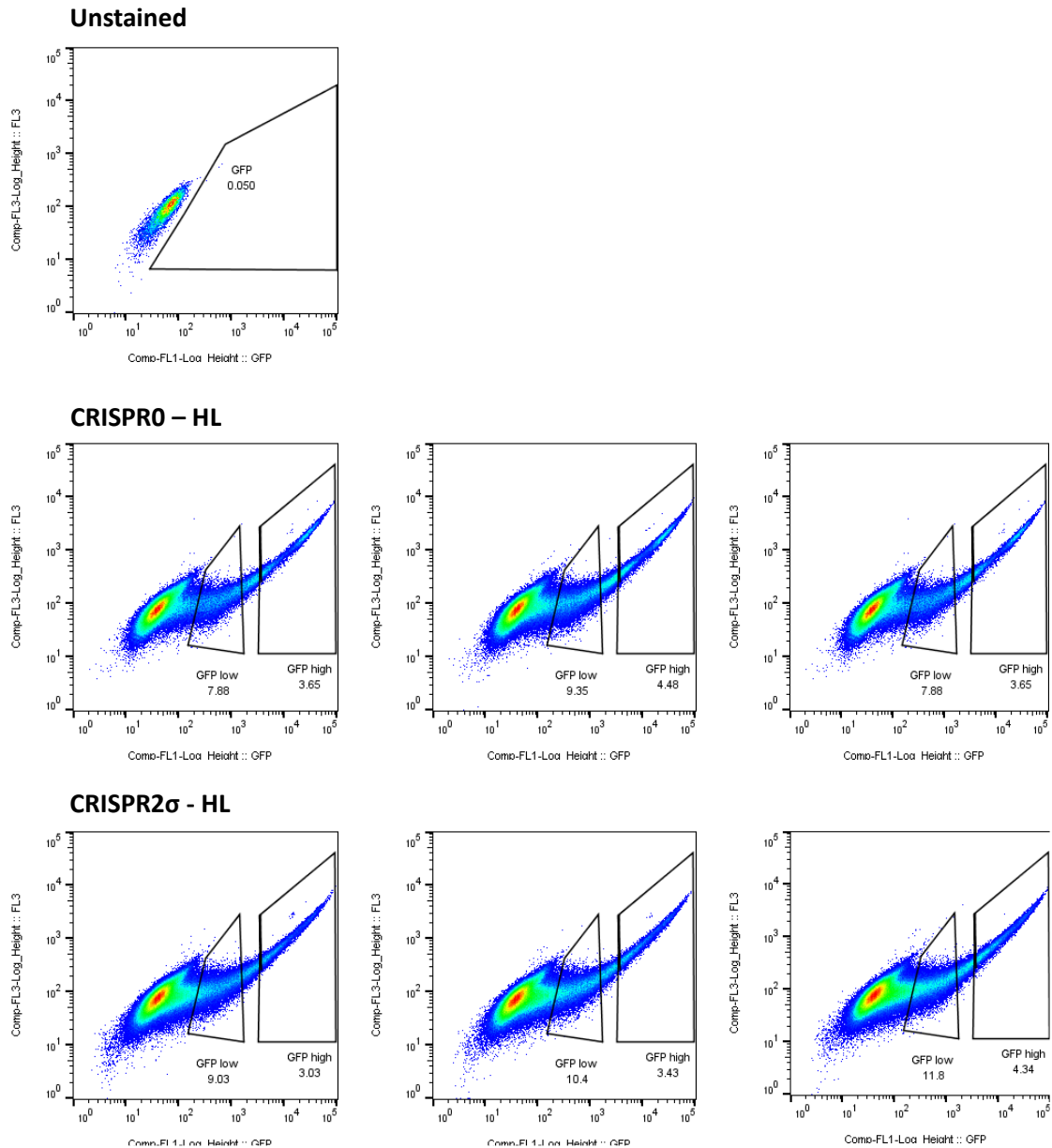


Figure 5.23. FACS analysis of DUP2i myoblasts electroporated by NEON.

DUP2i myoblasts were electroporated with 1 μ g CRISPR0 and CRISPR2 σ plasmids, which co-express CRISPR/Cas9 components and EGFP. GFP-positive cells were gated according to their intensity of fluorescence (high or low). The mean percentage of GFP-positive cells for each gated population were 8.37 \pm 0.49 (CRISPR0 Low), 10.41 \pm 0.8 (CRISPR2 σ Low), 3.92 % \pm 0.28 (CRISPR0 High) and 3.60% \pm 0.39 (CRISPR2 σ High).

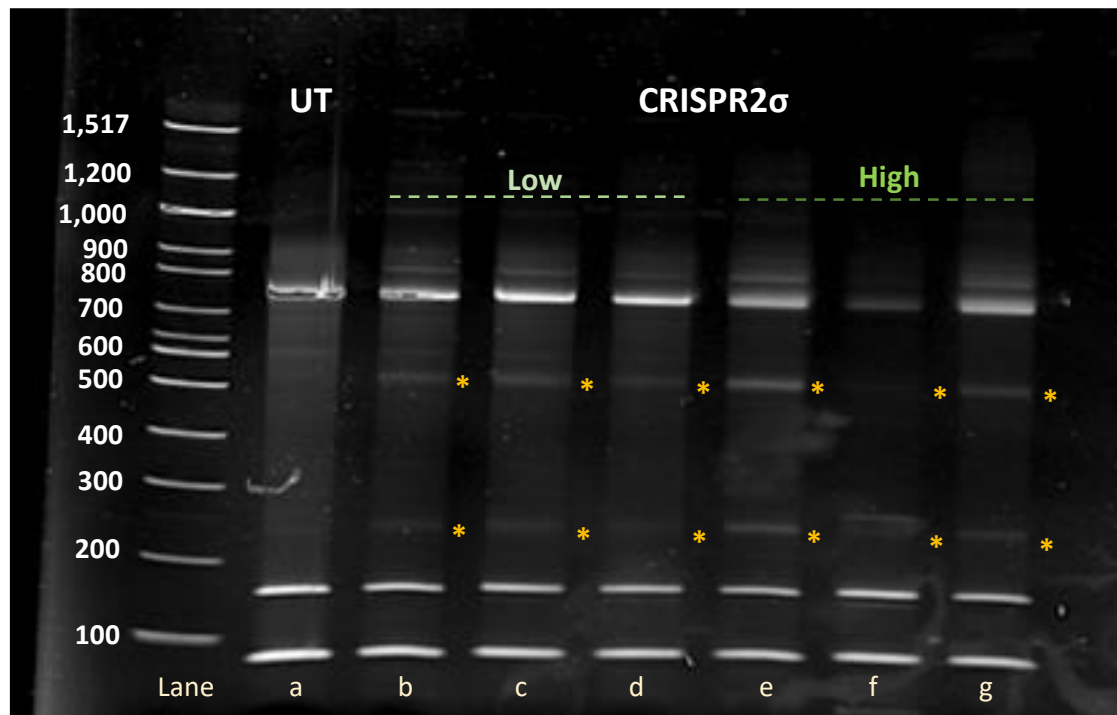


Figure 5.24. T7 assay on electroporated DUP2i myoblasts (NEON) expressing low and high CRISPR/Cas9 levels.

T7 assay was performed on DNA derived from the pool of DUP2i patient-derived myoblasts electroporated with CRISPR2 σ and expressing either low or high GFP levels. Lanes a-d represent the amplicons derived from non-treated cells (a) and cells expressing low levels of CRISPR2 σ (b-d), respectively. Lanes e-g represent instead cells strongly expressing CRISPR2 σ . Cleaved bands of the expected molecular weight (511 bp,*) were visible in all samples apart for the one corresponding to untreated cells, as expected. However, the average editing efficiency was lower in cells barely expressing the nuclease (Low population, mean efficiency = 19.8 \pm 1.12) as compared to cells with an high expression of CRISPR/Cas9 (High population, mean efficiency = 28.0 \pm 3.01). A 100 bp ladder (NEB) was run alongside the samples.

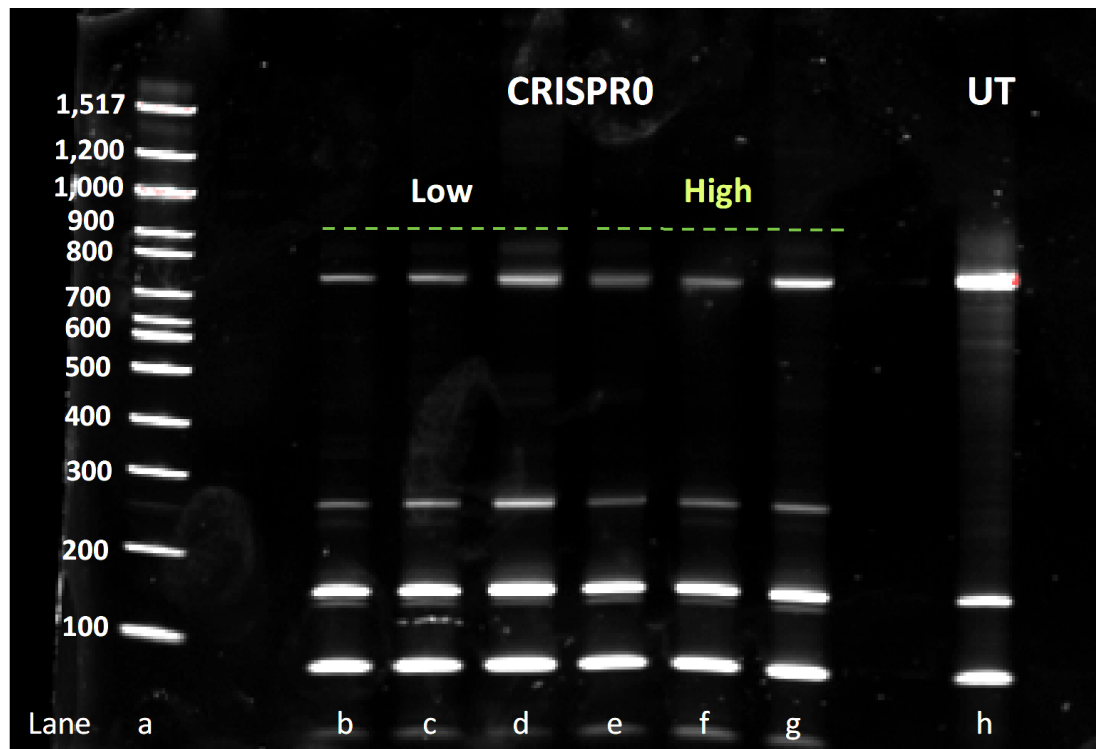


Figure 5.25. T7 assay on electroporated DUP2i myoblasts (NEON) expressing low and high CRISPR/Cas9 levels.

T7 assay was performed on DNA derived from the pool of DUP2i patient-derived myoblasts electroporated with CRISPR0 and expressing either low or high GFP levels. Lane h represents the amplicons derived from non-treated cells. Cells expressing low levels of CRISPR2 σ are shown in lanes b-d, while lanes e-g represent instead cells strongly expressing CRISPR2 σ . As expected, in all samples only the full length amplicon whose size was expected to be 728 bp was detected. A 100 bp ladder (NEB) was run alongside the samples (lane a).

c.ii). Verification of dystrophin transcript repair following CRISPR/Cas9 activity

Sorted cells were further expanded and allowed to fuse into myotubes from which I extracted RNA and protein following the protocols mentioned in sections 2.3.10 and 2.3.19. I analyzed dystrophin transcript via qPCR (section 2.3.13.b) and analyzed the data by the $\Delta\Delta C_t$ method (section 2.3.14). I observed a reduction of the mutated transcript in both populations. This reduction was greater in the cell pool expressing high levels of Cas9 than the pool where the nuclease expression was low. Moreover, both data sets (Low and High populations) were statistically significant according to Kruskal-Wallis test (p.values equal to 0.028 (High) and 0.0005 (Low)). I applied the Mann-Whitney test to each experimental set analyzed by Kruskal-Wallis test and I confirmed that only the population expressing CRISPR2 σ statistically differ from the control samples (i.e. untreated cells and cells expressing CRISPR0 (**Fig.5.26.b and 5.26.c**)). Moreover, I observed that the mean transcript correction of the Low and High cell populations expressing CRISPR2 σ (0.58 \pm 0.03) concurred with the mean level of transcript correction detected in the total pool of GFP-positive cells (0.62 \pm 0.10).

c.iii) Verification of dystrophin repair following CRISPR/Cas9 activity

Finally, I ran western blot assay (section 2.3.22.b) to quantify the extent of dystrophin restoration in both the Low and High cell populations. Positive control was represented by protein extracted from myotubes originating from H2K 2B4 cells⁵³⁰, while as negative controls I used the protein obtained from non-treated samples and from samples expressing CRISPR0. Reflecting the findings at the genomic and transcript levels, wild-type dystrophin was barely detectable in cells poorly expressing Cas9, while it was clearly present in cells where the nuclease levels were higher (**Fig.5.27**). This was in agreement with the hypothesis that the levels of editing increase in line with the numbers of Cas9 copies present in the cells. The mean percentages of dystrophin correction were 25.23 \pm 0.65% and 38.6% \pm 0.88 for the Low and High populations, respectively.

However, the quantification of dystrophin bands was complicated both by their close proximity and the multitude of non-specific bands present on the PVDF membrane. The former issue was related to an inefficient protein run within the polyacrylamide gel (section 2.3.21), while the identification of non-specific bands was likely due to the over-amplification of the signal because of the detection method I used (section 2.3.22.b). Future work will be focused on repeating the western blot to confirm the outcome indicated above.

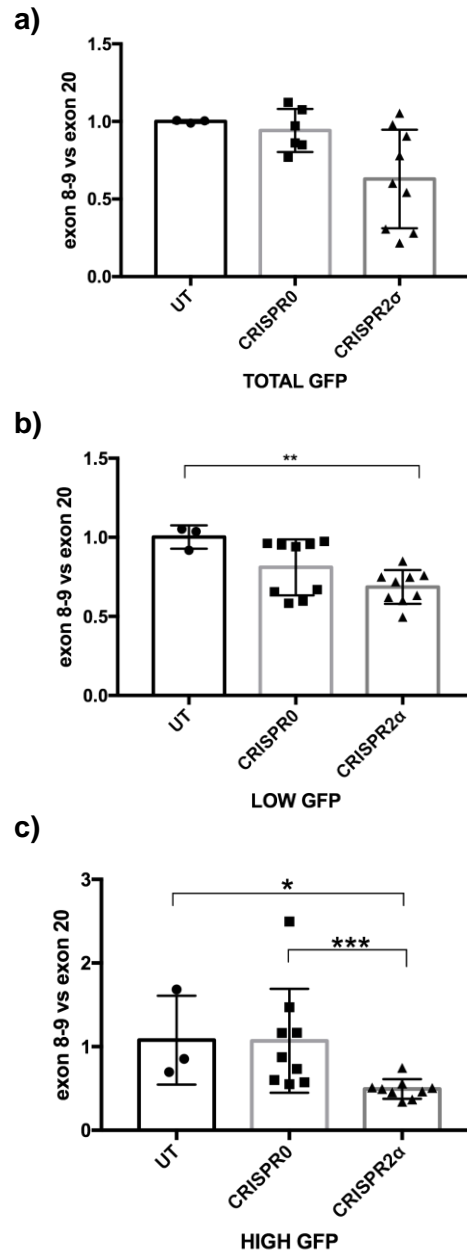


Figure 5.26. Quantification of dystrophin transcript correction in electroporated DUP2i cells.

qPCR was performed to assess the level of restored dystrophin transcript in the total pool of DUP2i cells expressing the CRISPR/Cas9 (a) and in DUP2i cells expressing low (b) and high (c) nuclease levels. UT= untreated cells (negative control); CRISPR0 = cells expressing CRISPR0 plasmid (negative control), CRISPR2σ = cells expressing CRISPR2σ plasmid. The reduction of the mutated transcript compared to each negative control reached significance only when the Low and High pools were considered singularly. a) Kruskal-Wallis test not significant ($p=0.0552$). b) Kruskal-Wallis test significant ($p=0.028$) and Mann-Whitney test significant when UT was compared to CRISPR2σ ($p=0.0091$). c) Kruskal-Wallis test significant ($p=0.0005$), Mann-Whitney test significant when both UT and CRISPR0 were compared to CRISPR2σ ($p=0.0182$ and $p=0.0008$). Error bars represent the standard deviation.

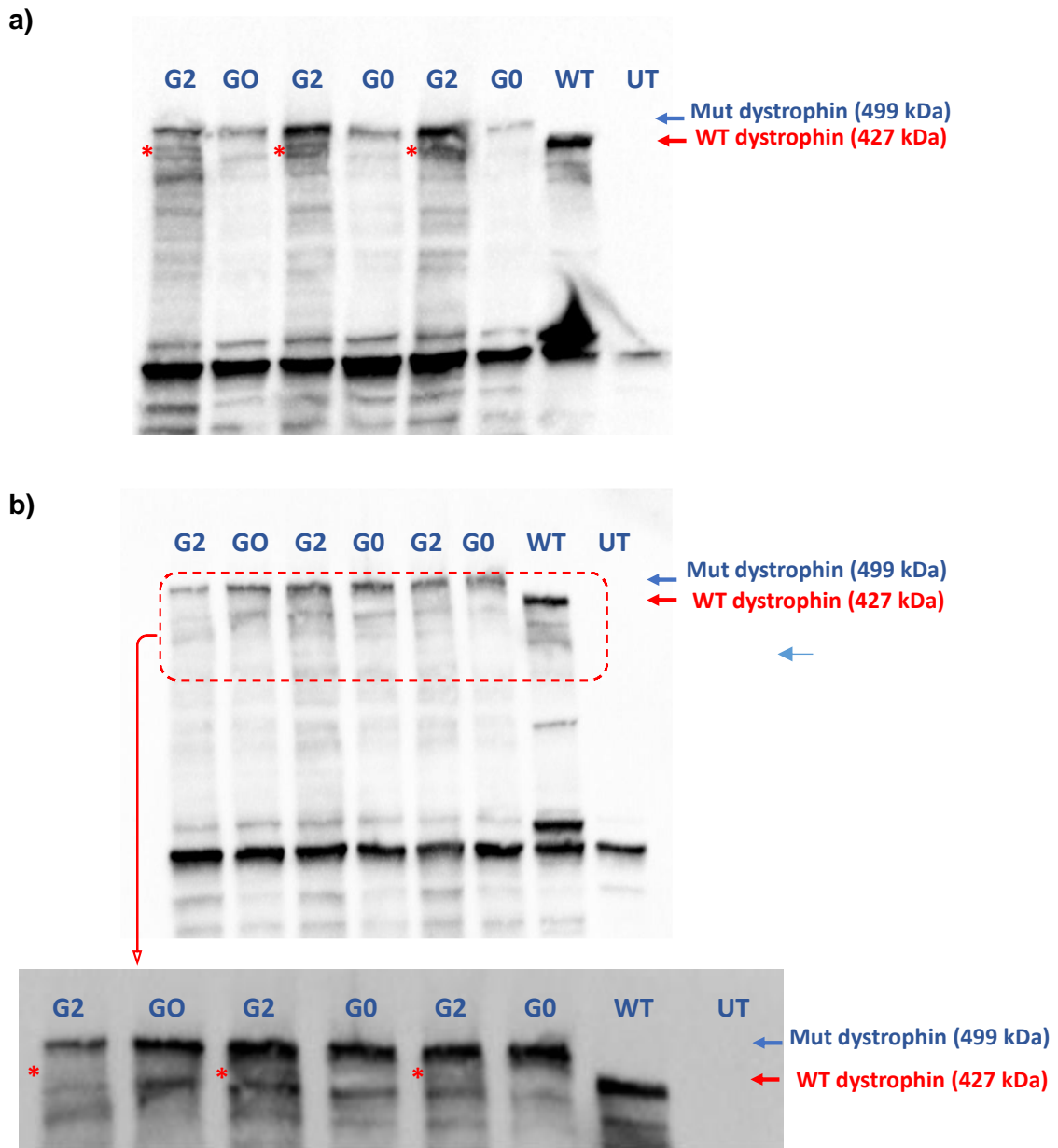


Figure 5.27. Validation of dystrophin repair in electroporated DUP2i myoblasts expressing either High or Low CRISPR/Cas9 levels.

Western blot was performed to verify dystrophin protein correction in transduced DUP2i cells expressing different levels of CRISPR/Cas9. UT = protein derived from untreated cells (protein not detected); G0 = protein derived from cells electroporated with CRISPR0 plasmid, negative control; G2 = protein derived from cells electroporated with CRISPR2 σ plasmid; WT = protein derived from the murine H2K 2B4 cells, expressing wild-type dystrophin, positive control. a) represents High CRISPR/Cas9 expression, while b) represents Low CRISPR/Cas9 expression. Wet blot followed by signal amplification allowed to observe a greater level of wild-type dystrophin isoform in a) (High) compared to b) (Low). Also, wild-type dystrophin appeared in cells electroporated with CRISPR2 σ plasmid but not in the negative controls, where only the higher molecular weight dystrophin isoform (499 kDa) was detected. The percentages of corrected dystrophin for each experimental replicate were 40.3%, 38.2% and 37.3% (High panel) and 24.0%, 26.2% and 25.5% (Low panel). Red stars and arrow indicate wild-type dystrophin (427 kDa), while blue arrow indicate the mutated isoform (499 kDa).

5.4 DISCUSSION

In this chapter, I demonstrated that a single CRISPR/Cas9 nuclease can correct duplications arising in the *DMD* gene and restore the mutated dystrophin protein in patient-derived myoblasts.

To investigate dystrophin repair at protein level, I took advantage of a patient-derived myoblast cell line that was extremely myogenic (90% fusion index). This allowed me to obtain terminally differentiated myotubes, essential for dystrophin protein analysis. The chosen cell line had a duplication spanning *DMD* exons 3-16, which does not disrupt the *DMD* reading frame (in-frame mutation), (**table 2.1**, section 2.2.2). This feature was advantageous as it allowed to assess the production of the duplicated dystrophin isoform expression, as an ideal internal control. When I had to assess the extent of dystrophin repair at protein level, therefore, I could quantify the extent of dystrophin correction referring to the mutated dystrophin isoform. Moreover, I could use the detection of the mutated dystrophin via western blot as a way to assess the reliability of the protocol I used to study dystrophin protein.

I initially confirmed the efficacy of my approach by using integrating lentiviral particles expressing CRISPR, and later validated dystrophin correction by means of a non-viral approach based on nuclear electroporation.

I chose to test lentiviral vectors first, as I considered them advantageous for proving the efficiency of my study design *in vitro*. In fact, these vectors can target almost every cell type and integrate in their genome, thus ensuring a continuous expression of the nuclease⁵³¹. I demonstrated that the low level of CRISPR/Cas9 nuclease expressed *in vitro* by these vectors was sufficient to target and restore the duplication within *DMD* gene, thus allowing the production of a wild-type dystrophin protein. The tropism of lentiviral vectors for both dividing and non-dividing tissues would make these vectors a good option to target muscle *in vivo*^{266,532,533}. However, their random genome integration *in vivo* might result in adverse effects which include the risk of oncogenesis⁵³⁴. Based on this consideration, I decided to repeat my experiments with an alternative non-viral strategy as I thought that, for future in-vivo applications, the use of other vectors whose transgenes mainly remain episomal upon host infection (such as adenoviral vectors and AAVs^{535,536}) would be preferable.

I initially explored if a transient CRISPR/Cas9 expression could have been sufficient to restore dystrophin. To understand this, I looked for a non-viral method to deliver and express the nuclease *in vitro* in patient-derived myoblasts.

The choice of a non-viral method was challenged by the low efficiency of plasmid transfection to primary myoblasts. In fact, in contrast to immortalized cells^{530,537}, very few studies managed to obtain high (above 60%) efficiency of transfection in primary human myoblasts with non-viral methods^{538,539}. I screened three commercial transfection reagents (Lipofectamine2000, TurboFect and GeneJuice), which all resulted in a variable but poor transfection efficiency (around 5% or lower). This was even lower than the efficiency shown in work done on mouse myoblasts^{540,541} (around 10%).

Considering that experimental parameters (such as the presence/absence of serum during transfection and the ratio of transfection reagent and DNA) might affect the outcome of the experiment⁵²⁴, I performed all the transfections in serum-free medium (i.e. OptiMem), and screened different volumes of transfection reagents. I thought that the serum-free transfection should have increased the transfection efficiency (as shown by Dodds *et al.*⁵⁴²) especially for the Lipofectamine transfection. In fact, serum protein (as albumin) can interact with either the DNA or the liposome formulation. This would have reduced the recognition and binding of the membrane receptors through which liposomes enter the cell⁵⁴³ and consequently the efficiency of transfection. In cells transfected with Lipofectamine, my results were consistent with those of Jackson *et al.*⁵⁴⁴ who transfected murine myoblasts with different Lipofectamine:DNA ratios (2:1, 3:1 and 5:1). I confirmed that a Lipofectamine:DNA ratio ranging from 3:1 to 4:1 is more efficient (around 5%) than the other two ratios of about 2:1 and 6:1, as these resulted in an extremely low transfection efficiency (less than 1%) (section 5.3.3.a).

I next investigated nuclear electroporation, which I previously explored in the attempt to deliver CRISPR/Cas9 to primary fibroblasts. I used two different devices (Amaxa and NEON). I observed a considerable amount of cell death (about 80-90%) following Amaxa electroporation and therefore I did not pursue this strategy further. However, there was less cell death when I electroporated the cells with NEON (30-40%).

I observed that the cell death I saw following NEON electroporation was higher than what I saw in cells transduced with lentiviral vectors at 0.4 MOI. Despite that, the percentage of electroporated cells expressing CRISPR/Cas9 was on average three times higher than that observed in transduced cells (**Figures 5.5 and 5.20**). I therefore decided to use NEON as a non-viral method to test the efficiency of the episomal Cas9 expression to correct dystrophin duplications.

Based on the data I obtained from primary fibroblasts, for both CRISPR/Cas9 transduction and electroporation, I decided to rely on transgenes co-expressing Cas9 and EGFP to facilitate the selection of the cells expressing the nuclease. I realized that, in myoblasts, the isolation of the total pool of EGFP-positive cells was sufficient to

observe genomic cleavage via T7 assay. This result differed from what I observed in fibroblasts, as in these cells I could not detect genomic targeting unless I isolated the top 20% of EGFP-positive cells. I hypothesized that this difference might have been possibly due to a more open chromatin conformation in myoblasts (in which a low level of dystrophin transcript has been observed⁵⁴⁵) as opposed to fibroblasts. A more open chromatin, in fact, would be more permissive in allowing Cas9-mediated genome targeting⁵⁴⁶.

Interestingly, when I compared the outcomes obtained upon cells transduction and electroporation, I saw that despite its episomal expression, CRISPR/Cas9 delivered via NEON electroporation corrected dystrophin to a similar extent (54% +/-7.11) as the constitutive Cas9 expression of integrating lentiviral vectors (59.9% +/- 6.74). I noticed that, other than for their targeting efficiency, the two methods differed in the extent of CRISPR/Cas9 expression. The fluorescence profile obtained by FACS data, in fact, allowed me to observe that none of the transduced cells expressed high levels of the CRISPR/Cas9 transgene (**Fig.5.5**), whereas following NEON electroporation, I could clearly distinguish very bright cells (i.e. cells expressing high levels of the nuclease) (**Fig.5.19**). By analyzing the level of dystrophin correction obtained in cells expressing low and high levels of the nuclease, I observed that the amount of wild-type dystrophin protein was higher in cells expressing the higher level of Cas9, while it was barely detectable in cells in which Cas9 was poorly expressed. Nevertheless, when I analyzed transduced cells expressing Cas9 at low levels, I observed an high level of dystrophin repair. Based on these results I hypothesized that the results obtained by CRISPR/Cas9 transduction were likely the result of the constant nuclease expression in the cells. This effect was therefore equivalent as having a transient but higher CRISPR/Cas9 expression. I therefore concluded that, as long as CRISPR/Cas9 is expressed at a sufficiently high level, it does not need to be constantly present within the cell for targeting the genome. For this reason, I would recommend the use of a non-integrating transgene that could sustain a strong but transient CRISPR/Cas9 expression *in vivo*.

To summarize, the experiments detailed in this chapter show that:

- Low doses of integrating lentiviral vectors can trigger CRISPR/Cas9 expression in primary human myoblasts.
- Among the non-viral alternatives tested for CRISPR/Cas9 delivery in patient-derived myoblasts, the transient transfection reagents have a limited efficiency (below 5%), whereas NEON electroporation had the highest efficiency (about 15%).

- Single CRISPR/Cas9 nucleases expressed either by integrating lentiviral vectors or NEON electroporation can efficiently target small dystrophin duplications and restore dystrophin protein expression.
- Even a transient CRISPR/Cas9 expression is sufficient to restore dystrophin in a patient carrying a duplication *in vitro*, to a similar extent as its continuous expression.

As I demonstrated the applicability of a single nuclease approach in repairing small duplications, the next step (focus of Chapter 6) will be the investigation of its applicability in repairing larger dystrophin duplications.

Chapter 6. Addressing single CRISPR/Cas9 nucleases to large dystrophin duplications in patient-derived induced pluripotent stem cells

6.1 INTRODUCTION

6.1.1. The advent of induced pluripotent stem cells

Induced pluripotent stem cells (iPSCs) are a type of pluripotent stem cells derived from terminally differentiated cells. iPSCs were created in 2006 by Takahashi and Yamanaka who, introducing specific transcription factors (Oct4, Sox2, cMyc and Klf4) into skin-derived cells, managed to reprogram these cells into pluripotent cells behaving like embryonic stem cells (ESCs)⁵⁴⁷.

Human iPSCs (hiPSCs) can be established from a wide variety of somatic cells (like skin-derived fibroblasts, keratinocytes, blood and urine-derived cells)^{547–551}. Similarly to ESCs, iPSCs can be expanded indefinitely and differentiated into almost any type of cell. However, in contrast to ESCs that require blastocyst isolation, iPSCs can be obtained from the easily accessible cells mentioned above with only minimally invasive procedures. Since individual-specific iPSCs can be obtained to generate multiple cell lineages *in vitro*, they represent a great potential for the development of personalized cell therapies aimed to replace lost or damaged tissues. The use of iPSCs as a therapy would also eliminate the risk of immune response often exacerbated upon the introduction of non-self biological molecules.

Considering these advantages, in the years following the initial iPSC generation, the published protocol was replicated by many laboratories and optimized in the attempt to improve both efficiency and safety of the process^{552–556}. In fact, one of the main challenges that comes with the iPSCs technology is the low rate of somatic cell reprogramming (below 1%), hypothesized to be due to the epigenetic memory retained by the original somatic cell⁵⁵⁷. Among the vectors initially screened, integrating vectors that could be excised from the genome via a Cre-Lox cassette generated iPSCs with an efficiency varying from 0.5% to 1.5%⁵⁵⁸, which was even lower by using adenoviral vectors⁵⁵⁶. Nowadays, the reprogramming of fibroblasts is commonly achieved using the Sendai virus (a non-integrating vector completely removed from transduced cells after ten passaging steps)^{559,560}, or by the expression of the reprogramming factors such as mRNAs⁵⁶¹.

Newer protocols also are devoid of the c-Myc transcription factor, as the use of this proto-oncogene increased the risk of tumorigenicity that would be detrimental for the clinical application of the iPSC technology. Teratoma formation in immunodeficient mice injected with iPSCs, considered as a check-point to ensure the success of iPSC derivation⁵⁶², was shown to be diminished or absent in case of iPSCs generated using a protocol devoid of c-Myc^{563,564}.

In culture, the generation of iPSCs is generally associated with the identification of embryoid bodies. These are three-dimensional ball-shaped structures similar to those obtained upon culturing human ESCs and are formed by a core of mitotically active cells surrounded by terminally differentiated cells⁵⁶⁵. Once embryoid bodies are formed *in vitro*, iPSCs have to be assessed through molecular check-points which include the analysis of their karyotype, qPCR and either immunocytochemistry or FACS analysis. Karyotype analysis is needed to screen for chromosomal abnormalities and rearrangements that can be generated during the reprogramming⁵⁶⁶, while the other techniques are needed to verify the expression of specific nuclear or surface pluripotency markers such as Nanog, Sox2, Oct4, SSEA4 or TRA-1-60^{567–569}, for example. Once iPSCs have been established, specific differentiation protocols can be finally put in place to drive iPSCs towards the cell type of interest.

6.1.2. Mimicking developmental myogenesis to induce myogenic specification in iPSCs

To induce myogenic differentiation of iPSCs and obtain myogenic precursors, the conditions present during embryonic development have to be recapitulated starting from the mesodermal lineage which, as reviewed by Kodaka *et al.*, also generates cardiomyocytes, osteocytes, endothelial and vascular smooth muscle cells and dermis⁵⁷⁰. The region from which myogenic cells arise during embryonic development is the dermomyotome, the epithelial layer which defines the dorsal part of the somites originated from the mesoderm, underneath the ectoderm⁵⁷¹. The dermomyotome is characterized by the expression of the Pax3 and Pax7 transcription factors. This population of cells is maintained by the secretion of bone morphogenetic protein 4 (BMP4)^{572,573}. Upon receiving signals generated by the neural tube and notochord such as Wnt and Sonic hedgehog (Shh)^{574,575}, and the transient activation of Notch signalling⁵⁷⁶, Pax3/Pax7-positive muscle precursor cells are converted to cells expressing MyoD and Myf5 transcription factors which mark myoblast specification⁵⁷⁷. Myoblasts then migrate below the dermomyotome, giving rise to the myotome. The subsequent expression of Mrf4 and myogenin lead to the myoblasts fusing with each other to form terminally-differentiated muscle fibers⁵⁷⁷.

Some of the approaches developed to induce iPSCs to undergo myogenic differentiation rely on the overexpression of myogenic factors (such as MyoD)^{578–581}. The forced expression of MyoD was pioneered by Dekel *et al.*⁵⁸² in early '90s and later replicated by Warren *et al.*⁵⁶¹ and Tanaka *et al.*⁵⁸³ who, by using either a conventional transfection or a transposon-based approach to allow a stable genomic integration, respectively obtained 40% and up to 90% of myogenic cells expressing markers of terminal differentiation, such as myogenin and myosin heavy chain.

An alternative to MyoD overexpression is represented by the forced expression of Pax3 and Pax7 transcription factors, markers expressed by the myogenic progenitors⁵⁸⁴. However, these markers are also present in precursors of the neuronal lineage⁵⁸⁵ and therefore a further purification step is needed to obtain a pure myoblasts lineage.

Once the myogenic differentiation is triggered, iPSCs have to be cultured with media supplemented by amino-acids, growth factors (such as FGF, IGF, HGF) and other stage-specific factors whose concentrations mimic the embryonic gradients of morphogens (such as Wnt, Notch ligands, Shh, inhibitors of BMP4)^{579,586}.

Xu *et al.*⁵⁸⁷ identified a cocktail able to induce robust myogenic differentiation in iPSCs. This contained an inhibitor of the glycogen synthase kinase 3 β (GSK-3 β) enzyme which, by stabilizing β -catenin, acts as an activator of the Wnt pathway⁵⁸⁸, and also included bFGF (involved in cell proliferation)⁵⁸⁹.

A similar mixture was also used by Chal *et al.*^{579,590}, who developed a transgene-free approach to induce myogenic differentiation. His approach exploits chemicals able to trigger the simultaneous activation and inhibition of Wnt- and BMP pathway, respectively, thus recreating the pre-somitic mesoderm stage observed *in vivo*. One of such molecules is CHIR99021, a potent activator of the canonical Wnt pathway and LDN193189, inhibitor of the BMP pathway.

Once the myogenic specification is initiated and myoprogenitors are obtained, terminal differentiation is generally triggered according to the protocol proposed by Borchin *et al.*⁵⁹¹, by using media containing a low concentration of serum. These media are also generally supplemented with IST (insulin/ selenite/ transferrin), a compound that aids cell toxicity by reducing the levels of reactive oxygen species.

6.2 AIMS

To demonstrate the applicability of a single CRISPR/Cas9 nuclease approach to repair a large dystrophin duplication, a patient-derived iPSC line was chosen as this cell line harbored a dystrophin duplication spanning dystrophin exons 3-41.

These cells will be used for the completion of the following chapter's aims, also represented in **Fig.6.1**:

- To electroporate CRISPR/Cas9 plasmids in DMD iPSCs cells
- To induce the myogenic differentiation of DMD iPSCs treated with CRISPR/Cas9 in order to obtain iPSC-derived myoblast-like cells in which I will require to demonstrate dystrophin repair at genomic, transcript and protein level
- To generate iPSCs-derived myoblast-like cells from untreated DMD iPSCs and electroporate them with CRISPR/Cas9 plasmids
- To assess dystrophin duplication removal in CRISPR-Cas9 treated iPSC-derived myogenic cells

Dystrophin correction in iPSCs

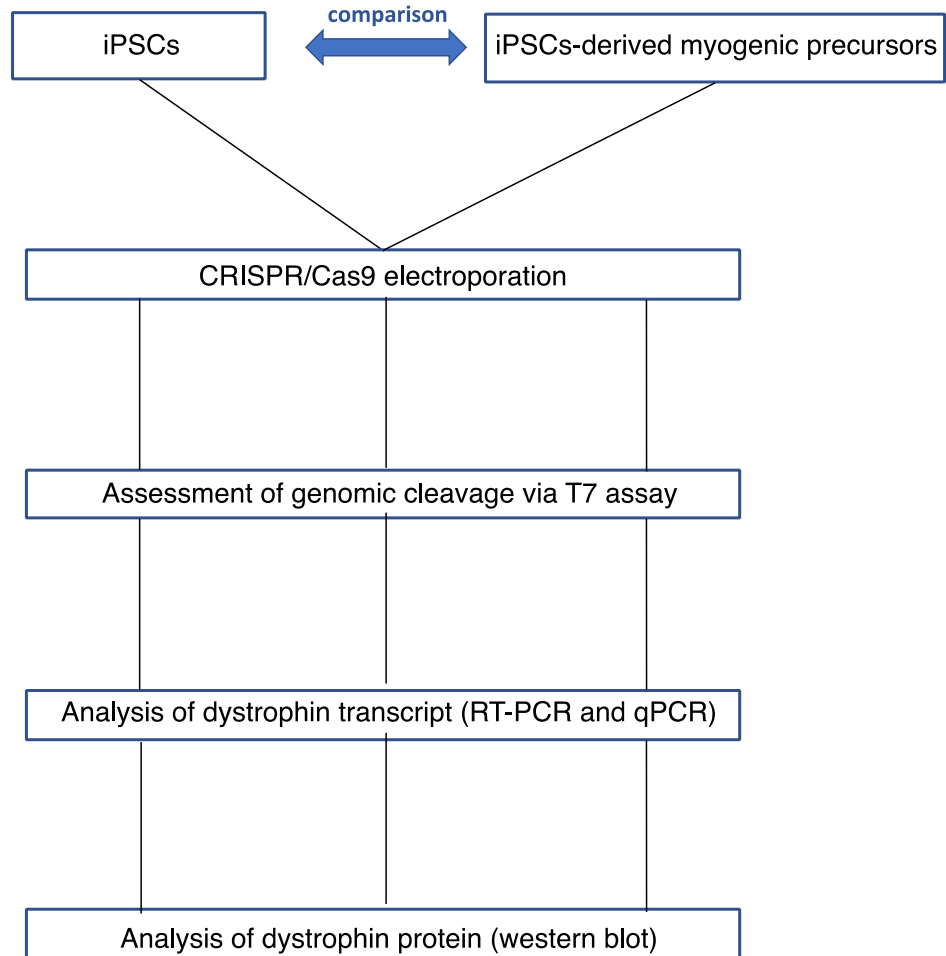


Figure 6.1. Schematic of the experimental steps required for restoring dystrophin in iPSCs.

6.3 RESULTS

6.3.1 NEON Electroporation of iPSCs with CRISPR2 σ

I cultured DUP3 iPSCs (duplication of exons 3-41) (**table 2.1**, section 2.2.2) in mTeSR1 medium (section 2.2.7) and, when they reached 5×10^5 cells, I detached them from the plate by means of the TRYPLE™ Express reagent (section 2.2.8). I electroporated these cells with the CRISPR2 σ plasmid using the NEON device. I tested three different electroporation parameters (conditions 1-3) (section 2.2.12.c) to identify one resulting in a good balance between efficiency of delivery and cell viability. I performed each electroporation twice (biological replicate). As a negative control, I used non-treated DUP3 iPSCs.

None of the conditions I tested (including condition 1 which was successful for the electroporation of patient-derived myoblasts) (section 2.2.12.b), allowed me to detect either cells viability or EGFP-expression. I observed that almost 90% of the iPSCs present in the plate 48 hours post-electroporation were dead and those that survived to the electric shock did not express EGFP (**Fig.6.2**).

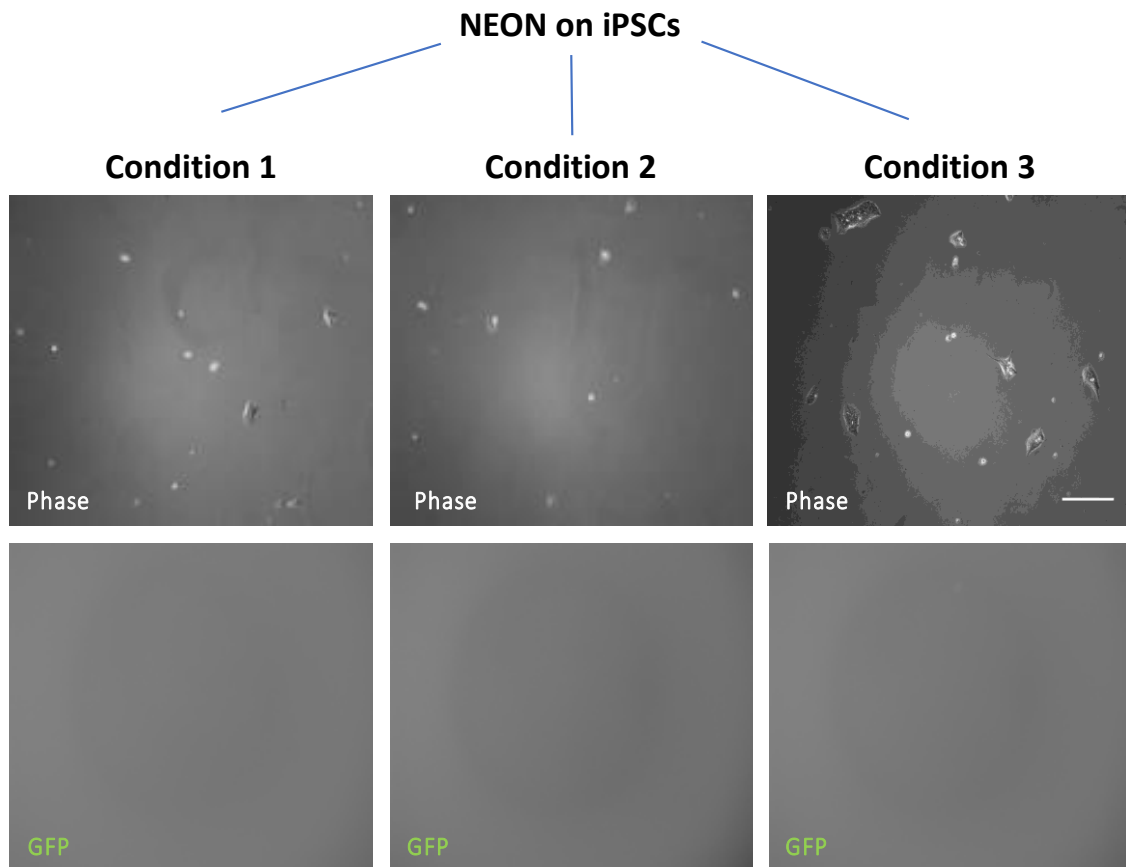


Figure 6.2. Electroporation of DUP3 iPSCs with NEON electroporation system.

A DMD iPSC line (DUP3) with an in-frame duplication of dystrophin exons 3-41 was used for electroporation of CRISPR2 σ plasmid. Cells were electroporated with the plasmid using the following conditions: 1400 V, 20 milliseconds, 1 pulse (condition 1), 1050 V, 30 milliseconds, 2 pulses (condition 2), 1100 V, 30 milliseconds, 1 pulse (condition 3). GFP fluorescence was assessed 48 hours post-electroporation by means of an Olympus IX Inverted microscope (magnification 10X). Unexpectedly, none of the tested conditions resulted in successful transfection as indicated by the death of majority of the cells and the lack of GFP expression in the remaining cells. Images were taken at a magnification of 10X. Scale bar = 500 μ m.

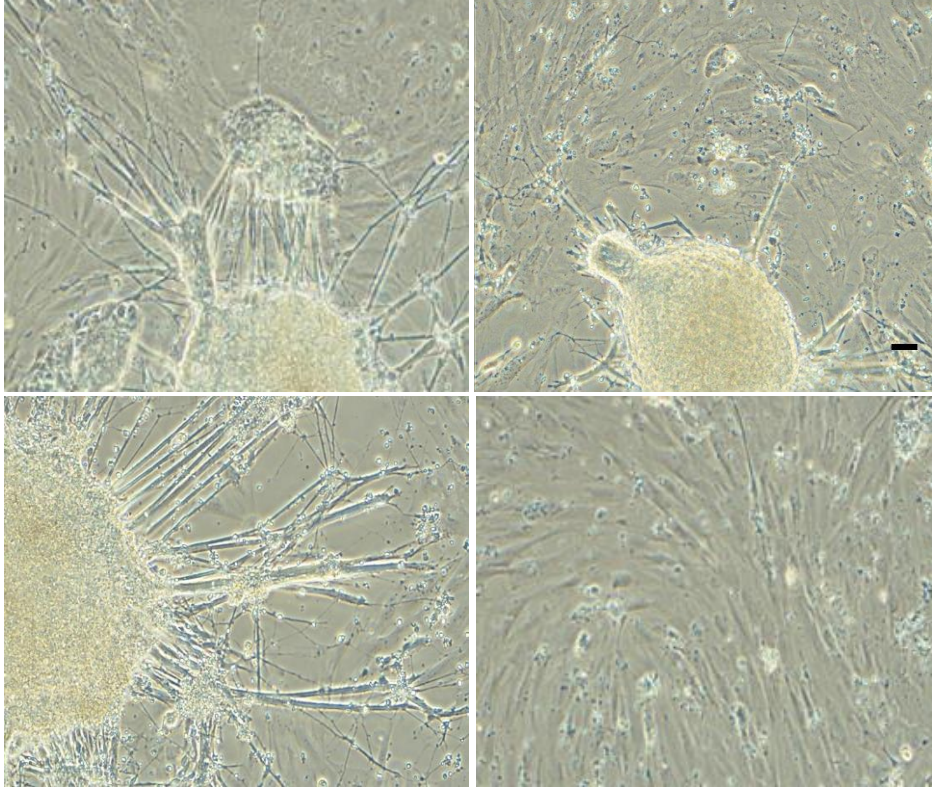
6.3.2. Differentiation of iPSCs into myoblasts

As I could not observe CRISPR/Cas9 expression after electroporating DUP3 iPSCs, I decided to induce these cells towards the myogenic lineage and repeat the NEON electroporation on the resulting myoblast-like cells, as I did in section 2.2.12.b on DUP2i myoblasts.

The protocol I used to induce DUP3 towards skeletal muscle lineage was the transgene-free protocol described by Chal *et al.*⁵⁹⁰.

I seeded DUP3 iPSCs at three plating densities: low (LD, **Fig.6.3**), medium (MD, **Fig.6.4**) and high (HD, **Fig.6.5**). I did this twice (technical replicates). Each time, I induced myogenic induction (section 2.2.7) by using the Wnt-activator CHIR99021 at a concentration of either 3 μ M or 6 μ M. I did this considering that both seeding density and CHIR99021 concentration in the culture medium affect the number of cells that initiate the differentiation program^{579,580}. Moreover, the testing of different conditions allowed me to identify the best association of density and CHIR99021 concentration for differentiating my iPSC line, as different iPSCs vary in their differentiation capacity⁵⁹². I imaged cells at the completion of differentiation protocol (day 25) and I observed that, generally, the cell population cultured with 6 μ M CHIR99021 contained more cells with the elongated myoblast-like morphology than those treated with 3 μ M CHIR99021. Within the group of cells cultured with 6 μ M CHIR99021, cells seeded at a low density visually appeared as the most homogeneous cell type (**Fig.6.3, bottom panel**).

LD 3uM CHIR



LD 6uM CHIR

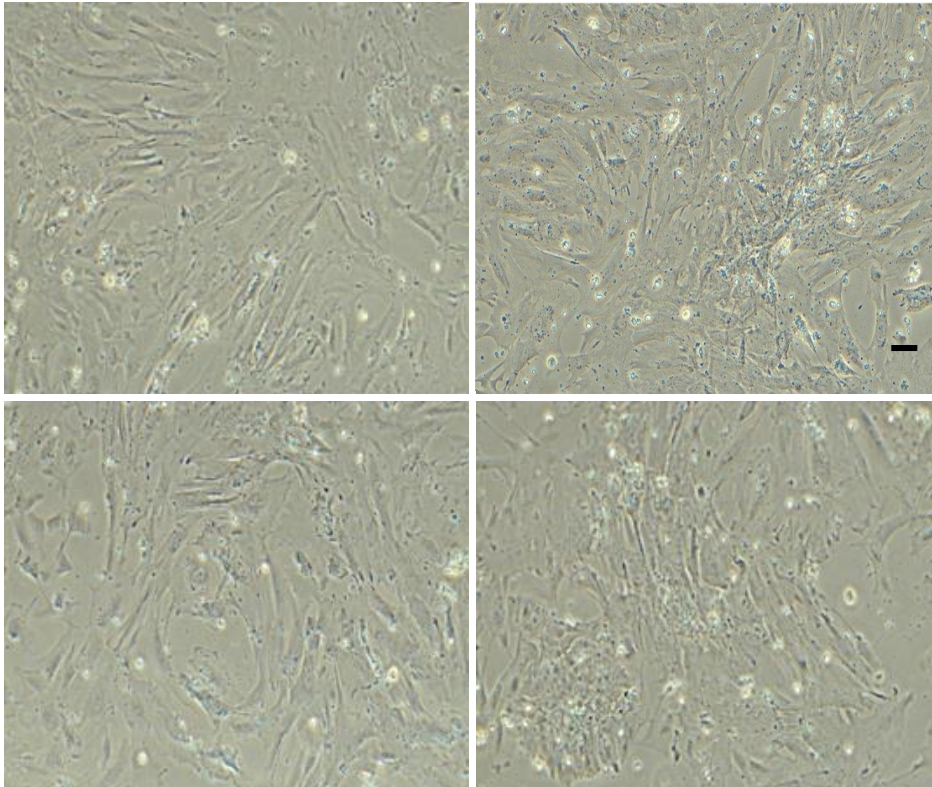
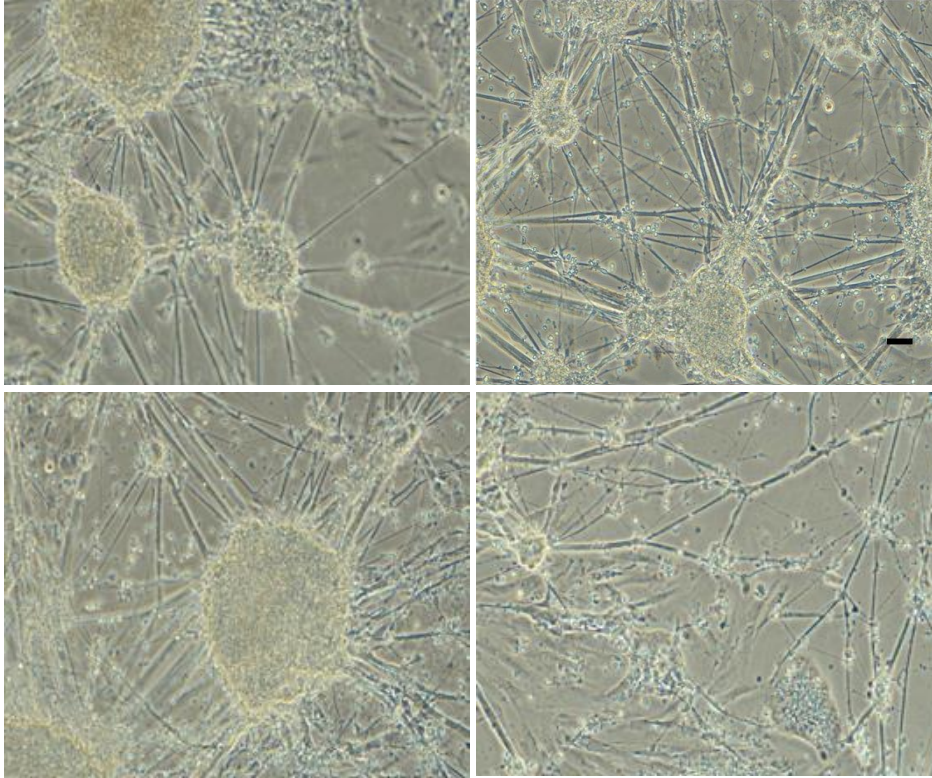


Figure 6.3. Myogenic differentiation of DUP3 iPSCs cells seeded at low density.

DUP3 iPSCs cells were induced towards myogenic specification according to the protocol detailed by Chal *et al*⁶⁹⁰. Cells seeded at low density (LD) were either treated with 3 μ M (top panel) or 6 μ M (bottom panel) and imaged at completion of the differentiation protocol (day 25). Pictures were acquired by a white light microscope at a magnification of 40X. Scale bar = 200 μ m.

MD 3 μ M CHIR



MD 6 μ M CHIR

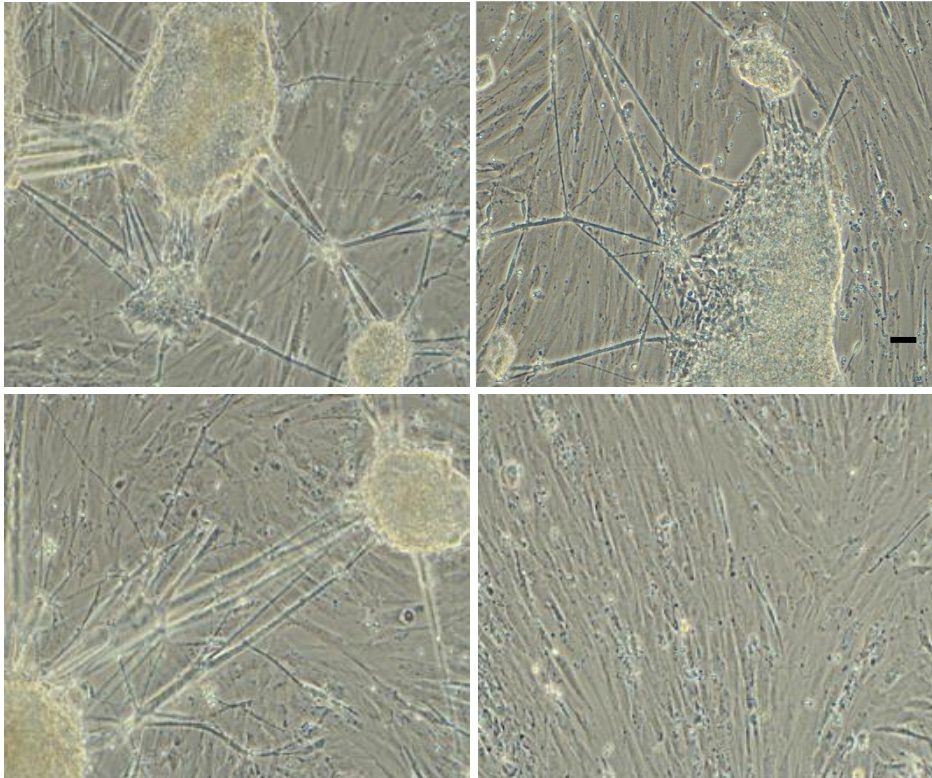
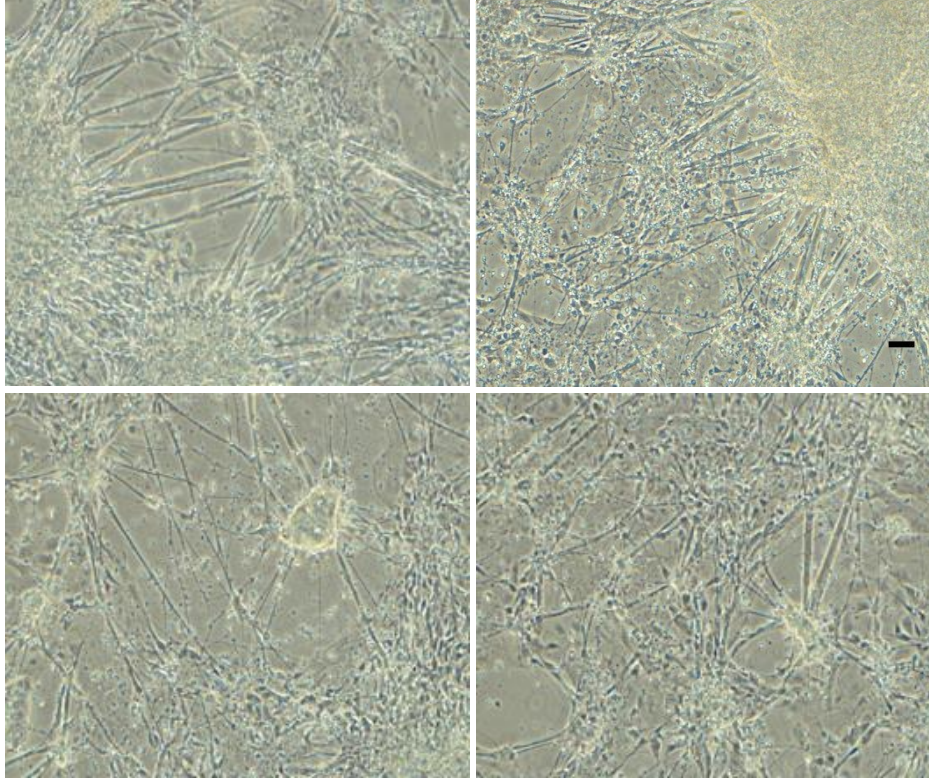


Figure 6.4. Myogenic differentiation of DUP3 iPSCs cells seeded at medium density.

DUP3 iPSCs cells were induced towards myogenic specification according to the protocol detailed by Chal *et al*⁶⁹⁰. Cells seeded at medium density (MD) were either treated with 3 μ M (top panel) or 6 μ M (bottom panel) and imaged at completion of the differentiation protocol (day 25). Images were acquired by a white light microscope at a magnification of 40X. Scale bar = 200 μ m.

HD 3uM CHIR



HD 6uM CHIR

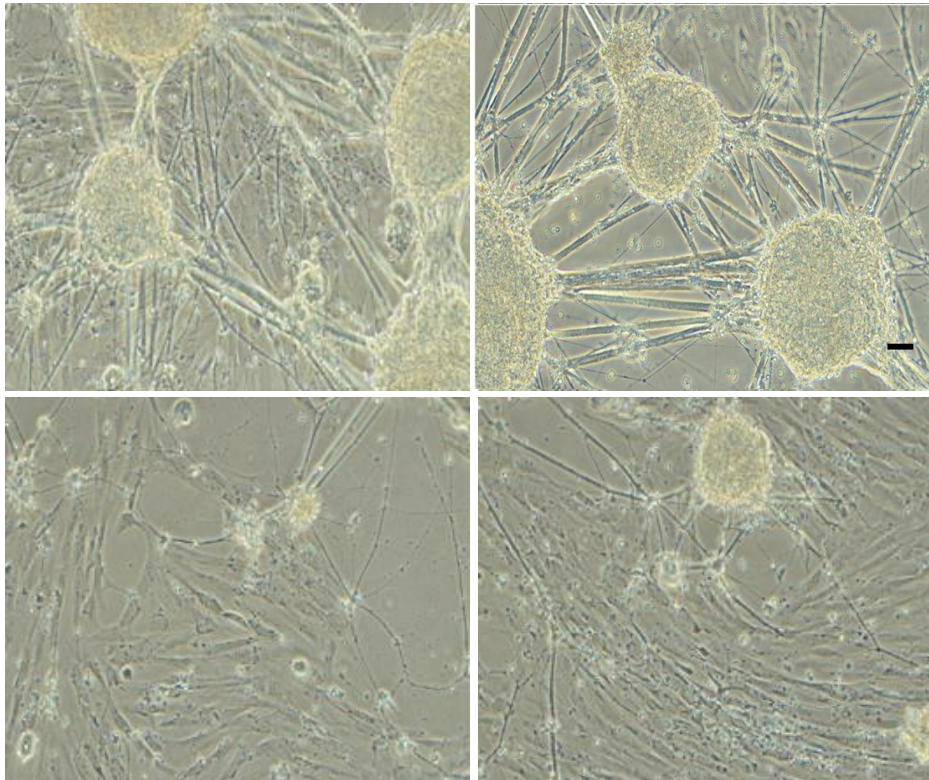


Figure 6.5. Myogenic differentiation of DUP3 iPSCs cells seeded at high density.

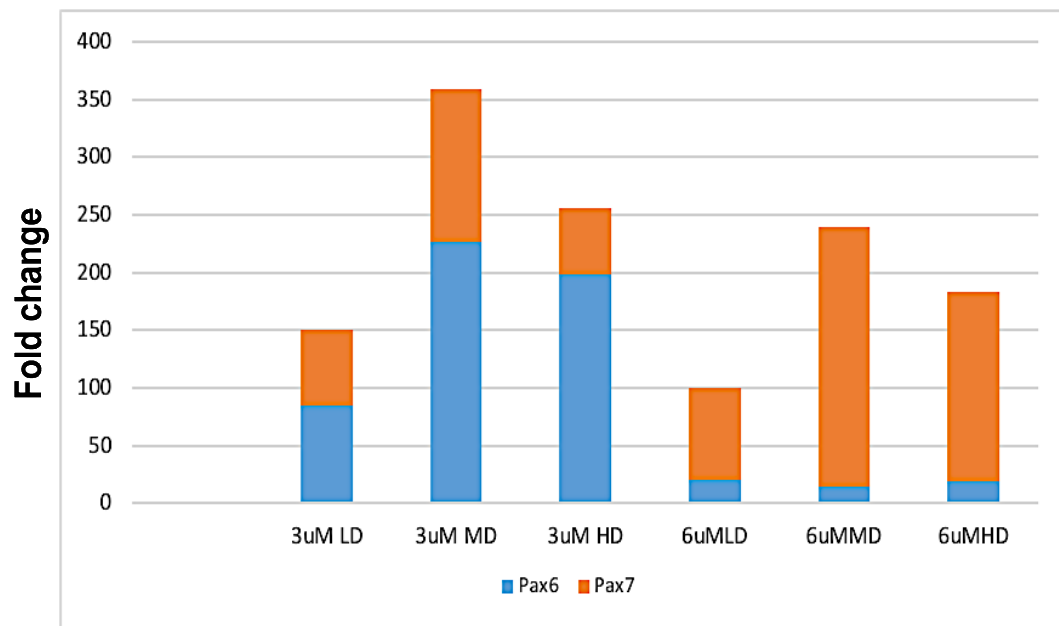
DUP3 iPSCs cells were induced towards myogenic specification according to the protocol detailed by Chal *et al*⁶⁹⁰. Cells seeded at high density (HD) were either treated with 3 μ M (top panel) or 6 μ M (bottom panel) and imaged at completion of the differentiation protocol (day 25). Pictures were acquired by a white light microscope at a magnification of 40X. Scale bar = 200 μ m.

6.3.3. Analysis of myogenic markers expression following iPSCs myogenic induction

Each replicate of the differentiated DUP3 of section 6.3.2 were passaged 1:2. I extracted RNA (section 2.3.10) from each of the four resulting samples and used it to quantify Pax7 expression via qPCR (section 2.3.13.a), as shown in fig (**Fig.2.14**). I did this to accurately define which cell density and CHIR99021 concentration resulted in the highest percentage of myogenic precursors expressing the myogenic marker Pax7. However, I considered that Pax7 is also expressed in precursors of the neuronal lineage^{593,594}. Therefore, I used the same qPCR protocol to also quantify the expression of the neuronal marker Pax6. This allowed me to estimate the degree of neuronal differentiation in each of the analyzed samples. I normalized both Pax7 and Pax6 expression data against the expression of the constitutive SDHA gene, coding for the succinate dehydrogenase complex flavoprotein subunit A⁵⁹⁵. The qPCR primers I used are detailed in section 2.3.13.a.

As negative controls, I included the RNA extracted from wild-type iPSCs (provided by Dr. Jinhong Meng) and undifferentiated DUP3 iPSCs (**Fig.2.14**). As they were not induced towards myogenic or neuronal differentiation, these iPSCs should not express Pax7 or Pax6. I loaded each of the four samples three times in the qPCR plate (technical replicates).

I analysed qPCR data using the $\Delta\Delta C_t$ method (section 2.3.14). I observed that, as a general trend, cells cultured with 6 μ M CHIR99021 had an high Pax7 expression and a low Pax6 expression. Within this group, cells having the highest Pax7 and lowest Pax6 expression were cells seeded at a medium density and cultured with 6 μ M CHIR99021. Specifically, Pax7 and Pax6 expression was 224.94 and 14.8 times higher than undifferentiated DUP3 iPSCs, respectively (**Fig.6.6**). I then checked by qPCR the expression of myogenin and myosin heavy chain (MF20) in cells cultured with 6 μ M CHIR99021. I observed that the expression of MF20 was generally low, indicating a small number of terminally differentiated cells. Myogenin, instead, was highly expressed in cells seeded at medium an high density. The highest myogenin expression (152.40 times higher than in undifferentiated DUP3 iPSCs) was detected in cells seeded at medium density (**Fig.6.7**). I therefore expanded these cells and used them for NEON electroporation (section 2.2.12.b) with CRISPR0 and CRISPR2 σ plasmids.



Differentiation conditions

Figure 6.6. Pax7 and Pax6 expression upon myogenic differentiation of DMD-derived iPSCs.

The expression of Pax7 and Pax6 upon myogenic induction in iPSCs was normalized to the expression of the constitutive SDHA gene. Fold changes of expression were obtained by comparison with the original DUP3 iPSCs, prior to their differentiation.

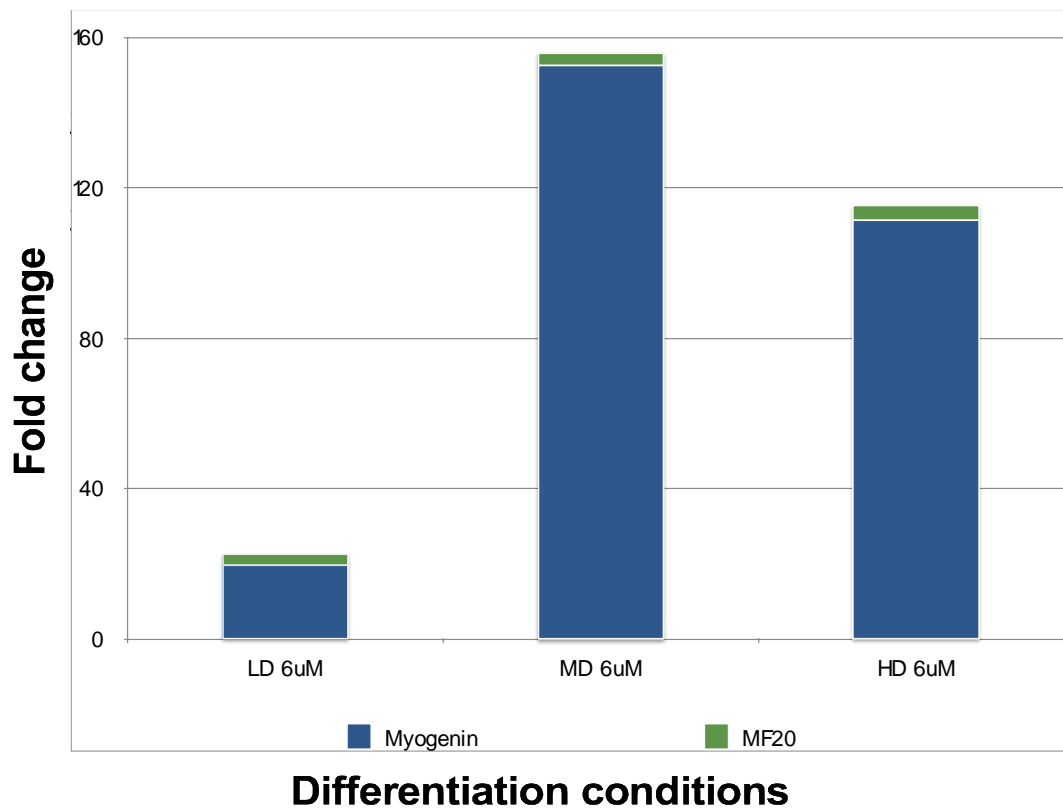


Figure 6.7. Myogenin and myosin expression upon myogenic differentiation of DMD-derived iPSCs cultured with 6 μ M CHIR99021.

The expression of myogenin and MF20 upon myogenic induction in iPSCs cultured with 6 μ M CHIR99021 was normalized to the expression of the constitutive SDHA gene. Fold changes of expression were obtained by comparison with the original DUP3 iPSCs, prior to their differentiation.

6.3.4. NEON electroporation of myoblast-like cells derived from iPSCs

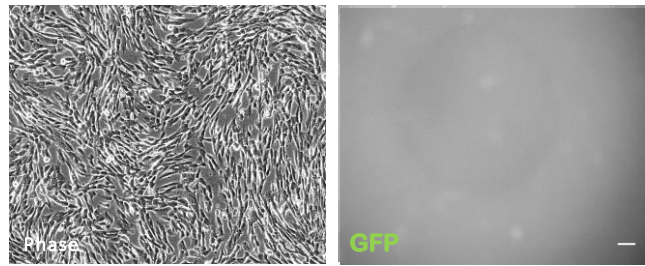
I electroporated DUP3 iPSCs-derived myoblast-like cells with CRISPR0 and CRISPR2 σ plasmids using the electroporation condition 1, as this worked well for the electroporation of DUP2i myoblasts (section 2.2.12.b). The experiment was performed three times (biological replicates). Negative controls were represented by cells electroporated with CRISPR0 and non-treated cells (**table 6.2**).

Control (+/-)	Electroporated plasmid	Function
-	CRISPR0	Monitor non-specific cleavage
-	None	Monitor background bands

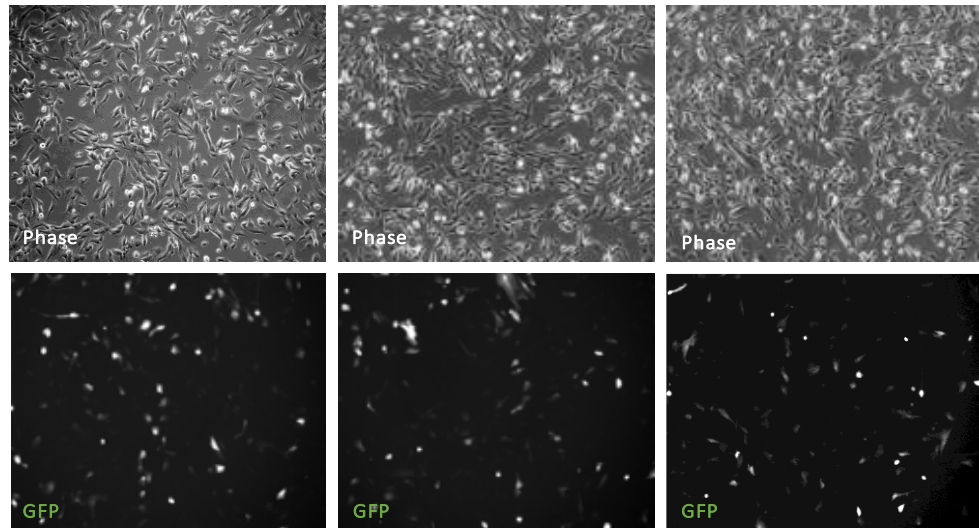
Table 6.1. List of controls used for the electroporation of DUP3-derived myoblast-like cells with CRISPR/Cas9 plasmids.

After 48 hours, I observed more than 50% fluorescent cells expressing EGFP by microscopic analysis (**Fig.6.8**) (section 2.3.23) and I isolated them via FACS (section 2.2.14.a). According to FACS data, the percentage of cells expressing EGFP following CRISPR2 σ electroporation was the highest (33.23 +/- 0.37) that I have achieved in any cell type I electroporated (DUP1, DUP2, DUP2i, **table 2.1** in section 2.2.2). Moreover, I observed that a considerably higher number of cells expressed CRISPR2 σ (33.23 +/- 0.37) as opposed to CRISPR0 (13.9 +/- 0.56) (**Fig.6.9**). Such difference was also found upon microscopic analysis (57.71 +/- 2.61 cells expressing CRISPR2 σ versus 40.91 +/- 1.51% cells expressing CRISPR0). I noticed that via microscopic analysis I detected an higher percentage of EGFP-positive cells than via FACS. I attributed this outcome to the fact that part of the cells imaged at the microscope were fluorescent but rounded dead cells that lied above healthy cells. The inclusion of these in the manual counting might have biased the quantification by overestimating the actual number of live fluorescent cells in the plate.

UT



CRISPR0



CRISPR2σ

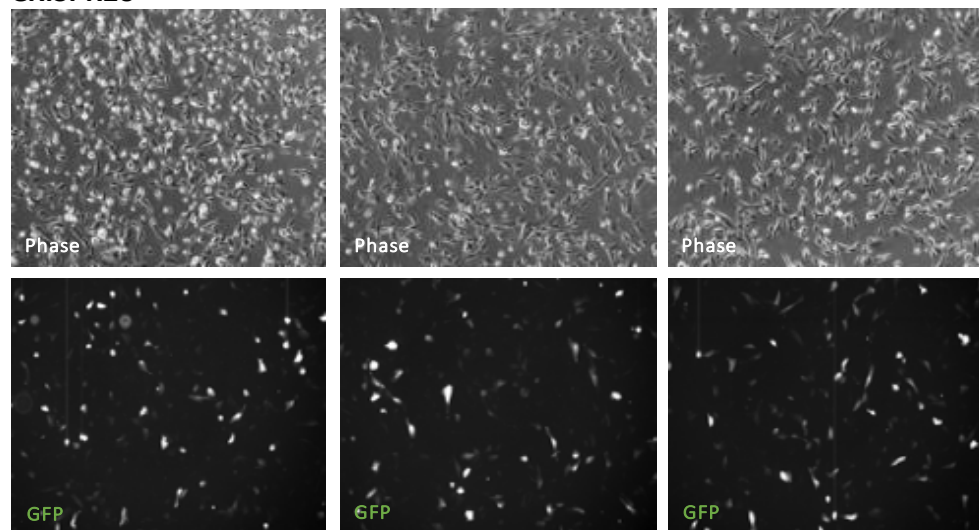


Figure 6.8. NEON electroporation of DUP3 iPSCs-derived myoblasts.

Myoblasts derived from the DUP3 iPSCs cell line carrying a dystrophin duplication in dystrophin exons 3-41 was electroporated with the CRISPR0 and CRISPR2σ with the parameters previously chosen for condition 1 (1 pulse, 1400 V, 20 milliseconds). Almost 60% of cells expressed CRISPR2σ (57.71 +/- 2.61) while fewer cells (40.91 +/- 1.51%) expressed CRISPR0. Images were taken 48 hours post-electroporation by a camera associated to the inverted fluorescent microscope Olympus IX at a magnification of 5X.

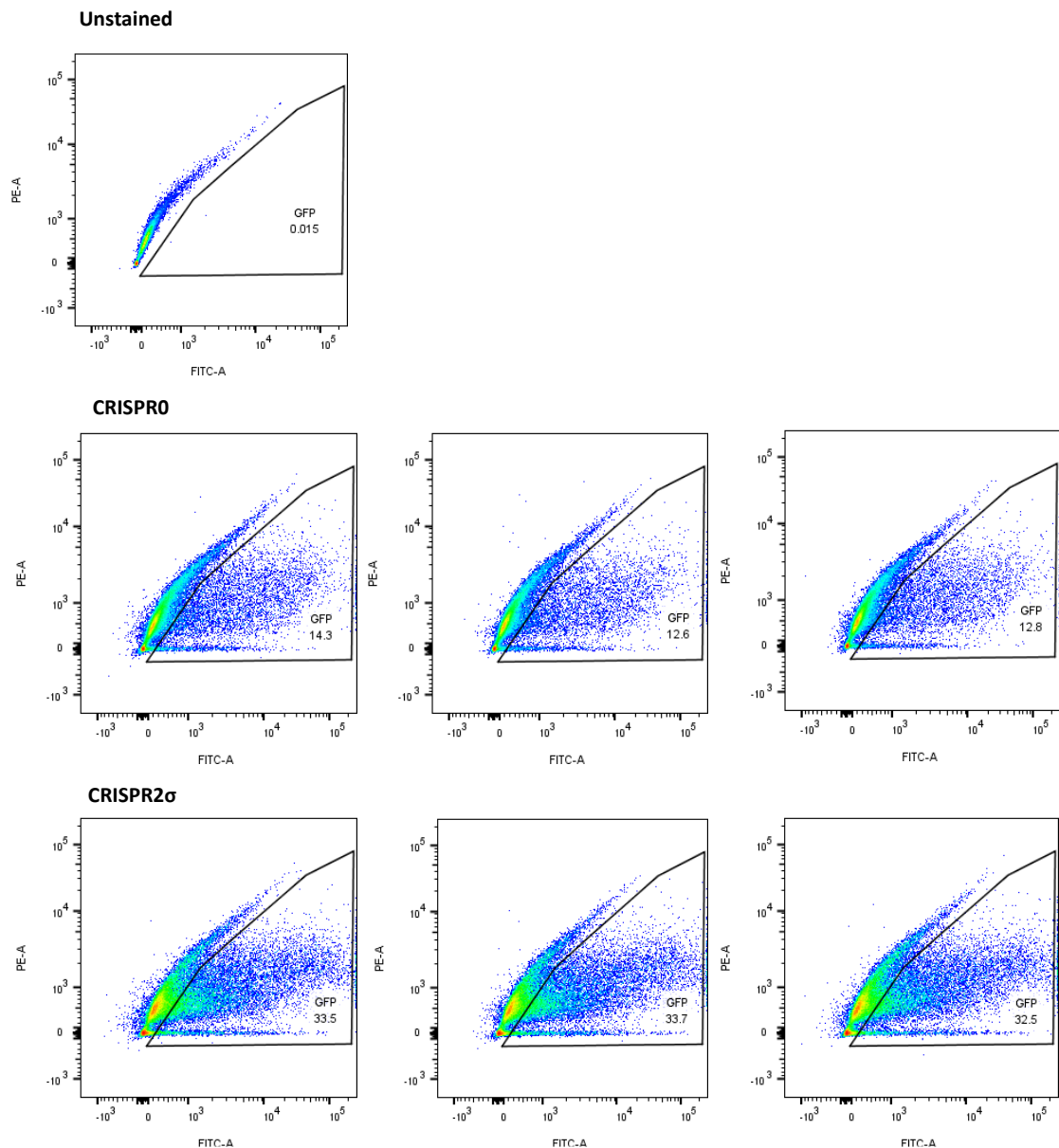


Figure 6.9. FACS analysis of electroporated iPSCs-derived patient myoblasts.

Myoblasts derived from the iPSCs cell line carrying a dystrophin duplication in dystrophin exons 3-41 and electroporated with the CRISPR0 and CRISPR2 σ plasmids were sorted by the FACSaria cell sorter. Consistently with the fluorescent microscope analysis, a greater percentage of cells expressed CRISPR2 σ (33.23 \pm 0.37) while less cells expressed CRISPR0 (13.9 \pm 0.56).

6.3.5. T7 assay on electroporated and FACS sorted iPSC-derived myoblast-like cells

I expanded FACS sorted cells, extracted their DNA (section 2.3.1) and amplified it as required for the T7 assay (sections 2.3.4 and 2.3.7), which I performed to monitor the CRISPR/Cas9 targeting efficiency in DUP3- myoblast-like cells. When I ran T7 assay products on a 4-20% TBE gel (section 2.3.8), I observed cleaved bands of the expected molecular size (511 bp and 217 bp) in samples obtained from cells electroporated with CRISPR2 σ (**Fig.6.10**). The cleavage efficiencies measured in cells electroporated with CRISPR2 σ were 19.7%, 16.7% and 17.3% (mean efficiency 17.9 \pm 0.92). However, I observed non-specific background bands (whose molecular weight was similar to the lower CRISPR2 σ cleaved fragment) also in samples derived from CRISPR0 electroporation. As such, I could not infer that these values of CRISPR/Cas9 were reliable.

6.3.6. Inducing terminal differentiation of iPSCs-derived myoblast-like cells

To assess dystrophin repair at transcript and protein level, I had to culture the electroporated and sorted cells in M2 medium (section 2.2.6) to generate myotubes expressing dystrophin. However, as the myogenic differentiation of iPSCs did not result in a pure population, I performed FACS analysis to screen for the ErbB3 marker, expressed in myogenic cells (section 2.2.14.b). Unexpectedly, a very limited number of cells expressed the ErbB3 marker (**Fig.6.11.**), indicating that this cell pool contained only a few myoblast-like cells that could have sustained myogenic differentiation. Therefore, I could not obtain RNA and proteins.

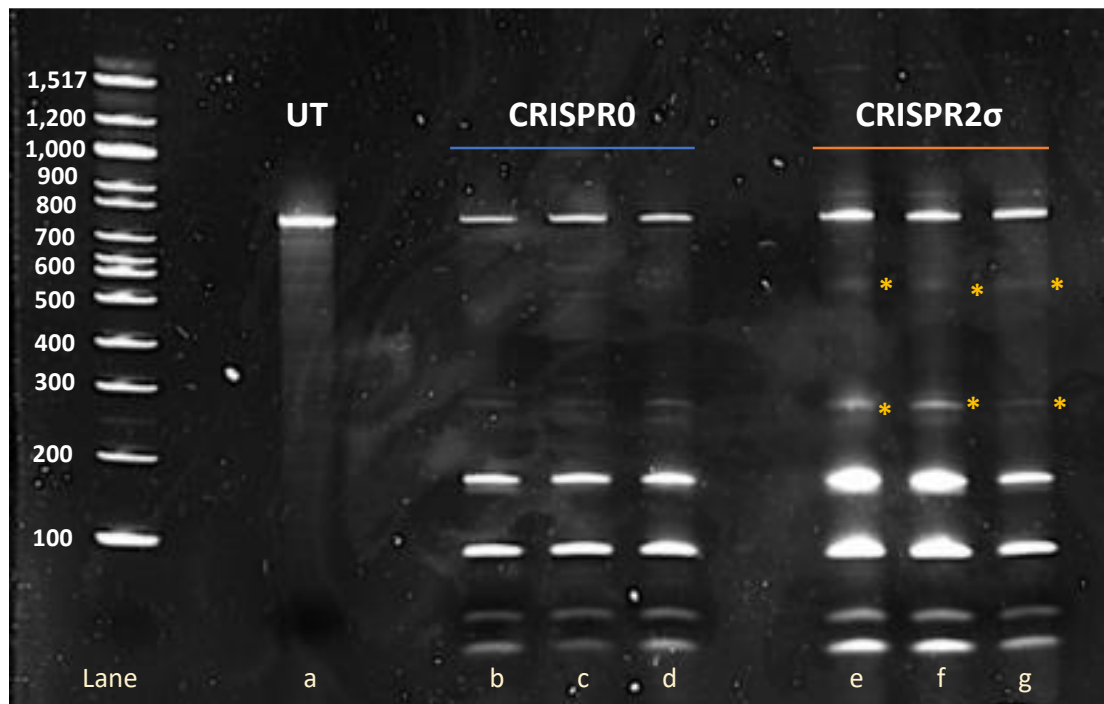


Figure 6.10. T7 assay on electroporated myoblast-like cells derived from DUP3 iPSCs.

T7 assay was performed on DNA derived from myoblast-like cells derived from DUP3 iPSCs and electroporated with CRISPR/Cas9 plasmids (NEON device). Cleaved bands of the expected molecular weight (511 bp and 217 bp) were observed in cells electroporated with CRISPR2σ (lanes e-g). However, none of the cleaved bands was detected in the untreated sample (lane a) and in samples electroporated with the CRISPR0 plasmid (lanes b-d). The molecular weight of each full length and cleaved band was assessed by using the 100 bp ladder (NEB).

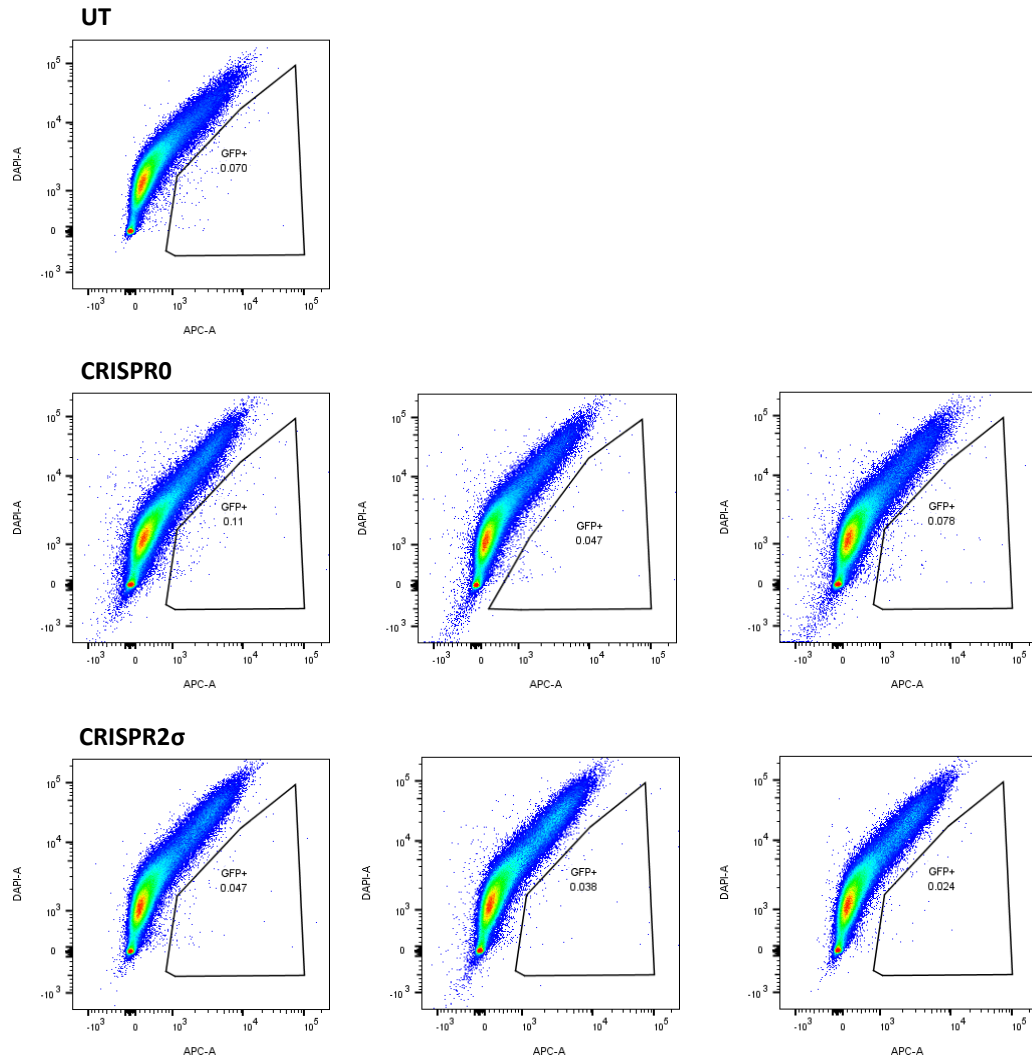


Figure 6.11. Selection of iPSCs-derived myoblast-like cells expressing *ErbB3*.

Once the DUP3-derived myoblast-like cells expressing CRISPR/Cas9 plasmids were isolated, they were further selected according to the expression of the myogenic marker *ErbB3*. A very low percentage of cells (below 1%) expressed this marker and was isolated by flow-cytometry. UT= untreated cells; CRISPR0= cells previously electroporated with CRISPR0; CRISPR2σ= cells previously electroporated with CRISPR2σ.

6.4 DISCUSSION

In this chapter I aimed to investigate the ability of a single CRISPR/Cas9 nuclease approach to restore considerably large dystrophin duplications in a DMD iPSC line.

The use of CRISPR/Cas9 to remove large-sized genomic regions was reported a few years ago by Young *et al.*, who generated an internally deleted dystrophin devoided of up to 725 kb in DMD iPSCs⁵⁹⁶. However, his approach was not applied to *DMD* duplications and was based on the use of two different nucleases.

Conversely, Long *et al.*⁵⁹⁷ restored dystrophin in iPSC-derived cardiomyocytes carrying a duplication of *DMD* exons 55-59 using the single nuclease approach I used in this thesis. Nevertheless, his approach was applied to a multi-exon duplication whose size was limited compared to the one present in the cell line I chose, spanning exons 3-41 (about 1000 kb). Achieving dystrophin repair in this cell line would therefore allow me to demonstrate for the first time that CRISPR/Cas9 can target and correct genomic duplication larger than 1 million bp.

I chose to electroporate CRISPR/Cas9 nuclease both in iPSCs and in myoblast-like cells derived from iPSCs (duplication of *DMD* exons 3-41) to compare the efficiency of editing in these two cell types.

I considered that the study of both cell types would be advantageous for my study. On one hand, the benefit of using gene-edited iPSC clones is their capability of unlimited expansion and the possibility to use them as isogenic control alongside the non-treated iPSC cells⁵⁹⁸. Furthermore, iPSCs can generate not only skeletal muscle, but also cardiomyocytes and neuronal cells, all expressing dystrophin and functionally compromised in Duchenne muscular dystrophy (section 1.2). Therefore, the removal of *DMD* duplication at the iPSC stage would allow me to obtain cells expressing wild-type dystrophin that I could differentiate towards the lineages mentioned above. In this way, I could assess the efficiency of dystrophin protein repair across multiple cell types.

On the other hand, the CRISPR/Cas9 delivery to iPSC-derived myoblast-like cells would allow me to determine how the editing efficiency varies across myoblasts carrying specific mutations. Moreover, by comparing the results obtained from the nuclease expression in iPSCs and iPSCs-derived myoblast-like cells, I could examine the efficiency of CRISPR/Cas9 in two separate cell types genotypically identical. Finally, the use of iPSC-derived myoblast-like cells would be beneficial also from a technical perspective. In fact, once the myogenic cells have been produced in culture, they can be easily expanded and stored for further applications. This would allow me to save time and lower the costs associated with the iPSCs differentiation following each experiment.

I observed that the CRISPR/Cas9 delivery in these two cell types resulted in clearly different outcomes. The electroporation of cells at the iPSC stage gave rise to a considerable cell death, even though two of the three parameters I tested were shown to be successful in iPSCs in the publications of Wang *et al.*, Howden *et al.* and Ihry *et al.*^{599–601}. According to the recent study by Ihry *et al.*, high levels of CRISPR/Cas9 expression in iPSCs are associated with a considerable cell death, by virtue of the Cas9-mediated activation of the p53 pathway⁶⁰¹. However, if the cell death I observed had been derived by an high level of CRISPR/Cas9 expression, I would have expected it to occur later in time and not immediately after electroporation. I therefore hypothesized that the cell death might be due to the unsuitable electroporation conditions for these particular patient-derived iPSCs.

In contrast, iPSC-derived myoblast-like cells survived the electroporation and expressed high levels of CRISPR/Cas9. I noticed that, despite having been manipulated first by electroporation and then by sorting, the iPSC-derived cells maintained their proliferative ability, so that I could quickly expand them to assess CRISPR/Cas9 efficiency at the genomic level. It is likely that such considerable proliferation was due to the protocol I used for their differentiation, published by Chal *et al.*⁵⁹⁰. This protocol, in fact, gives rise not only to mature muscle fibers, but also to satellite cell-like myogenic precursors. This provides an advantage over primary myoblasts when cells have to be kept in culture for a long period of time, as the proliferative ability of primary myoblasts is limited compared to satellite cells⁶⁰².

In this thesis, iPSC-derived myogenic precursors treated with CRISPR/Cas9 would be ideally maintained in the proliferative state until their number was sufficient for DNA, RNA and protein studies. Only at that point cells would be induced towards terminal differentiation. The same approach would be also suitable if *ex vivo* cell therapies were considered, as it would allow the generation of a considerable number of cells to be transplanted *in vivo* (even though cells with a limited proliferative ability *in vitro* have been shown to have a better survival post engraftment⁶⁰³).

I observed that the efficiency of the CRISPR-mediated genomic cleavage in iPSC-derived myoblast-like cells was similar to that which occurred in the immortalized DUP2i myoblasts (duplication of *DMD* exons 3-16, **table 2.1**, section 2.2.2) expressing low levels of CRISPR/Cas9 (around 20%). In order to assess the efficiency of dystrophin correction I had to integrate the genomic analysis with dystrophin transcript and protein data. However, neither of these are available at present. In fact, I could not generate myotubes from which to extract RNA and protein, as a very limited amount of the myogenic precursor cells expressed the myogenic surface marker ErbB3 according to FACS analysis. This outcome lead me to hypothesize that contaminant cells were

present in culture, potentially generated by experimental variables occurred during the iPSC differentiation. In support of this hypothesis, I noticed that almost all the differentiation conditions I tested (apart for cells seeded at low density and treated with 6 μ M CHIR99021) visually appeared to be a mixture of myogenic and neuronal cells. This was also observed by Hosoyama *et al.*⁶⁰⁴ in the presence of the differentiation conditions I used.

Since the outcome of FACS analysis limited the information that I could derive in regards to the efficiency of CRISPR/Cas9 for correcting large dystrophin duplications, for future work I would recommend the selection of a pure myoblast population based on myoblast specific markers^{605,606} before CRISPR/Cas9 delivery. In this way, contaminant cells will be eliminated and I could obtain a population of myogenic cells able to sustain terminal differentiation. This approach nevertheless is not devoid of risks. Cells would have to be kept in culture for the relatively long period associated with the CRISPR/Cas9 electroporation, followed by sorting and expansion that would be needed before inducing terminal differentiation. This might lead to cells senescence or death⁶⁰⁷, compromising the success of the experiment.

To summarize, in this chapter I observed that:

- DUP3 (duplication of *DMD* exons 3-41) iPSC-derived myoblast-like cells (but not DUP3 iPSCs) can express CRISPR/Cas9 following NEON electroporation.
- CRISPR/Cas9 can target dystrophin genomic region in myoblast-like cells derived from DUP3 iPSCs.
- The isolation of a pure myoblasts cell population is a necessary requirement for the analysis of dystrophin transcript and protein, so that the CRISPR/Cas9 ability to repair large dystrophin duplications can be confirmed.

Chapter 7. General discussion

7.1 THE ROLE OF CRISPR/CAS9 AS A THERAPEUTIC STRATEGY FOR DMD

Duchenne Muscular Dystrophy is caused by genetic mutations resulting in the absence of the intracellular protein dystrophin, and is one of the most debilitating and severe neuromuscular conditions that leads to death early in life¹³. In the years which followed the discovery of the *DMD* gene in 1985, clinicians and scientists undertook extensive investigations to understand the DMD pathogenesis (section 1.2) and, since then, many therapeutic approaches have been developed in the attempt to delay DMD progression (section 1.6). However, none of the proposed strategies operates at the genomic level to restore a wild-type *DMD* gene, as they act either by mitigating the symptoms of the disease, or by compensating for the consequences of the missing/defective protein.

As an example, the mutation-independent approach based on AAV-mediated delivery of mini/micro-dystrophin constructs can be used for any type of DMD mutation and potentially be exploited to ameliorate disease severity (section 1.6.2.d). However, as the dystrophin transgene expressed by the AAVs is internally deleted, it is expected to be only partially functional. This shorter dystrophin isoform would not be able to totally compensate for the absence of dystrophin. Therefore, it is expected that, in the best scenario, its expression would result in a milder Becker-like phenotype (section 1.2). Similarly, exon-skipping (section 1.6.2.b) would be able to restore the dystrophin reading frame by skipping selected exons, but would not lead to the production of a wild-type dystrophin. Furthermore, exon-skipping is only applicable to deletions, small duplications and nonsense mutations disrupting dystrophin reading frame, but is not applicable to large duplications spanning multiple exons.

Finally, most of these approaches would require the repeated administration of the therapeutic agent for a life-long maintenance of the beneficial effect, with the exception of the AAV, for which repeated administration is currently not feasible due to immunological reasons.

An alternative to these therapies has recently been found in engineered nucleases and in particular in CRISPR/Cas9, a bacterial-derived nuclease that can potentially target and cleave any genomic region that needs to be edited³⁷⁵. The main advantage of this strategy is that any modification introduced in the target DNA would be permanent. Furthermore, depending on which repair pathway prevails after that the nuclease cuts the target DNA (i.e. HDR or NHEJ), CRISPR/Cas9 can be exploited to repair any type

of mutation, from smaller point mutations to larger deletions and duplications (section 1.8).

Long *et al.*⁶⁰⁸, for example, were among the first to demonstrate CRISPR/Cas9-mediated dystrophin restoration via HDR in the *mdx* murine model (which carries a point mutation in exon 23)²¹⁷. By injecting *mdx* zygotes with both CRISPR/Cas9 components (sgRNA and Cas9, section 1.8) and a DNA template carrying the wild-type exon 23, he demonstrated variable degrees of dystrophin correction (from 2% to 100%) in three different skeletal muscles (quadriceps, soleus and diaphragm) and cardiac muscle of mice examined in their adulthood.

The approach proposed one year later by Ousterout *et al.*, instead, relied on the NHEJ repair mechanism to partially correct dystrophin both *in vitro* in immortalized DMD myoblasts and *in vivo* in immunodeficient mice⁶⁰⁹. His strategy (defined as multiplexed approach) exploited the simultaneous delivery of multiple sgRNAs either to introduce large dystrophin deletions or insertion/deletions (indels, section 1.7) that would restore the dystrophin reading frame⁶⁰⁹.

However, the use of multiple sgRNAs increases the probability of observing cleavage at unintended sites and chromosomal translocations between on-target and off-target sites⁶¹⁰.

In the present study, I developed a simpler NHEJ-based CRISPR/Cas9 strategy that would not require multiplexing and tested its efficacy *in vitro*, as a proof-of-concept for the generation of future *in vivo* therapeutic approaches.

Specifically, the strategy I designed was based on the use of a single CRISPR/Cas9 nuclease to correct dystrophin duplications (section 1.9), that occur in 10-15% of DMD patients¹. The hypothesis behind the project was that directing the nuclease to a specific genomic region within the duplication would result in two genomic cuts (one for each target sequence). This would cause the elimination of the intervening duplicated segment and, upon ligation of the two cleaved ends by NHEJ, the restoration of the *DMD* gene (**Fig.1.9**). This approach is theoretically applicable to all the patients carrying *DMD* duplications, as it can be used for correcting both small and large duplications.

The main advantage of my approach is that DMD patients treated with CRISPR/Ca9 would not produce a Becker-like dystrophin, but a fully functional dystrophin that would be expressed in a proportion of nuclei, resembling the situation of DMD carriers⁶¹¹. Depending on the extent of somatic cell correction, this could potentially be advantageous as most DMD carrier typically have no substantial progression of clinical symptoms.

7.2 CRISPR/CAS9 AND DYSTROPHIN EDITING: THE IMPORTANCE OF THE STUDY DESIGN

As a first step, I had to select a genomic region of the *DMD* gene to be targeted by the nuclease. I decided to target an intronic region. The reason for this choice was linked to the fact that NHEJ, the prevalent repair mechanism of DNA breaks in the absence of a DNA template, is an imprecise mechanism which could randomly introduce insertion and/or deletions at the break site. If these would occur within an exon, they could potentially disrupt the *DMD* gene reading frame⁶¹². Alternatively, indels could result in the addition of new amino-acid sequences to the dystrophin protein. This would generate new epitopes that could alter the protein structure⁶¹². Both cases would result in a non-functional or altered protein.

I then had to consider which cell type was the best to demonstrate my experimental hypothesis *in vitro*. I demonstrated that CRISPR/Cas9 nuclease can target both small and larger genomic duplications in primary fibroblasts, primary and immortalized myoblasts and iPSCs. I initially expressed CRISPR/Cas9 in primary fibroblasts derived from patients carrying dystrophin duplications as these were readily available to me through our BioBank. Since dystrophin protein is not detectable in fibroblasts⁴⁶⁸, I had to induce them towards the myogenic lineage through the forced expression of the MyoD transcription factor (section 4.1.2). However, I could not generate enough myoblast-like cells to be terminally differentiated into myotubes expressing dystrophin⁴⁹⁰ (section 4.3.5), and therefore I only used fibroblasts to confirm genomic targeting upon CRISPR/Cas9 treatment. Instead, I identified a patient-derived myoblast cell line carrying a duplication of *DMD* exons 3-16 with an extremely high myogenicity (fusion index of 90%) which allowed me to generate terminally differentiated myotubes so that I could assess dystrophin repair at protein level. Thanks to this highly myogenic myoblast cell line, I was able to show that CRISPR/Cas9 is capable of correcting small dystrophin duplications (sections 5.3.2.c, 5.3.3.bii and 5.3.3.ciii).

I also used iPSCs (duplication of *DMD* exons 3-41), a cell model that offers considerable advantages for diseases affecting multiple tissues as for DMD⁶¹³. Despite the fact that I did not succeed in expressing CRISPR/Cas9 in iPSCs, I would consider this cell type for future studies. In fact, iPSCs can be used to derive not only myoblasts, but also cardiomyocytes and neurons (which are also compromised in DMD but impossible to obtain other than at post-mortem). This feature makes iPSCs appealing not only for demonstrating genetic correction of different cell types *in vitro*, but also for *in vitro* disease modelling aimed to provide new molecular insights into the pathogenic mechanisms of the disease⁶¹³.

Moreover, if *ex vivo* cell therapies based on transplantation of edited patient-derived cells were considered, the indefinite proliferative ability of iPSCs would generate a larger number of cells to transplant than primary myoblasts or mesoangioblasts.

Alternatively, despite not used for my study, I would consider the use of urine-derived stem cells, recently identified by Falzarano *et al.*⁶¹⁴. Urine-derived stem cells can be coaxed down the myogenic lineage and generate myogenic precursor cells⁶¹⁴ but, unlike iPSCs, they maintain their proliferative ability without generating tumours or teratomas⁶¹⁵. Furthermore, unlike fibroblasts, myoblasts and iPSCs, urine-derived stem cells are easy to be accessed without any discomfort.

The use of a stable versus a transient transgene expression was a further parameter that I had to consider in my experimental design. Initially, I chose integrating lentiviral vectors for demonstrating the efficacy of my approach *in vitro*. The reason for such choice is that CRISPR/Cas9 transgenes expressed by lentiviral vectors are integrated into the DNA and persistently expressed. This results in an overload of CRISPR/Cas9 nuclease, thus increasing the chances of cleaving DNA at both duplicated sites and, as a consequence, of restoring as much dystrophin as possible. Moreover, lentiviral vectors can accommodate transgenes as large as 11 kb⁴⁰⁶. As such, these vectors are suitable for hosting the all-in-one plasmids that express not only the CRISPR/Cas9 components, but also selection markers (antibiotic resistance and/or fluorescent proteins) to facilitate the readout of CRISPR/Cas9 expression in the treated cells in an *in vitro* system. Thanks to the permanent co-expression of Cas9 and EGFP by means of lentiviral vectors, I could select transduced cells and demonstrate that the approach I designed is successful to correct dystrophin duplications *in vitro*.

The single CRISPR/Cas9 nuclease approach trialled in my study has been independently validated *in vitro* by the work of Lattanzi *et al.*⁶¹⁶, who succeeded in removing the duplication of dystrophin exon 2 (the most common dystrophin duplication)⁷⁸. Similarly, Wojtal *et al.*³⁸⁹ demonstrated the removal of a dystrophin duplication by using a single nuclease not only *in vitro* in patient-derived myoblasts, but also *in vivo* in a mouse model carrying a dystrophin duplication of exons 18-30. In both studies CRISPR/Cas9 was expressed by the LentiCRISPRv2 vector and transduced cells were selected exploiting puromycin selection.

Kimura *et al.*⁵³³ showed that lentiviral vectors delivered *in vivo* in the tibialis anterior of *mdx*^{4cv} dystrophic mice (nonsense mutation in dystrophin exon 53)⁶¹⁷ were able to target and integrate in satellite cells. This represents a major advantage of these vectors, as satellite cells are the muscle precursor stem cells that are activated upon muscle damage to produce new myoblasts. If CRISPR/Cas9 would repair DMD gene

in satellite cell's DNA, each newly formed myoblast would express a wild-type and stable dystrophin which would contribute to the integrity of the myofibers with which it would fuse. Similarly, the new myofibers generated by the regenerative process would not undergo degeneration⁶¹⁸. Moreover, the muscle regeneration process itself would be improved by correcting dystrophin in satellite cells, as regeneration process is affected by dystrophin loss in this cell type¹³⁹.

Despite this important achievement, an intra-muscular delivery route on the line of Kimura *et al.*⁵³³ could not be used for pathologies like DMD which affects muscles body-wide. A systemic delivery would be necessary, but lentiviral particles delivered systemically are known to be sequestered by the liver and rapidly cleared from the circulation following the interferon response, thus hampering successful systemic muscle targeting⁶¹⁹.

Moreover, although lentivirus can integrate in both dividing and non-dividing tissues such as muscle²⁶⁴, their ability to integrate within the host genome would increase the chances of off-targeting due to the constant presence of Cas9, and also the risk of tumorigenesis due to insertional mutagenesis²⁷². Taken together, these limitations led me to consider alternative non-integrating delivery systems which I could eventually use for the delivery of the CRISPR/Cas9 *in vivo* model, which is the future goal of the project.

I therefore investigated if an episomal (and thus transient) CRISPR/Cas9 expression could be exploited to repair dystrophin duplications *in vitro*. I hypothesized that a provisional CRISPR/Cas9 expression might, in fact, not be sufficient for achieving a sufficient level of duplication removal to be detectable as a restored dystrophin protein. I attempted transient CRISPR/Cas9 expression *in vitro* in patient-derived myoblasts (duplication of *DMD* exons 3-16) either by transfecting these cells by commercial transfection reagents or by electroporation. The use of transfection reagents resulted in a very low level of CRISPR/Cas9 expression (below 5 %, section 5.3.3.a) and was therefore not pursued further. Electroporation, instead, allowed the nuclease expression at even higher levels than those achieved with lentiviral vectors at the optimal titre of 0.4 MOI (sections 5.3.3.b.1 and 5.3.2), despite the nuclease being expressed by two strong promoters (i.e. EF1 σ core and CMV)⁶²⁰. Moreover, I observed that even though Cas9 was serially diluted with every cell division following electroporation, its transient but higher expression resulted in a similar extent of dystrophin protein correction as its constitutive expression (sections 5.3.2.c and 5.3.3.biii). These promising results indicated that the genomic integration of transgene DNA is not necessary to ensure a high level of editing, and that also an episomal CRISPR/Cas9 expression in skeletal muscle *in vivo* might be efficient in restoring dystrophin protein. An approach based on the transient CRISPR/Cas9 expression

would be safer when moving *in vivo*, as it would avoid the oncogenic risk associated with a random transgene integration into patient's genome, and could be potentially repeated multiple times.

7.3 IN VITRO CRISPR/CAS9 GOES IN VIVO: FUTURE PERSPECTIVES

7.3.1 Genome-wide analysis of off-target effects

The next step following on from my project would be testing CRISPR/Cas9 efficacy *in vivo*. However, safety checks, such as the assessment of off-target effects derived from the nuclease, would be needed first.

In fact, the degree of nuclease specificity in recognizing its target is a well-documented issue: off-target sites sharing a certain degree of homology with the target site can indeed be found either in non-coding and coding regions of genes. As a result of nonspecific nuclease activity, these genes might be inactivated or mutated by the introduction of insertions and/or deletions⁶²¹. Similarly, the introduction of indels in off-target sites located within regulatory regions such as enhancers/repressors can, in turn, disrupt the homeostasis of several other genes⁶²². Tumorigenesis might also occur if perturbed genes are involved in controlling cancerous cells either as oncogenes or tumour suppressor-genes.

Most of the *in silico* tools for the prediction of potential off-target sites developed up to date are not completely accurate. The off-target sites listed by the algorithm of Tsai *et al.*⁶²³, for example, complemented those predicted by the different algorithm used by Frock *et al.*⁶²⁴ and vice-versa. This suggested that neither tool could predict all of the off-targets with absolute certainty (even though the most recent tool developed by Tsai offers an off-target profile resembling the one obtained following whole genome sequencing)⁶²⁵.

The chances of having non-specific editing might be reduced using Cas9 isolated from different organisms, as these would rely on PAM motifs (the sequences that defines each nuclease target) less common than the trinucleotide 5'-NGG-3'. As an example, the Cas9 nuclease isolated from *S. aureus* recognizes the more rare 5'-NNGRR-3' motif, where N and R stays for any nucleotide and a purine, respectively⁶²⁶.

Finally, as *in silico* tools are developed on reference genomes, they do not consider the huge number of sites (up to 5 million) that have been estimated to vary in the genetic background of each individual⁶²⁷. Lessard *et al.* recently pointed out that single nucleotide polymorphisms or copy number variation specific for each person might either alter, create or destroy the PAM motif⁶²⁷. Furthermore, these can as well disrupt

on-target sequences and/or generate additional off-target sites⁶²⁵. All these scenarios would affect the final outcome of the therapeutic approach *in vivo*, resulting either in inefficient treatment or severe adverse effects⁶²⁷. To guarantee the best nuclease design, whole genome sequencing of individual patients would be recommended⁶²⁵, although at present this is not a very practical solution as it would be time-consuming and expensive.

Not even the combinatorial approach of *in-silico* predictions and whole-genome sequencing analysis might reflect what actually occurs upon CRISPR/Cas9 targeting *in vivo*. These tools, in fact, do not account for the epigenetic status of the target cells, which determines the degree of chromatin accessibility to the nuclease^{628,629}. In fact, different cell types have specific chromatin conformation in accordance with their expressed genes. This might limit the ability of Cas9 to access any genomic region with the same efficiency⁵⁴⁶.

While alternative strategies and optimization of the off-target effects analysis are ongoing, an ideal nuclease to be tested *in vivo* would be expected to induce the therapeutic benefit with a minimal profile of off-target effects.

7.3.2. Considerations for future CRISPR/Cas9-based therapies

The choice of which system to use for CRISPR/Cas9 delivery *in vivo* is a fundamental factor that has to be considered depending on which cells/tissues have to be targeted. For example, the delivery of CRISPR/Cas9 via electroporation (which I tested *in vitro*) could be applied for *ex vivo* cell-based therapies (section 1.6.2.c), but cannot be used to target muscles body-wide *in vivo*. Although many other non-viral strategies are being continuously improved and adapted for the delivery of the nuclease in different formats, for targeting muscle *in vivo* I would opt for the well characterized AAV vectors that are in use in several clinical trials (section 1.6.2.d).

Several advantages related to AAV vectors allowed me to consider them as the best choice for *in vivo* CRISPR/Cas9 delivery.

One factor is that the transgene expressed by AAV vectors remains mostly episomal. This would result in a CRISPR/Cas9 expression limited in time, overcoming the risk of insertional mutagenesis associated with the integrating vectors²⁷².

Also, despite having a limited packaging ability which would require the substitution of the *Streptococcus pyogenes* Cas9 with the smaller Cas9 protein derived from *Staphylococcus aureus*⁶³⁰, AAV vectors have instead a relatively low immunogenicity and are much better-tolerated when delivered systemically, as this is the only administration route that would allow the targeting of muscles body-wide. This feature

does not apply to adenoviral vectors, which are instead associated with an high immunogenic risk²⁶³.

Finally, as each AAV serotype has a preferential tissue tropism, different AAV vectors can be used to target the different tissues and organs affected in DMD^{275,276}. AAVs of serotypes 6, 8 and 9 successfully transduce heart and muscle *in vivo* in the *mdx* mouse model, as independently observed by several groups^{387,631–633}.

However, the works published by Arnett *et al.* and Tabebordbar *et al.*^{387,634} underlined the poor tropism of AAVs towards satellite cells, in which a level of editing less than 5% was observed^{387,634}. This might limit the extent of dystrophin correction that might be achieved, underlining the need to engineer AAV vectors so that an efficient targeting of satellite cells would be achieved. In fact, an ideal therapeutic solution would consider the editing of dystrophin both in satellite cells (section 1.5.1 and 1.6.1.c) and in myofibers to obtain a maximal level of protein repair.

Once the vector for expressing CRISPR/Cas9 and a delivery route have been identified, a further criterion to be considered *in vivo* is the age at which the nuclease is administered. I believe that providing the nuclease to DMD patients at the appearance of symptoms (i.e. at an early stage of the disease) would be more beneficial than later in time, considering that more muscle tissue would be available and preserved from fibrosis. Moreover, once the AAV targeting of satellite cells will be possible, the muscle regeneration observed at the initial stages of the disease will allow a quicker replacement of the damaged fibres with fully functional ones⁶³⁵.

An additional question to be answered is how much functional dystrophin is needed to restore muscle function. Van Putten *et al.* investigated how much dystrophin is required to ameliorate the phenotype of the *mdx-Xist^{Ahs}* mice, a disease model which exhibit minimal but varying dystrophin levels as it derives from the breeding of *Xist^{Ahs}* females (carrying a mutation in the *Xist* gene, responsible for X-inactivation⁶³⁶) and *mdx* males. She showed that very low levels of wild-type dystrophin (starting from a minimum of 4%) are enough to improve performance (measured as endurance), muscle survival as well as cardiac and motor function of *mdx-Xist^{Ahs}* mice, even though at least 20% of dystrophin expression had to be achieved before she could assess the full protection of muscle against eccentric contractions^{637–639}. Similar results were reported by Godfrey *et al.* Her work on *mdx* mice showed that as few as 15% dystrophin homogeneously expressed in the muscle fibres within the muscle is enough to protect against muscle damage caused by eccentric exercise⁶⁴⁰, similarly to what seen upon an higher (73%) but heterogenous dystrophin expression across muscle fibres (only 65% of dystrophin

positive fibres)⁶⁴¹. These findings indicated that, even though it is effective at very low levels, dystrophin has to be expressed in the majority of the muscle fibres to allow a functional improvement.

However, mice are not bipedal and therefore they distribute force across four limbs. As a consequence, muscle damage in response to load is expected to be less in mice than humans. The study published by Neri *et al.* showed that, in humans, a level of dystrophin expression between 29% (which is almost double the amount reported by Godfrey *et al.* ⁶⁴⁰) and 57 % of wild-type dystrophin is needed to completely prevent muscular dystrophy when expressed in the majority of muscle fibres within the muscle⁶⁴².

In light of these findings, the single nuclease approach developed in my study could be successful *in vivo* if resulting in levels of restored dystrophin expression higher or similar to those mentioned by Neri *et al.*

7.3.3. *In vivo* CRISPR/Cas9 immunity: further considerations

A recent study by Charlesworth *et al.* demonstrated that, in healthy human adults, anti-Cas9 immunity exists to Cas9 nucleases derived both from *Streptococcus pyogenes* and *Staphylococcus aureus*, as these bacteria frequently colonize and cause diseases in humans⁶⁴³. Even though the humoral response towards the intracellular Cas9 would not be harmful for Cas9-expressing cells, these can be killed by the cellular immune response, which triggers the secretion of interferon- γ by the CD8+ T lymphocytes⁶⁴⁴. The risk of a strong immune response might be mitigated by the use of low immunogenic vectors such as AAV. However, the pro-inflammatory environment which is characteristic of dystrophic skeletal muscle⁶⁴⁵ might exacerbate Cas9 immunity and increase the extent of nuclease destruction. Therefore, the use of immune-suppressors during the time of nuclease expression might be needed. Alternatively, a minimal therapeutic level of Cas9 expression might increase the tolerance of Cas9 and limit the triggering of the immune response⁶⁴⁶. To ensure a low but still effective CRISPR/Cas9 expression in the clinical context, I would modulate the nuclease expression by means of a tissue specific and inducible promoter.

Tissue specificity would restrict the Cas9 expression to the tissue of interest (i.e. skeletal muscle), similarly to what was recently shown by Bengtsson *et al.*⁶³³, who used a custom-designed⁶⁴⁷ muscle-specific promoter to limit the expression of Cas9 to skeletal and cardiac muscles of *mdx*^{4cv} mice (nonsense mutation in dystrophin exon 53). The possibility to express Cas9 only in desired tissues would also limit potential off-target effects (section 1.8) that might arise if targeting occurs in tissues or cells having a different epigenetic landscape.

The use of an inducible promoter, instead, would allow me to control the timing of the nuclease expression, so that it could be maintained only until required. This would reduce the risk of off-target effects and lower the chances of triggering immune response. Ideally, I would put Cas9 under the control of a promoter that could be induced by means of a drug. In this way, patients would receive the drug for a limited period of time required for Cas9 to be expressed and, following drug clearance, Cas9 expression would be silenced.

7.4 CONCLUDING REMARKS

The scientific approach developed in this thesis allowed me to confirm that single CRISPR/Cas9 nucleases can correct dystrophin duplications *in vitro* in multiple cell types, providing the basis for *in vivo* therapeutic approaches tailored to the 10-15% of patients carrying DMD duplications.

Moreover, I could validate this approach both by using viral and non-viral strategies and the comparison of these strategies allowed me to infer that a considerable but transient expression of Cas9 can be as effective as its continuous but poor expression. Therefore, *in vivo* therapeutic designs should aim to a “hit-and-run” expression of Cas9, so that long term risk of constitutively active endonucleases would be reduced.

Further optimizations aimed to increase the safety of this approach are for sure needed before translating this approach into clinic for DMD. However, I believe that the scientific progress will speed up this process and will reveal the widespread therapeutic applicability of this tool. In the field of DMD, the refinements of the CRISPR/Cas9 editing tool is likely to proceed in conjunction with the rapidly evolving AAV gene therapies. I believe that it is realistic to expect that, with new discoveries on improving safety and specificity, genome editing for DMD including duplication will be achieved in the future.

BIBLIOGRAPHY

1. Bladen, C. L. *et al.* The TREAT-NMD DMD Global Database: Analysis of More than 7,000 Duchenne Muscular Dystrophy Mutations. *Hum. Mutat.* **36**, 395–402 (2015).
2. Hu, X. Y., Ray, P. N., Murphy, E. G., Thompson, M. W. & Worton, R. G. Duplicational mutation at the Duchenne muscular dystrophy locus: its frequency, distribution, origin, and phenotype/genotype correlation. *Am. J. Hum. Genet.* **46**, 682–95 (1990).
3. Guiraud, S. *et al.* The Pathogenesis and Therapy of Muscular Dystrophies. *Annu. Rev. Genomics Hum. Genet.* **16**, 281–308 (2015).
4. Stankiewicz, P. & Lupski, J. R. Structural Variation in the Human Genome and its Role in Disease. *Annu. Rev. Med.* **61**, 437–455 (2010).
5. Harteveld, C. L., Refaldi, C., Cassinerio, E., Cappellini, M. D. & Giordano, P. C. Segmental duplications involving the α -globin gene cluster are causing β -thalassemia intermedia phenotypes in β -thalassemia heterozygous patients. *Blood Cells, Mol. Dis.* **40**, 312–316 (2008).
6. Nigro, V. & Piluso, G. Spectrum of muscular dystrophies associated with sarcolemmal-protein genetic defects. *Biochim. Biophys. Acta* (2014). doi:10.1016/j.bbadis.2014.07.023
7. Mora, M. Fibrous-adipose replacement in skeletal muscle biopsy. *Eur. Heart J.* **10**, 103–104 (1989).
8. Mercuri, E. & Muntoni, F. Muscular dystrophies. *Lancet* **381**, 845–60 (2013).
9. Bell, C. *The nervous system of the human body: embracing the papers delivered to the Royal Society on the subject of the nerves.* (Longman Rees Orme Brown and Green, 1830).
10. Conte & Conte, G. Scrofolo del sistema muscolare. *Ann. Clin. dell' Osp. degli 'Incurabili di Napoli* **2**, (1836).
11. Meryon, E. On Granular and Fatty Degeneration of the Voluntary Muscles. *Med. Chir. Trans.* **35**, 73–84.1 (1852).
12. Duchenne, G.-B. (1806-1875). A. du texte. De la paralysie musculaire pseudo-hypertrophique, ou paralysie myo-sclérosique / par le Dr Duchenne (de Boulogne). (1868).
13. Emery, A. Duchenne muscular dystrophy or Meryon's disease. *Lancet* **357**,

1529 (2001).

14. Chelly, J. *et al.* De novo DNA microdeletion in a girl with Turner syndrome and Duchenne muscular dystrophy. *Hum. Genet.* **74**, 193–6 (1986).
15. Pena, S. D., Karpati, G., Carpenter, S. & Fraser, F. C. The clinical consequences of X-chromosome inactivation: Duchenne muscular dystrophy in one of monozygotic twins. *J. Neurol. Sci.* **79**, 337–44 (1987).
16. Pegoraro, E. *et al.* Detection of new paternal dystrophin gene mutations in isolated cases of dystrophinopathy in females. *Am. J. Hum. Genet.* **54**, 989–1003 (1994).
17. Ricotti, V. *et al.* Long-term benefits and adverse effects of intermittent versus daily glucocorticoids in boys with Duchenne muscular dystrophy. *J. Neurol. Neurosurg. Psychiatry* **84**, 698–705 (2013).
18. Zhu, Y. *et al.* Serum Enzyme Profiles Differentiate Five Types of Muscular Dystrophy. *Dis. Markers* **2015**, 1–7 (2015).
19. Mendell, J. R. *et al.* Evidence-based path to newborn screening for duchenne muscular dystrophy. *Ann. Neurol.* **71**, 304–313 (2012).
20. Ebashi, S., Toyokura, Y., Momoi, H. & Sugita, H. High Creatine Phosphokinase Activity of Sera with Progressive Muscular Dystrophy. *J. Biochem.* **46**, 103–104 (1959).
21. Echegaray, M. & Rivera, M. A. Role of Creatine Kinase Isoenzymes on Muscular and Cardiorespiratory Endurance. *Sport. Med.* **31**, 919–934 (2001).
22. Zatz, M., Shapiro, L. J., Campion, D. S., Oda, E. & Kaback, M. M. Serum pyruvate-kinase (PK) and creatine-phosphokinase (CPK) in progressive muscular dystrophies. *J. Neurol. Sci.* **36**, 349–362 (1978).
23. Moat, S. J., Bradley, D. M., Salmon, R., Clarke, A. & Hartley, L. Newborn bloodspot screening for Duchenne muscular dystrophy: 21 years experience in Wales (UK). *Eur. J. Hum. Genet.* **21**, 1049–53 (2013).
24. Clinical Lecture ON PSEUDO-HYPERTROPHIC MUSCULAR PARALYSIS. *Lancet* **114**, 73–75 (1879).
25. Kornegay, J. N. *et al.* The Paradox of Muscle Hypertrophy in Muscular Dystrophy. *Phys. Med. Rehabil. Clin. N. Am.* **23**, 149–172 (2012).
26. Torriani, M. *et al.* Lower leg muscle involvement in Duchenne muscular dystrophy: an MR imaging and spectroscopy study. *Skeletal Radiol.* **41**, 437–45

(2012).

27. Mummery, C. J., Copeland, S. & Rose, M. R. in *The Cochrane Database of Systematic Reviews* (ed. Mummery, C. J.) CD003278 (John Wiley & Sons, Ltd, 2003). doi:10.1002/14651858.CD003278
28. Scher, D. M. & Mubarak, S. J. Surgical prevention of foot deformity in patients with Duchenne muscular dystrophy. *J. Pediatr. Orthop.* **22**, 384–91
29. Mazzone, E. *et al.* Functional changes in Duchenne muscular dystrophy: a 12-month longitudinal cohort study. *Neurology* **77**, 250–6 (2011).
30. Kinal M, Main M, Mercuri E & Muntoni F. Evolution of abnormal postures in Duchenne muscular dystrophy. *Ann Indian Acad Neurol* **10**, (2007).
31. Shneerson, J. M. & Simonds, A. K. Noninvasive ventilation for chest wall and neuromuscular disorders. *Eur. Respir. J.* **20**, 480–7 (2002).
32. Kim, E. Y. *et al.* Correlation of serum creatine kinase level with pulmonary function in duchenne muscular dystrophy. *Ann. Rehabil. Med.* **41**, 306–312 (2017).
33. Wagner, M. B., Vignos, P. J. & Carlozzi, C. Duchenne muscular dystrophy: A study of wrist and hand function. *Muscle Nerve* **12**, 236–244 (1989).
34. Johnson, E. R., Jr, W. M. F., Lieberman, J. S., Fowler, W. M. J. & Lieberman, J. S. *Contractures in neuromuscular disease. Archives of Physical Medicine and Rehabilitation* **73**, (1992).
35. Fayssoil, A., Abasse, S. & Silverston, K. Cardiac Involvement Classification and Therapeutic Management in Patients with Duchenne Muscular Dystrophy. *J. Neuromuscul. Dis.* **4**, 17–23 (2017).
36. Abe, Y. Myocardial fibrosis of the left ventricular posterior wall can be a target for early detection of cardiac involvement in patients with Duchenne muscular dystrophy. *J. Cardiol.* **66**, 203–4 (2015).
37. Duboc, D. *et al.* Perindopril preventive treatment on mortality in Duchenne muscular dystrophy: 10 years' follow-up. *Am. Heart J.* **154**, 596–602 (2007).
38. Silva, M. C. *et al.* Myocardial Fibrosis Progression in Duchenne and Becker Muscular Dystrophy. *JAMA Cardiol.* **2**, 190 (2017).
39. Anand, A. *et al.* Dystrophin induced cognitive impairment: mechanisms, models and therapeutic strategies. *Ann. Neurosci.* **22**, 108–18 (2015).

40. Ash, A., Booth-Wynne, L. & Anthony, K. Brain involvement in Duchenne muscular dystrophy: a role for dystrophin isoform Dp71 in cell migration and proliferation. *Neuromuscul. Disord.* **27**, S114–S115 (2017).
41. Ricotti, V., Roland, J., Roberts, G. & Muntoni, F. Dystrophin and the brain. doi:10.1111/j.1469-8749.2010.03836.x
42. Felisari, G. *et al.* Loss of Dp140 dystrophin isoform and intellectual impairment in Duchenne dystrophy. *Neurology* **55**, 559–64 (2000).
43. Daoud, F. *et al.* Analysis of Dp71 contribution in the severity of mental retardation through comparison of Duchenne and Becker patients differing by mutation consequences on Dp71 expression. *Hum. Mol. Genet.* **18**, 3779–3794 (2009).
44. Pane, M. *et al.* Attention Deficit Hyperactivity Disorder and Cognitive Function in Duchenne Muscular Dystrophy: Phenotype-Genotype Correlation. *J. Pediatr.* **161**, 705–709.e1 (2012).
45. Filippo, T. Di, Parisi, L. & Roccella, M. Psychological aspects in children affected by duchenne de boulogne muscular dystrophy. *Ment. Illn.* **4**, e5 (2012).
46. Lo Cascio, C. M. *et al.* Gastrointestinal Dysfunction in Patients with Duchenne Muscular Dystrophy. *PLoS One* **11**, e0163779 (2016).
47. Loufrani, L. *et al.* Flow (shear stress)-induced endothelium-dependent dilation is altered in mice lacking the gene encoding for dystrophin. *Circulation* **103**, 864–70 (2001).
48. Rodríguez-Cruz, M. *et al.* Evidence of Insulin Resistance and Other Metabolic Alterations in Boys with Duchenne or Becker Muscular Dystrophy. *Int. J. Endocrinol.* **2015**, 867273 (2015).
49. Matsumura, T., Saito, T., Fujimura, H., Shinno, S. & Sakoda, S. A longitudinal cause-of-death analysis of patients with Duchenne muscular dystrophy. *Rinsho Shinkeigaku* **51**, 743–750 (2011).
50. Motoki, T. *et al.* Treatable renal failure found in non-ambulatory Duchenne muscular dystrophy patients. *Neuromuscul. Disord.* **25**, 754–757 (2015).
51. Alhaswani, Z., Thomas, C. & Roper, H. Renal impairment in two patients with Duchenne muscular dystrophy. *Neuromuscul. Disord.* **26**, S122 (2016).
52. Simonds, A. K., Muntoni, F., Heather, S. & Fielding, S. Impact of nasal ventilation on survival in hypercapnic Duchenne muscular dystrophy. *Thorax*

53, 949–52 (1998).

53. Eagle, M. *et al.* Managing Duchenne muscular dystrophy – The additive effect of spinal surgery and home nocturnal ventilation in improving survival. *Neuromuscul. Disord.* **17**, 470–475 (2007).
54. Eagle, M. *et al.* Survival in Duchenne muscular dystrophy: improvements in life expectancy since 1967 and the impact of home nocturnal ventilation. *Neuromuscul. Disord.* **12**, 926–9 (2002).
55. Rall, S. & Grimm, T. Survival in Duchenne muscular dystrophy. *Acta Myol. myopathies cardiomyopathies Off. J. Mediterr. Soc. Myol.* **31**, 117–20 (2012).
56. BECKER, P. E. & KIENER, F. [A new x-chromosomal muscular dystrophy]. *Arch. Psychiatr. Nervenkr. Z. Gesamte Neurol. Psychiatr.* **193**, 427–48 (1955).
57. BECKER, P. E. Two families of benign sex-linked recessive muscular dystrophy. *Rev. Can. Biol.* **21**, 551–66
58. Kunkel, L. M. & Hoffman, E. P. Duchenne/Becker muscular dystrophy: a short overview of the gene, the protein, and current diagnostics. *Br. Med. Bull.* **45**, 630–43 (1989).
59. Kunkel, L. M., Monaco, A. P., Middlesworth, W., Ochs, H. D. & Latt, S. A. Specific cloning of DNA fragments absent from the DNA of a male patient with an X chromosome deletion. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4778–82 (1985).
60. Davies, K. E. *et al.* Linkage analysis of two cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome. *Nucleic Acids Res.* **1**, (1983).
61. Worton, R. G., Duff, C., Sylvester, J. E., Schmickel, R. D. & Willard, H. F. Duchenne muscular dystrophy involving translocation of the *dmd* gene next to ribosomal RNA genes. *Science* **224**, 1447–9 (1984).
62. Nudel, U. *et al.* Duchenne muscular dystrophy gene product is not identical in muscle and brain. *Nature* **337**, 76–78 (1989).
63. Lidov, H. G. W., Byers, T. J., Watkins, S. C. & Kunkel, L. M. Localization of dystrophin to postsynaptic regions of central nervous system cortical neurons. *Nature* **348**, 725–728 (1990).
64. Holder, E., Maeda, M. & Bies, R. D. Expression and regulation of the dystrophin Purkinje promoter in human skeletal muscle, heart, and brain. *Hum. Genet.* **97**, 232–9 (1996).

65. Torelli, S., Ferlini, A., Obici, L., Sewry, C. & Muntoni, F. Expression, regulation and localisation of dystrophin isoforms in human foetal skeletal and cardiac muscle. *Neuromuscul. Disord.* **9**, 541–551 (1999).
66. Koenig, M. *et al.* Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* **50**, 509–17 (1987).
67. Pillers, D.-A. M. *et al.* Dystrophin expression in the human retina is required for normal function as defined by electroretinography. *Nat. Genet.* **4**, 82–86 (1993).
68. Lidov, H. G., Selig, S. & Kunkel, L. M. Dp140: a novel 140 kDa CNS transcript from the dystrophin locus. *Hum. Mol. Genet.* **4**, 329–35 (1995).
69. Byers, T. J., Lidov, H. G. W. & Kunkel, L. M. An alternative dystrophin transcript specific to peripheral nerve. *Nat. Genet.* **4**, 77–81 (1993).
70. Bar, S. *et al.* A novel product of the Duchenne muscular dystrophy gene which greatly differs from the known isoforms in its structure and tissue distribution. *Biochem. J.* **272**, 557–60 (1990).
71. Rapaport, D. *et al.* Characterization and cell type distribution of a novel, major transcript of the Duchenne Muscular Dystrophy gene. *Differentiation* **49**, 187–193 (1992).
72. Feener, C. A., Koenig, M. & Kunkel, L. M. Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy terminus. *Nature* **338**, 509–511 (1989).
73. Ferlini, A., Neri, M. & Gualandi, F. The medical genetics of dystrophinopathies: molecular genetic diagnosis and its impact on clinical practice. *Neuromuscul. Disord.* **23**, 4–14 (2013).
74. Hu, X. Y., Burghes, A. H., Bulman, D. E., Ray, P. N. & Worton, R. G. Evidence for mutation by unequal sister chromatid exchange in the Duchenne muscular dystrophy gene. *Am. J. Hum. Genet.* **44**, 855–63 (1989).
75. Hu, X. Y., Ray, P. N., Murphy, E. G., Thompson, M. W. & Worton, R. G. Duplicational mutation at the Duchenne muscular dystrophy locus: its frequency, distribution, origin, and phenotype/genotype correlation. *Am. J. Hum. Genet.* **46**, 682–95 (1990).
76. Hu, X. Y., Ray, P. N. & Worton, R. G. Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J.* **10**, 2471–7 (1991).

77. Grimm, T. *et al.* On the origin of deletions and point mutations in Duchenne muscular dystrophy: most deletions arise in oogenesis and most point mutations result from events in spermatogenesis. *J. Med. Genet.* **31**, 183–186 (1994).
78. White, S. J. *et al.* Duplications in the DMD gene. *Hum. Mutat.* **27**, 938–45 (2006).
79. Den Dunnen, J. T. *et al.* Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am. J. Hum. Genet.* **45**, 835–47 (1989).
80. Hiraishi, Y., Kato, S., Ishihara, T. & Takano, T. Quantitative Southern blot analysis in the dystrophin gene of Japanese patients with Duchenne or Becker muscular dystrophy: a high frequency of duplications. *J. Med. Genet.* **29**, 897–901 (1992).
81. Galvagni, F. *et al.* A study on duplications of the dystrophin gene: evidence of a geographical difference in the distribution of breakpoints by intron. *Hum. Genet.* **94**, 83–7 (1994).
82. Angelini, C., Beggs, A. H., Hoffman, E. P., Fanin, M. & Kunkel, L. M. Enormous dystrophin in a patient with Becker muscular dystrophy. *Neurology* **40**, 808–12 (1990).
83. Gold, R., Kreutz, W., Bettecken, T., Reichmann, H. & Møller, C. R. A 400-kb tandem duplication within the dystrophin gene leads to severe Becker muscular dystrophy. *J. Neurol.* **241**, 331–334 (1994).
84. Wu, B. *et al.* Identification of a novel DMD duplication identified by a combination of MLPA and targeted exome sequencing. *Mol. Cytogenet.* **10**, 8 (2017).
85. Gualandi, F. *et al.* Transcriptional behavior of DMD gene duplications in DMD/BMD males. *Hum. Mutat.* **30**, E310–E319 (2009).
86. Ishmukhametova, A. *et al.* Dissecting the Structure and Mechanism of a Complex Duplication-Triplication Rearrangement in the *DMD* Gene. *Hum. Mutat.* **34**, 1080–1084 (2013).
87. López-Hernández, L. B. *et al.* A novel noncontiguous duplication in the DMD gene escapes the ‘reading-frame rule’. *J. Genet.* **93**, 225–9 (2014).
88. Monaco, A. P., Bertelson, C. J., Liechti-Gallati, S., Moser, H. & Kunkel, L. M. An explanation for the phenotypic differences between patients bearing partial

- deletions of the DMD locus. *Genomics* **2**, 90–5 (1988).
89. Muntoni, F., Strong, P. N. & Lewis, J. Transcription of the dystrophin gene in Duchenne muscular dystrophy muscle. **2522**, 95–98 (1989).
 90. Hoffman, E. P., Brown, R. H. & Kunkel, L. M. Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**, 919–28 (1987).
 91. Minetti, C., Beltrame, F., Marcenaro, G. & Bonilla, E. Dystrophin at the plasma membrane of human muscle fibers shows a costameric localization. *Neuromuscul. Disord.* **2**, 99–109 (1992).
 92. Kong, J. & Anderson, J. E. Dystrophin is required for organizing large acetylcholine receptor aggregates. *Brain Res.* **839**, 298–304 (1999).
 93. Law, D. J. & Tidball, J. G. Dystrophin deficiency is associated with myotendinous junction defects in pre necrotic and fully regenerated skeletal muscle. *Am. J. Pathol.* **142**, 1513–23 (1993).
 94. Way, M., Pope, B., Cross, R. A., Kendrick-Jones, J. & Weeds, A. G. Expression of the N-terminal domain of dystrophin in *E. coli* and demonstration of binding to F-actin. **301**, 243–245 (1992).
 95. Stone, M. R., O'Neill, A., Catino, D. & Bloch, R. J. Specific interaction of the actin-binding domain of dystrophin with intermediate filaments containing keratin 19. *Mol. Biol. Cell* **16**, 4280–93 (2005).
 96. Koenig, M. & Kunkel, L. M. Detailed analysis of the repeat domain of dystrophin reveals four potential hinge segments that may confer flexibility. *J. Biol. Chem.* **265**, 4560–6 (1990).
 97. Legrand, B., Giudice, E., Nicolas, A., Delalande, O. & Le Rumeur, E. Computational study of the human dystrophin repeats: interaction properties and molecular dynamics. *PLoS One* **6**, e23819 (2011).
 98. Blake, D. J. *et al.* Coiled-coil regions in the carboxy-terminal domains of dystrophin and related proteins: potentials for protein-protein interactions. *Trends Biochem. Sci.* **20**, 133–135 (1995).
 99. Belanto, J. J. *et al.* Microtubule binding distinguishes dystrophin from utrophin. *Proc. Natl. Acad. Sci.* **111**, 5723–5728 (2014).
 100. Amann, K. J., Renley, B. A. & Ervasti, J. M. A cluster of basic repeats in the dystrophin rod domain binds F-actin through an electrostatic interaction. *J. Biol. Chem.* **273**, 28419–23 (1998).

101. Adams, M. E., Odom, G. L., Kim, M. J., Chamberlain, J. S. & Froehner, S. C. Syntrophin binds directly to multiple spectrin-like repeats in dystrophin and mediates binding of nNOS to repeats 16–17. *Hum. Mol. Genet.* **27**, 2978–2985 (2018).
102. Le Rumeur, E. *et al.* Interaction of Dystrophin Rod Domain with Membrane Phospholipids. *J. Biol. Chem.* **278**, 5993–6001 (2003).
103. Zhao, J. *et al.* Dystrophin contains multiple independent membrane-binding domains. *Hum. Mol. Genet.* **25**, 3647–3653 (2016).
104. Yamashita, K. *et al.* The 8th and 9th tandem spectrin-like repeats of utrophin cooperatively form a functional unit to interact with polarity-regulating kinase PAR-1b. *Biochem. Biophys. Res. Commun.* **391**, 812–817 (2010).
105. Ilsley, J. L., Sudol, M. & Winder, S. J. The interaction of dystrophin with beta-dystroglycan is regulated by tyrosine phosphorylation. *Cell. Signal.* **13**, 625–32 (2001).
106. Ilsley, J. L., Sudol, M. & Winder, S. J. The WW domain: linking cell signalling to the membrane cytoskeleton. *Cell. Signal.* **14**, 183–9 (2002).
107. Jung, D., Yang, B., Meyer, J., Chamberlain, J. S. & Campbell, K. P. Identification and Characterization of the Dystrophin Anchoring Site on α -Dystroglycan*.
108. Rentschler, S. *et al.* The WW Domain of Dystrophin Requires EF-Hands Region to Interact with β -Dystroglycan. *Biol. Chem.* **380**, 431–42 (1999).
109. Ponting, C. P., Blake, D. J., Davies, K. E., Kendrick-Jones, J. & Winder, S. J. ZZ and TAZ: new putative zinc fingers in dystrophin and other proteins. *Trends Biochem. Sci.* **21**, 11–13 (1996).
110. Ayalon, G., Davis, J. Q., Scotland, P. B. & Bennett, V. An Ankyrin-Based Mechanism for Functional Organization of Dystrophin and Dystroglycan. *Cell* **135**, 1189–1200 (2008).
111. Sadoulet-Puccio, H. M., Rajala, M. & Kunkel, L. M. Dystrobrevin and dystrophin: an interaction through coiled-coil motifs. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 12413–8 (1997).
112. Swiderski, K. *et al.* Phosphorylation within the cysteine-rich region of dystrophin enhances its association with β -dystroglycan and identifies a potential novel therapeutic target for skeletal muscle wasting. *Hum. Mol. Genet.* **23**, 6697–711

(2014).

113. Shemanko, C. S., Sanghera, J. S., Milner, R. E., Pelech, S. & Michalak, M. Phosphorylation of the carboxyl terminal region of dystrophin by mitogen-activated protein (MAP) kinase. *Molecular and Cellular Biochemistry* **152**, 63–70 (1995).
114. Footer, M. J., Kerssemakers, J. W. J., Theriot, J. A. & Dogterom, M. Direct measurement of force generation by actin filament polymerization using an optical trap. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2181–6 (2007).
115. Elhanany-Tamir, H. *et al.* Organelle positioning in muscles requires cooperation between two KASH proteins and microtubules. *J. Cell Biol.* **198**, 833–846 (2012).
116. Percival, J. M. & Froehner, S. C. Golgi Complex Organization in Skeletal Muscle: A Role for Golgi-Mediated Glycosylation in Muscular Dystrophies? *Traffic* **8**, 184–194 (2007).
117. Campbell, K. P. & Kahl, S. D. Association of dystrophin and an integral membrane glycoprotein. *Nature* **338**, 259–62 (1989).
118. Ibraghimov-Beskrovnaya, O. *et al.* Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **355**, 696–702 (1992).
119. Ervasti, J. M. & Campbell, K. P. A role for the dystrophin-glycoprotein complex as a transmembrane linker between laminin and actin. *J. Cell Biol.* **122**, 809–23 (1993).
120. Bowe, M. A., Mendis, D. B. & Fallon, J. R. The small leucine-rich repeat proteoglycan biglycan binds to alpha-dystroglycan and is upregulated in dystrophic muscle. *J. Cell Biol.* **148**, 801–10 (2000).
121. Miller, G., Wang, E. L., Nassar, K. L., Peter, A. K. & Crosbie, R. H. Structural and functional analysis of the sarcoglycan–sarco-span subcomplex. *Exp. Cell Res.* **313**, 639–651 (2007).
122. Yoshida, M. *et al.* Biochemical evidence for association of dystrobrevin with the sarcoglycan-sarco-span complex as a basis for understanding sarcoglycanopathy. *Hum. Mol. Genet.* **9**, 1033–1040 (2000).
123. Mercado, M. L. *et al.* Biglycan regulates the expression and sarcolemmal localization of dystrobrevin, syntrophin, and nNOS. *FASEB J.* **20**, 1724–1726 (2006).

124. Marshall, J. L. *et al.* Dystrophin and utrophin expression require sarcospan: loss of $\alpha 7$ integrin exacerbates a newly discovered muscle phenotype in sarcospan-null mice. *Hum. Mol. Genet.* **21**, 4378–4393 (2012).
125. Oak, S. A., Russo, K., Petrucci, T. C. & Jarrett, H. W. Mouse $\alpha 1$ -Syntrophin Binding to Grb2: Further Evidence of a Role for Syntrophin in Cell Signaling [†]. *Biochemistry* **40**, 11270–11278 (2001).
126. Balasubramanian, S., Fung, E. T. & Huganir, R. L. Characterization of the tyrosine phosphorylation and distribution of dystrobrevin isoforms. *FEBS Lett.* **432**, 133–140 (1998).
127. Grady, R. M. *et al.* Tyrosine-phosphorylated and nonphosphorylated isoforms of α -dystrobrevin. *J. Cell Biol.* **160**, 741–752 (2003).
128. Okumura, A., Nagai, K. & Okumura, N. Interaction of $\alpha 1$ -syntrophin with multiple isoforms of heterotrimeric G protein α subunits. *FEBS J.* **275**, 22–33 (2008).
129. Chen, Z., Hague, C., Hall, R. A. & Minneman, K. P. Syntrophins Regulate α_{1D} - Adrenergic Receptors through a PDZ Domain-mediated Interaction. *J. Biol. Chem.* **281**, 12414–12420 (2006).
130. Gee, S. H. *et al.* Interaction of muscle and brain sodium channels with multiple members of the syntrophin family of dystrophin-associated proteins. *J. Neurosci.* **18**, 128–37 (1998).
131. Fanning, A. S. & Anderson, J. M. PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane. *J. Clin. Invest.* **103**, 767–72 (1999).
132. Vandebrouck, A. *et al.* Regulation of capacitative calcium entries by $\alpha 1$ -syntrophin: association of TRPC1 with dystrophin complex and the PDZ domain of $\alpha 1$ -syntrophin. *FASEB J.* **21**, 608–617 (2007).
133. Madhavan, R. & Jarrett, H. W. Phosphorylation of dystrophin and alpha-syntrophin by Ca(2+)-calmodulin dependent protein kinase II. *Biochim. Biophys. Acta* **1434**, 260–74 (1999).
134. Hill, M., Wernig, A. & Goldspink, G. Muscle satellite (stem) cell activation during local tissue injury and repair. *J. Anat.* **203**, 89–99 (2003).
135. Kuang, S. *et al.* Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* **129**, 999–1010 (2007).

136. PARTRIDGE, T. A., GROUNDS, M. & SLOPER, J. C. Evidence of fusion between host and donor myoblasts in skeletal muscle grafts. *Nature* **273**, 306–308 (1978).
137. Decary, S. *et al.* Replicative potential and telomere length in human skeletal muscle: implications for satellite cell-mediated gene therapy. *Hum. Gene Ther.* **8**, 1429–38 (1997).
138. Boldrin, L., Zammit, P. S. & Morgan, J. E. Satellite cells from dystrophic muscle retain regenerative capacity. *Stem Cell Res.* **14**, 20–9 (2015).
139. Dumont, N. A. *et al.* Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division. *Nat. Med.* **21**, 1455–63 (2015).
140. Mann, C. J. *et al.* Aberrant repair and fibrosis development in skeletal muscle. *Skelet. Muscle* **1**, 21 (2011).
141. Dadgar, S. *et al.* Asynchronous remodeling is a driver of failed regeneration in Duchenne muscular dystrophy. *J. Cell Biol.* **207**, 139–58 (2014).
142. Spencer, M. J., Montecino-Rodriguez, E., Dorshkind, K. & Tidball, J. G. Helper (CD4+) and Cytotoxic (CD8+) T Cells Promote the Pathology of Dystrophin-Deficient Muscle. *Clin. Immunol.* **98**, 235–243 (2001).
143. Villalta, S. A., Nguyen, H. X., Deng, B., Gotoh, T. & Tidball, J. G. Shifts in macrophage phenotypes and macrophage competition for arginine metabolism affect the severity of muscle pathology in muscular dystrophy. *Hum. Mol. Genet.* **18**, 482–496 (2008).
144. Gorospe, J. R., Tharp, M., Demitsu, T. & Hoffman, E. P. Dystrophin-deficient myofibers are vulnerable to mast cell granule-induced necrosis. *Neuromuscul. Disord.* **4**, 325–33 (1994).
145. Porter, J. D. *et al.* A chronic inflammatory response dominates the skeletal muscle molecular signature in dystrophin-deficient mdx mice. *Hum. Mol. Genet.* **11**, 263–272 (2002).
146. Li, Y. *et al.* Transforming growth factor-beta1 induces the differentiation of myogenic cells into fibrotic cells in injured skeletal muscle: a key event in muscle fibrogenesis. *Am. J. Pathol.* **164**, 1007–19 (2004).
147. Melone, M. A. B., Peluso, G., Galderisi, U., Petillo, O. & Cotrufo, R. Increased expression of IGF-binding protein-5 in Duchenne Muscular Dystrophy (DMD) fibroblasts correlates with the fibroblast-induced downregulation of DMD myoblast growth: An in vitro analysis. *J. Cell. Physiol.* **185**, 143–153 (2000).

148. Stapleton, D. I. *et al.* Dysfunctional muscle and liver glycogen metabolism in mdx dystrophic mice. *PLoS One* **9**, e91514 (2014).
149. Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M. & Sweeney, H. L. Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3710–4 (1993).
150. Mallouk, N. *et al.* Elevated subsarcolemmal Ca²⁺ in mdx mouse skeletal muscle fibers detected with Ca²⁺-activated K⁺ channels. *Proc. Natl. Acad. Sci.* **97**, 4950–4955 (2000).
151. Millay, D. P. *et al.* Calcium influx is sufficient to induce muscular dystrophy through a TRPC-dependent mechanism. *Proc. Natl. Acad. Sci.* **106**, 19023 LP–19028 (2009).
152. Spencer, M. J. & Tidball, J. G. Calpain concentration is elevated although net calcium-dependent proteolysis is suppressed in dystrophin-deficient muscle. *Exp. Cell Res.* **203**, 107–14 (1992).
153. Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M. & Sweeney, H. L. Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3710–4 (1993).
154. Haycock, J. W., MacNeil, S., Jones, P., Harris, J. B. & Mantle, D. Oxidative damage to muscle protein in Duchenne muscular dystrophy. *Neuroreport* **8**, 357–61 (1996).
155. Vila, M. C. *et al.* Mitochondria mediate cell membrane repair and contribute to Duchenne muscular dystrophy. *Cell Death Differ.* **24**, 330–342 (2017).
156. Renjini, R., Gayathri, N., Nalini, A. & Srinivas Bharath, M. M. Oxidative Damage in Muscular Dystrophy Correlates with the Severity of the Pathology: Role of Glutathione Metabolism. *Neurochem. Res.* **37**, 885–898 (2012).
157. Plant, D. R. & Lynch, G. S. Depolarization-induced contraction and SR function in mechanically skinned muscle fibers from dystrophic mdx mice.
158. Bellinger, A. M., Mongillo, M. & Marks, A. R. Stressed out: the skeletal muscle ryanodine receptor as a target of stress. *J. Clin. Invest.* **118**, 445–453 (2008).
159. Gentil, C. *et al.* Variable phenotype of del45-55 Becker patients correlated with nNOS μ mislocalization and RYR1 hypernitrosylation. *Hum. Mol. Genet.* **21**, 3449–3460 (2012).
160. Froehner, S. C., Reed, S. M., Anderson, K. N., Huang, P. L. & Percival, J. M.

- Loss of nNOS inhibits compensatory muscle hypertrophy and exacerbates inflammation and eccentric contraction-induced damage in mdx mice. *Hum. Mol. Genet.* **24**, 492–505 (2015).
161. Busillo, J. M., Azzam, K. M. & Cidlowski, J. A. Glucocorticoids sensitize the innate immune system through regulation of the NLRP3 inflammasome. *J. Biol. Chem.* **286**, 38703–13 (2011).
 162. Mendell, J. R. *et al.* Randomized, double-blind six-month trial of prednisone in Duchenne's muscular dystrophy. *N. Engl. J. Med.* **320**, 1592–7 (1989).
 163. Ishikawa, Y. *et al.* Duchenne muscular dystrophy: survival by cardio-respiratory interventions. *Neuromuscul. Disord.* **21**, 47–51 (2011).
 164. Bushby, K., Muntoni, F., Urtizberea, A., Hughes, R. & Griggs, R. 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 (2004).
 165. Markham, L. W., Kinnett, K., Wong, B. L., Woodrow Benson, D. & Cripe, L. H. Corticosteroid treatment retards development of ventricular dysfunction in Duchenne muscular dystrophy. *Neuromuscul. Disord.* **18**, 365–70 (2008).
 166. Garrood, P., Eagle, M., Jardine, P. E., Bushby, K. & Straub, V. Myoglobinuria in boys with Duchenne muscular dystrophy on corticosteroid therapy. *Neuromuscul. Disord.* **18**, 71–73 (2008).
 167. Griggs, R. C. *et al.* Corticosteroids in Duchenne muscular dystrophy: Major variations in practice. *Muscle Nerve* **48**, 27–31 (2013).
 168. Griggs, R. C. *et al.* Efficacy and safety of deflazacort vs prednisone and placebo for Duchenne muscular dystrophy. *Neurology* **87**, 2123–2131 (2016).
 169. American Academy of Neurology., B. *et al.* *Neurology. Neurology* **90**, (Advanstar Communications, 2018).
 170. Wong, B. L. *et al.* Long-Term Outcome of Interdisciplinary Management of Patients with Duchenne Muscular Dystrophy Receiving Daily Glucocorticoid Treatment. *J. Pediatr.* **182**, 296–303.e1 (2017).
 171. Bogdanovich, S. *et al.* Functional improvement of dystrophic muscle by myostatin blockade. *Nature* **420**, 418–421 (2002).
 172. Tsuchida, K. Myostatin inhibition by a follistatin-derived peptide ameliorates the

- pathophysiology of muscular dystrophy model mice. *Acta Myol. myopathies cardiomyopathies Off. J. Mediterr. Soc. Myol.* **27**, 14–8 (2008).
173. Nakatani, M. *et al.* Transgenic expression of a myostatin inhibitor derived from follistatin increases skeletal muscle mass and ameliorates dystrophic pathology in *mdx* mice. *FASEB J.* **22**, 477–487 (2008).
 174. Mendell, J. R. *et al.* A Phase 1/2a Follistatin Gene Therapy Trial for Becker Muscular Dystrophy. *Mol. Ther.* **23**, 192–201 (2015).
 175. Wagner, K. R. *et al.* A phase I/II trial of MYO-029 in adult subjects with muscular dystrophy. *Ann. Neurol.* **63**, 561–571 (2008).
 176. Campbell, C. *et al.* Myostatin inhibitor ACE-031 treatment of ambulatory boys with Duchenne muscular dystrophy: Results of a randomized, placebo-controlled clinical trial. *Muscle Nerve* **55**, 458–464 (2017).
 177. Mariot, V. *et al.* Downregulation of myostatin pathway in neuromuscular diseases may explain challenges of anti-myostatin therapeutic approaches. *Nat. Commun.* **8**, 1859 (2017).
 178. Barton, E. R., Morris, L., Musaro, A., Rosenthal, N. & Sweeney, H. L. Muscle-specific expression of insulin-like growth factor I counters muscle decline in *mdx* mice. *J. Cell Biol.* **157**, 137–48 (2002).
 179. Shavlakadze, T., White, J., Hoh, J. F. Y., Rosenthal, N. & Grounds, M. D. Targeted expression of insulin-like growth factor-I reduces early myofiber necrosis in dystrophic *mdx* mice. *Mol. Ther.* **10**, 829–43 (2004).
 180. Gregorevic, P., Plant, D. R. & Lynch, G. S. Administration of insulin-like growth factor-I improves fatigue resistance of skeletal muscles from dystrophic *mdx* mice. *Muscle Nerve* **30**, 295–304 (2004).
 181. Schertzer, J. D., Gehrig, S. M., Ryall, J. G. & Lynch, G. S. Modulation of Insulin-like Growth Factor (IGF)-I and IGF-Binding Protein Interactions Enhances Skeletal Muscle Regeneration and Ameliorates the Dystrophic Pathology in *mdx* Mice. *Am. J. Pathol.* **171**, 1180–1188 (2007).
 182. Gehrig, S. M. *et al.* Therapeutic potential of PEGylated insulin-like growth factor I for skeletal muscle disease evaluated in two murine models of muscular dystrophy. *Growth Horm. IGF Res.* **22**, 69–75 (2012).
 183. Shi, S. *et al.* BMP antagonists enhance myogenic differentiation and ameliorate the dystrophic phenotype in a DMD mouse model. *Neurobiol. Dis.* **41**, 353–360 (2011).

184. Pozzoli, U. *et al.* Comparative analysis of the human dystrophin and utrophin gene structures. *Genetics* **160**, 793–8 (2002).
185. Lumeng, C. N. *et al.* Characterization of dystrophin and utrophin diversity in the mouse. *Hum. Mol. Genet.* **8**, 593–9 (1999).
186. Weir, A. P., Morgan, J. E. & Davies, K. E. A-utrophin up-regulation in mdx skeletal muscle is independent of regeneration. *Neuromuscul. Disord.* **14**, 19–23 (2004).
187. Deconinck, A. E. *et al.* Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell* **90**, 717–27 (1997).
188. Chakravarty, D., Chakraborti, S. & Chakrabarti, P. Flexibility in the N-terminal actin-binding domain: Clues from *in silico* mutations and molecular dynamics. *Proteins Struct. Funct. Bioinforma.* **83**, 696–710 (2015).
189. Li, D. *et al.* Sarcolemmal nNOS anchoring reveals a qualitative difference between dystrophin and utrophin. *J. Cell Sci.* **123**, 2008–2013 (2010).
190. Fisher, R. *et al.* Non-toxic ubiquitous over-expression of utrophin in the mdx mouse. *Neuromuscul. Disord.* **11**, 713–21 (2001).
191. Tinsley, J. M. *et al.* Amelioration of the dystrophic phenotype of mdx mice using a truncated utrophin transgene. *Nature* **384**, 349–353 (1996).
192. Tinsley, J. *et al.* Expression of full-length utrophin prevents muscular dystrophy in mdx mice. *Nat. Med.* **4**, 1441–1444 (1998).
193. Tivorsan Pharmaceuticals Granted FDA Orphan Drug Designation for Human Recombinant Biglycan in the Treatment of Duchenne Muscular Dystrophy | Business Wire. Available at: <https://www.businesswire.com/news/home/20160815005326/en/Tivorsan-Pharmaceuticals-Granted-FDA-Orphan-Drug-Designation>. (Accessed: 13th June 2018)
194. Squire, S. *et al.* Prevention of pathology in mdx mice by expression of utrophin: analysis using an inducible transgenic expression system. *Hum. Mol. Genet.* **11**, 3333–44 (2002).
195. Kennedy, T. L. *et al.* Utrophin influences mitochondrial pathology and oxidative stress in dystrophic muscle. *Skelet. Muscle* **7**, 22 (2017).
196. Péladeau, C., Adam, N. J. & Jasmin, B. J. Celecoxib treatment improves muscle function in mdx mice and increases utrophin A expression. *FASEB J.*

197. PoC Study to Assess Activity and Safety of SMT C1100 (Ezutromid) in Boys With DMD - Full Text View - ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT02858362?term=PhaseOut+DMD&rank=1>. (Accessed: 12th June 2018)
198. Thomas, G. D. Functional muscle ischemia in Duchenne and Becker muscular dystrophy. *Front. Physiol.* **4**, 381 (2013).
199. Wehling, M., Spencer, M. J. & Tidball, J. G. A nitric oxide synthase transgene ameliorates muscular dystrophy in mdx mice. *J. Cell Biol.* **155**, 123–132 (2001).
200. VOISIN, V. *et al.* -arginine improves dystrophic phenotype in mice. *Neurobiol. Dis.* **20**, 123–130 (2005).
201. Sciorati, C. *et al.* A dual acting compound releasing nitric oxide (NO) and ibuprofen, NCX 320, shows significant therapeutic effects in a mouse model of muscular dystrophy. *Pharmacol. Res.* **64**, 210–217 (2011).
202. Adamo, C. M. *et al.* Sildenafil reverses cardiac dysfunction in the mdx mouse model of Duchenne muscular dystrophy. *Proc. Natl. Acad. Sci.* **107**, 19079–19083 (2010).
203. Nelson, M. D. *et al.* PDE5 inhibition alleviates functional muscle ischemia in boys with Duchenne muscular dystrophy. *Neurology* **82**, 2085–2091 (2014).
204. Leung, D. G. *et al.* Sildenafil does not improve cardiomyopathy in Duchenne/Becker muscular dystrophy. *Ann. Neurol.* **76**, 541–549 (2014).
205. Hodges, B. L. *et al.* Altered expression of the $\alpha 7 \beta 1$ integrin in human and murine muscular dystrophies. *J. Cell Sci.* **110 (Pt 22)**, 2873–81 (1997).
206. Guo, C. *et al.* Absence of $\alpha 7$ integrin in dystrophin-deficient mice causes a myopathy similar to Duchenne muscular dystrophy. *Hum. Mol. Genet.* **15**, 989–998 (2006).
207. Sunada, Y., Edgar, T. S., Lotz, B. P., Rust, R. S. & Campbell, K. P. Merosin-negative congenital muscular dystrophy associated with extensive brain abnormalities. *Neurology* **45**, 2084–9 (1995).
208. Liu, J., Burkin, D. J. & Kaufman, S. J. Increasing $\alpha 7 \beta 1$ -integrin promotes muscle cell proliferation, adhesion, and resistance to apoptosis without changing gene expression. *Am. J. Physiol. Cell Physiol.* **294**, C627–40 (2008).

209. Burkin, D. J. *et al.* Transgenic expression of $\alpha 7 \beta 1$ integrin maintains muscle integrity, increases regenerative capacity, promotes hypertrophy, and reduces cardiomyopathy in dystrophic mice. *Am. J. Pathol.* **166**, 253–63 (2005).
210. Burkin, D. J., Wallace, G. Q., Nicol, K. J., Kaufman, D. J. & Kaufman, S. J. Enhanced expression of the $\alpha 7 \beta 1$ integrin reduces muscular dystrophy and restores viability in dystrophic mice. *J. Cell Biol.* **152**, 1207–18 (2001).
211. Rooney, J. E., Gurpur, P. B., Yablonka-Reuveni, Z. & Burkin, D. J. Laminin-111 restores regenerative capacity in a mouse model for $\alpha 7$ integrin congenital myopathy. *Am. J. Pathol.* **174**, 256–64 (2009).
212. Rooney, J. E., Knapp, J. R., Hodges, B. L., Wuebbles, R. D. & Burkin, D. J. Laminin-111 protein therapy reduces muscle pathology and improves viability of a mouse model of merosin-deficient congenital muscular dystrophy. *Am. J. Pathol.* **180**, 1593–602 (2012).
213. Gawlik, K. I., Oliveira, B. M. & Durbeej, M. Transgenic Expression of Laminin $\alpha 1$ Chain Does Not Prevent Muscle Disease in the mdx Mouse Model for Duchenne Muscular Dystrophy. *Am. J. Pathol.* **178**, 1728–1737 (2011).
214. PALMER, E., WILHELM, J. M. & SHERMAN, F. Phenotypic suppression of nonsense mutants in yeast by aminoglycoside antibiotics. *Nature* **277**, 148–150 (1979).
215. SINGH, A., URSIC, D. & DAVIES, J. Phenotypic suppression and misreading in *Saccharomyces cerevisiae*. *Nature* **277**, 146–148 (1979).
216. Barton-Davis, E. R., Cordier, L., Shoturma, D. I., Leland, S. E. & Sweeney, H. L. Aminoglycoside antibiotics restore dystrophin function to skeletal muscles of mdx mice. *J. Clin. Invest.* **104**, 375–81 (1999).
217. Sicinski, P. *et al.* The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science* **244**, 1578–80 (1989).
218. Malik, V. *et al.* Gentamicin-induced readthrough of stop codons in Duchenne muscular dystrophy. *Ann. Neurol.* **67**, NA-NA (2010).
219. Tablan, O. C., Reyes, M. P., Rintelmann, W. F. & Lerner, A. M. Renal and auditory toxicity of high-dose, prolonged therapy with gentamicin and tobramycin in pseudomonas endocarditis. *J. Infect. Dis.* **149**, 257–63 (1984).
220. McDonald, C. M. *et al.* Ataluren in patients with nonsense mutation Duchenne

- muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **390**, 1489–1498 (2017).
221. Sazani, P. & Kole, R. Therapeutic potential of antisense oligonucleotides as modulators of alternative splicing. *Journal of Clinical Investigation* **112**, 481–486 (2003).
222. Rodino-Klapac, L. R., Mendell, J. R. & Sahenk, Z. Update on the treatment of Duchenne muscular dystrophy. *Current neurology and neuroscience reports* **13**, 332 (2013).
223. Duncley, M. G., Manoharan, M., Villiet, P., Eperon, I. C. & Dickson, G. Modification of splicing in the dystrophin gene in cultured Mdx muscle cells by antisense oligoribonucleotides. *Hum. Mol. Genet.* **7**, 1083–1090 (1998).
224. Lu, Q. L. *et al.* Functional amounts of dystrophin produced by skipping the mutated exon in the mdx dystrophic mouse. *Nat. Med.* **9**, 1009–14 (2003).
225. Goyenvalle, A. *et al.* Prevention of dystrophic pathology in severely affected dystrophin/utrophin-deficient mice by morpholino-oligomer-mediated exon-skipping. *Mol. Ther.* **18**, 198–205 (2010).
226. Malerba, A. *et al.* Chronic systemic therapy with low-dose morpholino oligomers ameliorates the pathology and normalizes locomotor behavior in mdx mice. *Mol. Ther.* **19**, 345–354 (2011).
227. Aoki, Y. *et al.* Bodywide skipping of exons 45-55 in dystrophic mdx52 mice by systemic antisense delivery. *Proceedings of the National Academy of Sciences* **109**, 13763–13768 (2012).
228. Yokota, T. *et al.* Extensive and prolonged restoration of dystrophin expression with vivo-morpholino-mediated multiple exon skipping in dystrophic dogs. *Nucleic Acid Ther.* **22**, 306–15 (2012).
229. Aartsma-Rus, A. *et al.* Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations. *Hum. Mutat.* **30**, 293–299 (2009).
230. van Deutekom, J. C. *et al.* Local dystrophin restoration with antisense oligonucleotide PRO051. *N. Engl. J. Med.* **357**, 2677–86 (2007).
231. Kinali, M. *et al.* Local restoration of dystrophin expression with the morpholino oligomer AVI-4658 in Duchenne muscular dystrophy: a single-blind, placebo-controlled, dose-escalation, proof-of-concept study. *Lancet. Neurol.* **8**, 918–28 (2009).

232. Aartsma-Rus, A. & Krieg, A. M. FDA Approves Eteplirsen for Duchenne Muscular Dystrophy: The Next Chapter in the Eteplirsen Saga. *Nucleic Acid Ther.* **27**, 1–3 (2017).
233. Komaki, H. *et al.* Exon 53 skipping of the dystrophin gene in patients with Duchenne muscular dystrophy by systemic administration of NS-065/NCNP-01: A phase 1, dose escalation, first-in-human study. *Neuromuscul. Disord.* **25**, S261–S262 (2015).
234. American Academy of Neurology., F. *et al.* *Neurology*. *Neurology* **90**, (Advanstar Communications, 2018).
235. An Extension Study to Evaluate Casimersen or Golodirsen in Patients With Duchenne Muscular Dystrophy - Full Text View - ClinicalTrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT03532542>. (Accessed: 18th June 2018)
236. Aartsma-Rus, A. & van Ommen, G.-J. B. Less is more: therapeutic exon skipping for Duchenne muscular dystrophy. *Lancet. Neurol.* **8**, 873–5 (2009).
237. Partridge, T. A., Morgan, J. E., Coulton, G. R., Hoffman, E. P. & Kunkel, L. M. Conversion of mdx myofibres from dystrophin-negative to -positive by injection of normal myoblasts. *Nature* **337**, 176–179 (1989).
238. Abujarour, R. *et al.* Myogenic Differentiation of Muscular Dystrophy-Specific Induced Pluripotent Stem Cells for Use in Drug Discovery. *Stem Cells Transl. Med.* **3**, 149–160 (2014).
239. Skuk, D. *et al.* Dystrophin expression in myofibers of Duchenne muscular dystrophy patients following intramuscular injections of normal myogenic cells. *Mol. Ther.* **9**, 475–82 (2004).
240. Fan, Y., Maley, M., Beilharz, M. & Grounds, M. Rapid death of injected myoblasts in myoblast transfer therapy. *Muscle Nerve* **19**, 853–60 (1996).
241. Dellavalle, A. *et al.* Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. *Nat. Cell Biol.* **9**, 255–267 (2007).
242. Meregalli, M. *et al.* Perspectives of stem cell therapy in Duchenne muscular dystrophy. *FEBS J.* **280**, 4251–4262 (2013).
243. Montarras, D. *et al.* Direct Isolation of Satellite Cells for Skeletal Muscle Regeneration. *Science (80-.)*. **309**, (2005).
244. Sacco, A., Doyonnas, R., Kraft, P., Vitorovic, S. & Blau, H. M. Self-renewal and

- expansion of single transplanted muscle stem cells. *Nature* **456**, 502–506 (2008).
245. Cerletti, M. *et al.* Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* **134**, 37–47 (2008).
 246. Pessina, P. *et al.* Stem Cell Reports Fibrogenic Cell Plasticity Blunts Tissue Regeneration and Aggravates Muscular Dystrophy. *Stem Cell Reports* **4**, 1046–1060 (2015).
 247. Minasi, M. G. *et al.* The meso-angioblast: a multipotent, self-renewing cell that originates from the dorsal aorta and differentiates into most mesodermal tissues. *Development* **129**, 2773–83 (2002).
 248. Galvez, B. G. *et al.* Complete repair of dystrophic skeletal muscle by mesoangioblasts with enhanced migration ability. *J. Cell Biol.* **174**, 231–43 (2006).
 249. Sampaolesi, M. *et al.* Mesoangioblast stem cells ameliorate muscle function in dystrophic dogs. *Nature* **444**, 574–579 (2006).
 250. Cossu, G. *et al.* Intra-arterial transplantation of HLA-matched donor mesoangioblasts in Duchenne muscular dystrophy. *EMBO Mol. Med.* **7**, 1513–1528 (2015).
 251. Ohkawa, S. *et al.* Magnetic targeting of human peripheral blood CD133+ cells for skeletal muscle regeneration. *Tissue Eng. Part C. Methods* **19**, 631–41 (2013).
 252. Meng, J. *et al.* Human Skeletal Muscle-derived CD133+ Cells Form Functional Satellite Cells After Intramuscular Transplantation in Immunodeficient Host Mice. *Mol. Ther.* **22**, 1008–1017 (2014).
 253. Torrente, Y. *et al.* Human circulating AC133+ stem cells restore dystrophin expression and ameliorate function in dystrophic skeletal muscle. *J. Clin. Invest.* **114**, 182–195 (2004).
 254. Negroni, E. *et al.* In vivo myogenic potential of human CD133+ muscle-derived stem cells: a quantitative study. *Mol. Ther.* **17**, 1771–8 (2009).
 255. Torrente, Y. *et al.* Autologous transplantation of muscle-derived CD133+ stem cells in Duchenne muscle patients. *Cell Transplant.* **16**, 563–77 (2007).
 256. Meng, J., Muntoni, F. & Morgan, J. CD133+ cells derived from skeletal muscles of Duchenne muscular dystrophy patients have a compromised myogenic and

- muscle regenerative capability. *Stem Cell Res.* **30**, 43–52 (2018).
257. Finer, M. & Glorioso, J. A brief account of viral vectors and their promise for gene therapy. *Gene Ther.* **24**, 1–2 (2017).
 258. Ahi, Y. S. & Mittal, S. K. Components of Adenovirus Genome Packaging. *Front. Microbiol.* **7**, 1503 (2016).
 259. Goosney, D. L. & Nemerow, G. R. Adenovirus Infection: Taking the Back Roads to Viral Entry. *Curr. Biol.* **13**, R99–R100 (2003).
 260. Ehrhardt, A. *et al.* A gene-deleted adenoviral vector results in phenotypic correction of canine hemophilia B without liver toxicity or thrombocytopenia. *Blood* **102**, 2403–2411 (2003).
 261. Van Deutekom, J. C. T. *et al.* *Extended Tropism of an Adenoviral Vector does not Circumvent the Maturation-dependent Transducibility of Mouse Skeletal Muscle.*
 262. Nalbantoglu, J., Pari, G., Karpati, G. & Holland, P. C. Expression of the Primary Coxsackie and Adenovirus Receptor Is Downregulated during Skeletal Muscle Maturation and Limits the Efficacy of Adenovirus-Mediated Gene Delivery to Muscle Cells. *Hum. Gene Ther.* **10**, 1009–1019 (1999).
 263. Muruve, D. A. The Innate Immune Response to Adenovirus Vectors. *Hum. Gene Ther.* **15**, 1157–1166 (2004).
 264. Connolly, J. B. Lentiviruses in gene therapy clinical research. *Gene Ther.* **9**, 1730–1734 (2002).
 265. Roberts, R. G., Coffey, A. J., Bobrow, M. & Bentley, D. R. Exon Structure of the Human Dystrophin Gene. *Genomics* **16**, 536–538 (1993).
 266. Counsell, J. R. *et al.* Lentiviral vectors can be used for full-length dystrophin gene therapy. *Sci. Rep.* **7**, 79 (2017).
 267. Peng, K.-W. *et al.* Organ distribution of gene expression after intravenous infusion of targeted and untargeted lentiviral vectors. *Gene Ther.* **8**, 1456–1463 (2001).
 268. Li, S. *et al.* Stable transduction of myogenic cells with lentiviral vectors expressing a minidystrophin. *Gene Ther.* **12**, 1099–1108 (2005).
 269. MacKenzie, T. C. *et al.* Transduction of satellite cells after prenatal intramuscular administration of lentiviral vectors. *J. Gene Med.* **7**, 50–58 (2005).

270. Talbot, G. E., Waddington, S. N., Bales, O., Tchen, R. C. & Antoniou, M. N. Desmin-regulated Lentiviral Vectors for Skeletal Muscle Gene Transfer. *Mol. Ther.* **18**, 601–608 (2010).
271. Kobinger, G. P., Louboutin, J.-P., Barton, E. R., Sweeney, H. L. & Wilson, J. M. Correction of the Dystrophic Phenotype by *In Vivo* Targeting of Muscle Progenitor Cells. *Hum. Gene Ther.* **14**, 1441–1449 (2003).
272. Howe, S. J. *et al.* Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J. Clin. Invest.* **118**, 3143–50 (2008).
273. Gregorevic, P. *et al.* Systemic delivery of genes to striated muscles using adeno-associated viral vectors. *Nat. Med.* **10**, 828–834 (2004).
274. Rivera, V. M. *et al.* Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer. *Blood* **105**, 1424–1430 (2005).
275. Zincarelli, C., Soltys, S., Rengo, G. & Rabinowitz, J. E. Analysis of AAV Serotypes 1–9 Mediated Gene Expression and Tropism in Mice After Systemic Injection. *Mol. Ther.* **16**, 1073–1080 (2008).
276. Zincarelli, C., Soltys, S., Rengo, G., Koch, W. J. & Rabinowitz, J. E. 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 (2010).
277. England, S. B. *et al.* Very mild muscular dystrophy associated with the deletion of 46% of dystrophin. *Nature* **343**, 180–182 (1990).
278. Watchko, J. *et al.* Adeno-associated virus vector-mediated minidystrophin gene therapy improves dystrophic muscle contractile function in mdx mice. *Hum. Gene Ther.* **13**, 1451–60 (2002).
279. Henderson, D. M., Belanto, J. J., Li, B., Heun-Johnson, H. & Ervasti, J. M. Internal deletion compromises the stability of dystrophin. *Hum. Mol. Genet.* **20**, 2955–63 (2011).
280. Zhang, Y. & Duan, D. Novel mini-dystrophin gene dual adeno-associated virus vectors restore neuronal nitric oxide synthase expression at the sarcolemma. *Hum. Gene Ther.* **23**, 98–103 (2012).
281. A Study to Evaluate the Safety and Tolerability of PF-06939926 Gene Therapy in Duchenne Muscular Dystrophy - Full Text View - ClinicalTrials.gov. Available

at: <https://clinicaltrials.gov/ct2/show/NCT03362502>. (Accessed: 8th April 2019)

282. Odom, G. L., Gregorevic, P., Allen, J. M. & Chamberlain, J. S. Gene Therapy of mdx Mice With Large Truncated Dystrophins Generated by Recombination Using rAAV6. *Mol. Ther.* **19**, 36–45 (2011).
283. Le Guiner, C. *et al.* Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat. Commun.* **8**, 16105 (2017).
284. Hakim, C. H. *et al.* A Five-Repeat Micro-Dystrophin Gene Ameliorated Dystrophic Phenotype in the Severe DBA/2J-mdx Model of Duchenne Muscular Dystrophy. *Mol. Ther. Methods Clin. Dev.* **6**, 216–230 (2017).
285. A Randomized, Double-blind, Placebo-controlled Study of SRP-9001 for Duchenne Muscular Dystrophy (DMD) - Full Text View - ClinicalTrials.gov. Available at:
<https://clinicaltrials.gov/ct2/show/NCT03769116?term=microdystrophin&rank=2>. (Accessed: 8th April 2019)
286. Microdystrophin Gene Transfer Study in Adolescents and Children With DMD - Full Text View - ClinicalTrials.gov. Available at:
<https://clinicaltrials.gov/ct2/show/NCT03368742?term=microdystrophin&rank=1>. (Accessed: 8th April 2019)
287. Vilenchik, M. M. & Knudson, A. G. Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12871–6 (2003).
288. Wyman, C. & Kanaar, R. DNA double-strand break repair: all's well that ends well. *Annu. Rev. Genet.* **40**, 363–83 (2006).
289. Stracker, T. H. & Petrini, J. H. J. The MRE11 complex: starting from the ends. *Nat. Rev. Mol. Cell Biol.* **12**, 90–103 (2011).
290. Limbo, O. *et al.* Ctp1 Is a Cell-Cycle-Regulated Protein that Functions with Mre11 Complex to Control Double-Strand Break Repair by Homologous Recombination. *Mol. Cell* **28**, 134–146 (2007).
291. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
292. Shivji, M. K. K. *et al.* The BRC repeats of human BRCA2 differentially regulate RAD51 binding on single- versus double-stranded DNA to stimulate strand exchange. *Proc. Natl. Acad. Sci.* **106**, 13254–13259 (2009).

293. Li, M. *et al.* The catalytic subunit of DNA-dependent protein kinase is required for cellular resistance to oxidative stress independent of DNA double-strand break repair. *Free Radic. Biol. Med.* **76**, 278–285 (2014).
294. Gu, J. *et al.* DNA-PKcs regulates a single-stranded DNA endonuclease activity of Artemis. *DNA Repair (Amst)*. **9**, 429–37 (2010).
295. Roy, S. *et al.* XRCC4/XLF Interaction Is Variably Required for DNA Repair and Is Not Required for Ligase IV Stimulation. *Mol. Cell. Biol.* **35**, 3017 LP-3028 (2015).
296. Williams, G. J. *et al.* Structural insights into NHEJ: Building up an integrated picture of the dynamic DSB repair super complex, one component and interaction at a time. *DNA Repair (Amst)*. **17**, 110–120 (2014).
297. Radhakrishnan, S. K. & Lees-Miller, S. P. DNA requirements for interaction of the C-terminal region of Ku80 with the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). *DNA Repair (Amst)*. **57**, 17–28 (2017).
298. Li, G. *et al.* Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts. *Int. J. Biochem. Cell Biol.* **99**, 154–160 (2018).
299. Pawelczak, K. S., Gavande, N. S., VanderVere-Carozza, P. S. & Turchi, J. J. Modulating DNA Repair Pathways to Improve Precision Genome Engineering. *ACS Chem. Biol.* **13**, 389–396 (2018).
300. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
301. Belfort, M. & Roberts, R. J. Homing endonucleases: keeping the house in order. *Nucleic Acids Res.* **25**, 3379–88 (1997).
302. Jurica, M. S. & Stoddard, B. L. Homing endonucleases: structure, function and evolution. *Cell. Mol. Life Sci.* **55**, 1304–1326 (1999).
303. Silva, G. *et al.* Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr. Gene Ther.* **11**, 11–27 (2011).
304. Heath, P. J., Stephens, K. M., Monnat, R. J. & Stoddard, B. L. The structure of I-Crel, a group I intron-encoded homing endonuclease. *Nat. Struct. Biol.* **4**, 468–76 (1997).

305. Duan, X., Gimble, F. S. & Quijcho, F. A. Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. *Cell* **89**, 555–64 (1997).
306. Thierry, A. & Dujon, B. Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes. *Nucleic Acids Res.* **20**, 5625–31 (1992).
307. Chevalier, B. S. & Stoddard, B. L. Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res.* **29**, 3757–74 (2001).
308. Pâques, F. & Duchateau, P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr. Gene Ther.* **7**, 49–66 (2007).
309. Chevalier, B. S. *et al.* Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell* **10**, 895–905 (2002).
310. Seligman, L. M. *et al.* Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res.* **30**, 3870–9 (2002).
311. Rosen, L. E. *et al.* Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.* **34**, 4791–800 (2006).
312. Smith, J. *et al.* A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.* **34**, e149 (2006).
313. Muñoz, I. G. *et al.* Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus. *Nucleic Acids Res.* **39**, 729–43 (2011).
314. Rousseau, J. *et al.* Endonucleases: Tools to correct the dystrophin gene. *J. Gene Med.* **13**, 522–537 (2011).
315. Chapdelaine, P., Pichavant, C., Rousseau, J., Pâques, F. & Tremblay, J. P. Meganucleases can restore the reading frame of a mutated dystrophin. *Gene Ther.* **17**, 846–58 (2010).
316. Li, L., Wu, L. P. & Chandrasegaran, S. Functional domains in Fok I restriction endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 4275–9 (1992).
317. Kim, Y. G. & Chandrasegaran, S. Chimeric restriction endonuclease. *Proc. Natl. Acad. Sci. U. S. A.* **91**, 883–7 (1994).
318. Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. U. S. A.* **93**,

1156–60 (1996).

319. Kim, Y. G., Smith, J., Durgesha, M. & Chandrasegaran, S. Chimeric restriction enzyme: Gal4 fusion to FokI cleavage domain. *Biol. Chem.* **379**, 489–95
320. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809–17 (1991).
321. Porteus, M. Creating zinc finger nucleases to manipulate the genome in a site-specific manner using a modular-assembly approach. *Cold Spring Harb. Protoc.* **2010**, pdb.top93 (2010).
322. Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 10570–5 (1998).
323. Mani, M., Smith, J., Kandavelou, K., Berg, J. M. & Chandrasegaran, S. Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage. *Biochem. Biophys. Res. Commun.* **334**, 1191–1197 (2005).
324. Beumer, K. J. *et al.* Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 19821–19826 (2008).
325. Santiago, Y. *et al.* Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 5809–5814 (2008).
326. Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4⁺ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* **26**, 808–816 (2008).
327. Foley, J. E. *et al.* Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. *Nat. Protoc.* **4**, 1855–1867 (2009).
328. Geurts, A. M. *et al.* Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* **325**, 433 (2009).
329. Cost, G. J. *et al.* BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol. Bioeng.* **105**, 330–340 (2010).
330. Ran, Y. *et al.* Zinc finger nuclease-mediated precision genome editing of an endogenous gene in hexaploid bread wheat (*Triticum aestivum*) using a DNA repair template. *Plant Biotechnol. J.* (2018). doi:10.1111/pbi.12941
331. Laoharawee, K. *et al.* Dose-Dependent Prevention of Metabolic and Neurologic

- Disease in Murine MPS II by ZFN-Mediated In Vivo Genome Editing. *Mol. Ther.* **26**, 1127–1136 (2018).
332. Ramirez, C. L. *et al.* Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods* **5**, 374–5 (2008).
 333. Ousterout, D. G. *et al.* 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 (2015).
 334. XU, Z. *et al.* Action modes of transcription activator-like effectors (TALEs) of *Xanthomonas* in plants. *J. Integr. Agric.* **16**, 2736–2745 (2017).
 335. Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509–12 (2009).
 336. Mussolino, C. & Cathomen, T. TALE nucleases: Tailored genome engineering made easy. *Current Opinion in Biotechnology* **23**, 644–650 (2012).
 337. Deng, D. *et al.* Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. *Science* **335**, 720–723 (2012).
 338. Gupta, R. M. & Musunuru, K. Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Invest.* **124**, 4154–61 (2014).
 339. Cermak, T., Doyle, E. & Christian, M. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic acids ...* **39**, e82 (2011).
 340. Ousterout, D. G. *et al.* Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol. Ther.* **21**, 1718–26 (2013).
 341. Holkers, M. *et al.* Differential integrity of TALE nuclease genes following adenoviral and lentiviral vector gene transfer into human cells. *Nucleic Acids Res.* **41**, (2013).
 342. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169**, 5429–33 (1987).
 343. Nakata, A., Amemura, M. & Makino, K. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *J. Bacteriol.* **171**, 3553–6 (1989).

344. van Soolingen, D., de Haas, P. E., Hermans, P. W., Groenen, P. M. & van Embden, J. D. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J. Clin. Microbiol.* **31**, 1987–95 (1993).
345. Groenen, P. M., Bunschoten, A. E., van Soolingen, D. & van Embden, J. D. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* **10**, 1057–65 (1993).
346. Mojica, F. J., Juez, G. & Rodríguez-Valera, F. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Mol. Microbiol.* **9**, 613–21 (1993).
347. Mojica, F. J., Ferrer, C., Juez, G. & Rodríguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Mol. Microbiol.* **17**, 85–93 (1995).
348. Mojica, F. J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36**, 244–6 (2000).
349. Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of a novel family of sequence repeats among prokaryotes. *OMICS* **6**, 23–33 (2002).
350. Jansen, R., Van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43**, 1565–1575 (2002).
351. Pourcel, C., Salvignol, G. & Vergnaud, G. 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 (2005).
352. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–82 (2005).
353. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–61 (2005).
354. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in

- prokaryotes. *Science* **315**, 1709–12 (2007).
355. Marraffini, L. A. & Sontheimer, E. J. CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science* (80-.). **322**, 1843–1845 (2008).
 356. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–4 (2008).
 357. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
 358. Shah, S. A., Erdmann, S., Mojica, F. J. M. & Garrett, R. A. Protospacer recognition motifs: mixed identities and functional diversity. *RNA Biol.* **10**, 891–9 (2013).
 359. Cencic, R. *et al.* Protospacer Adjacent Motif (PAM)-Distal Sequences Engage CRISPR Cas9 DNA Target Cleavage. *PLoS One* **9**, e109213 (2014).
 360. Pul, Ü. *et al.* Identification and characterization of *E. coli* CRISPR- *cas* promoters and their silencing by H-NS. *Mol. Microbiol.* **75**, 1495–1512 (2010).
 361. Makarova, K. S. *et al.* An updated evolutionary classification of CRISPR–Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
 362. Koonin, E. V, Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
 363. Cooper, L. A., Stringer, A. M. & Wade, J. T. Determining the Specificity of Cascade Binding, Interference, and Primed Adaptation In Vivo in the Escherichia coli Type I-E CRISPR-Cas System. *MBio* **9**, e02100-17 (2018).
 364. Chylinski, K., Le Rhun, A. & Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* **10**, 726–37 (2013).
 365. Nuñez, J. K., Lee, A. S. Y., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature* **519**, 193–198 (2015).
 366. Koo, Y., Jung, D. & Bae, E. Crystal Structure of Streptococcus pyogenes Csn2 Reveals Calcium-Dependent Conformational Changes in Its Tertiary and Quaternary Structure. *PLoS One* **7**, e33401 (2012).
 367. Heler, R. *et al.* Cas9 specifies functional viral targets during CRISPR–Cas adaptation. *Nature* **519**, 199–202 (2015).

368. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602–607 (2011).
369. Nicholson, A. W. Ribonuclease III mechanisms of double-stranded RNA cleavage. *Wiley Interdiscip. Rev. RNA* **5**, 31–48 (2014).
370. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
371. Szczelkun, M. D. *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 9798–803 (2014).
372. Gong, S., Yu, H. H., Johnson, K. A., Correspondence, D. W. T. & Taylor, D. W. DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. *CellReports* **22**, 359–371 (2018).
373. Jiang, F. *et al.* Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science (80-.).* **351**, 867–871 (2016).
374. Yang, M. *et al.* The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET. (2018).
doi:10.1016/j.celrep.2017.12.048
375. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–21 (2012).
376. Nishimasu, H. *et al.* Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell* **156**, 935–949 (2014).
377. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science (80-.).* **343**, 1247997–1247997 (2014).
378. Palermo, G. *et al.* Key role of the REC lobe during CRISPR–Cas9 activation by ‘sensing’, ‘regulating’, and ‘locking’ the catalytic HNH domain. *Q. Rev. Biophys.* **51**, e9 (2018).
379. Zuo, Z. & Liu, J. Structure and Dynamics of Cas9 HNH Domain Catalytic State. *Sci. Rep.* **7**, 17271 (2017).
380. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–6 (2013).
381. Wiles, M. V., Qin, W., Cheng, A. W. & Wang, H. CRISPR–Cas9-mediated genome editing and guide RNA design. *Mamm. Genome* **26**, 501–510 (2015).

382. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23 (2013).
383. Zhang, H., Cheng, Q.-X., Liu, A.-M., Zhao, G.-P. & Wang, J. A Novel and Efficient Method for Bacteria Genome Editing Employing both CRISPR/Cas9 and an Antibiotic Resistance Cassette. *Front. Microbiol.* **8**, 812 (2017).
384. Liu, X., Wu, S., Xu, J., Sui, C. & Wei, J. Application of CRISPR/Cas9 in plant biology. *Acta Pharm. Sin. B* **7**, 292–302 (2017).
385. Jao, L.-E., Wente, S. R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13904–9 (2013).
386. Li, D. *et al.* Heritable gene targeting in the mouse and rat using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 681–3 (2013).
387. Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407–11 (2016).
388. Waddington, S. N., Privolizzi, R., Karda, R. & O'Neill, H. C. A Broad Overview and Review of CRISPR-Cas Technology and Stem Cells. *Curr. stem cell reports* **2**, 9–20 (2016).
389. Wojtal, D. *et al.* Spell Checking Nature: Versatility of CRISPR/Cas9 for Developing Treatments for Inherited Disorders. *Am. J. Hum. Genet.* **98**, 90–101 (2016).
390. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23 (2013).
391. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).
392. Nakamura, K. *et al.* Generation of muscular dystrophy model rats with a CRISPR/Cas system. *Sci. Rep.* **4**, 5635 (2014).
393. Yu, H.-H. *et al.* Porcine Zygote Injection with Cas9/sgRNA Results in DMD-Modified Pig with Muscle Dystrophy. *Int. J. Mol. Sci.* **17**, (2016).
394. Hruscha, A. & Schmid, B. in *Methods in molecular biology (Clifton, N.J.)* **1254**, 341–350 (2015).
395. Birling, M.-C., Herault, Y. & Pavlovic, G. Modeling human disease in rodents by CRISPR/Cas9 genome editing. *Mamm. Genome* **28**, 291–301 (2017).

396. Kampmann, M. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem. Biol.* **13**, 406–416 (2018).
397. Hilton, I. B. *et al.* Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).
398. Ma, Y. *et al.* Live Visualization of HIV-1 Proviral DNA Using a Dual-Color-Labeled CRISPR System. *Anal. Chem.* **89**, 12896–12901 (2017).
399. Al-Allaf, F. A., Tolmachov, O. E., Zambetti, L. P., Tchetchelnitski, V. & Mehmet, H. Remarkable stability of an instability-prone lentiviral vector plasmid in *Escherichia coli* Stbl3. *3 Biotech* **3**, 61–70 (2013).
400. Froger, A. & Hall, J. E. Transformation of plasmid DNA into *E. coli* using the heat shock method. *J. Vis. Exp.* 253 (2007). doi:10.3791/253
401. Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87 (2014).
402. Heckl, D. *et al.* Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat. Biotechnol.* **32**, 941–946 (2014).
403. Sancak, Y. *et al.* The Rag GTPases Bind Raptor and Mediate Amino Acid Signaling to mTORC1. *Science* (80-.). **320**, 1496–1501 (2008).
404. Stewart, S. A. *et al.* Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501 (2003).
405. Matsuda, T. & Cepko, C. L. Electroporation and RNA interference in the rodent retina in vivo and in vitro. *Proc. Natl. Acad. Sci.* **101**, 16–22 (2004).
406. Tomás, H. A., Rodrigues, A. F., Alves, P. M. & Coroadinha, A. S. Lentiviral Gene Therapy Vectors : Challenges and Future Directions. (2011).
407. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
408. Kotsopoulou, E., Kim, V. N., Kingsman, A. J., Kingsman, S. M. & Mitrophanous, K. A. A Rev-Independent Human Immunodeficiency Virus Type 1 (HIV-1)-Based Vector That Exploits a Codon-Optimized HIV-1 gag-pol Gene. *JOURNAL OF VIROLOGY* **74**, (2000).
409. Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–7 (1996).

410. Barczak, W., Suchorska, W., Rubiś, B. & Kulcenty, K. Universal real-time PCR-based assay for lentiviral titration. *Mol. Biotechnol.* **57**, 195–200 (2015).
411. Mamchaoui, K. *et al.* Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet. Muscle* **1**, 34 (2011).
412. Claassen, D. A., Desler, M. M. & Rizzino, A. ROCK inhibition enhances the recovery and growth of cryopreserved human embryonic stem cells and human induced pluripotent stem cells. *Mol. Reprod. Dev.* **76**, 722–32 (2009).
413. Korbie, D. J. & Mattick, J. S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat. Protoc.* **3**, 1452–6 (2008).
414. Lorenz, T. C. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. *J. Vis. Exp.* e3998 (2012). doi:10.3791/3998
415. Qiu, P. *et al.* Mutation detection using Surveyor nuclease. *Biotechniques* **36**, 702–7 (2004).
416. Weterings, E. & Chen, D. J. The endless tale of non-homologous end-joining. *Cell Res.* **18**, 114–24 (2008).
417. McColl, R., Nkosi, M., Snyman, C. & Niesler, C. Analysis and quantification of in vitro myoblast fusion using the LADD Multiple Stain. *Biotechniques* **61**, (2016).
418. Ohno, S. *Evolution by gene duplication*. (Springer-Verlag, 1970).
419. Franke, M. *et al.* Formation of new chromatin domains determines pathogenicity of genomic duplications. *Nature* **538**, 265–269 (2016).
420. Katju, V. & Bergthorsson, U. Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Front. Genet.* **4**, 273 (2013).
421. Levinson, G. & Gutman, G. A. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**, 203–21 (1987).
422. Lehrman, M. A., Goldstein, J. L., Russell, D. W. & Brown, M. S. Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia. *Cell* **48**, 827–35 (1987).
423. Ramsey, J. & Schemske, D. W. PATHWAYS, MECHANISMS, AND RATES OF POLYPLOID FORMATION IN FLOWERING PLANTS. *Annu. Rev. Ecol. Syst.* **29**, 467–501 (1998).

424. Bogart, J. P. & Bi, K. Genetic and Genomic Interactions of Animals with Different Ploidy Levels. *Cytogenet. Genome Res.* **140**, 117–136 (2013).
425. Makino, T. & McLysaght, A. Ohnologs in the human genome are dosage balanced and frequently associated with disease. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 9270–4 (2010).
426. Mamedov, I. Z., Lebedev, Y. B., Sverdlov, E. D. & Mamedov, Z. Unusually long target site duplications flanking some of the long terminal repeats of human endogenous retrovirus K in the human genome. (2018).
doi:10.1099/vir.0.19717-0
427. Kaessmann, H., Vinckenbosch, N. & Long, M. RNA-based gene duplication: mechanistic and evolutionary insights. *Nat. Rev. Genet.* **10**, 19–31 (2009).
428. Bertelson, C. J. *et al.* Localisation of Xp21 meiotic exchange points in Duchenne muscular dystrophy families Human DNA was isolated from whole blood 531. *Journal of Medical Genetics* **23**, (1986).
429. den Dunnen, J. T., Bakker, E., Breteler, E. G. K., Pearson, P. L. & van Ommen, G. J. B. Direct detection of more than 50% of the Duchenne muscular dystrophy mutations by field inversion gels. *Nature* **329**, 640–642 (1987).
430. Hu, X. Y. *et al.* Partial gene duplication in Duchenne and Becker muscular dystrophies. *J. Med. Genet.* **25**, 369–76 (1988).
431. van Ommen, G.-J. B. Frequency of new copy number variation in humans. *Nat. Genet.* **37**, 333–334 (2005).
432. Tuffery-Giraud, S. *et al.* Genotype-phenotype analysis in 2,405 patients with a dystrophinopathy using the UMD-DMD database: a model of nationwide knowledgebase. *Hum. Mutat.* **30**, 934–945 (2009).
433. Yang, J. *et al.* MLPA-based genotype–phenotype analysis in 1053 Chinese patients with DMD/BMD. *BMC Med. Genet.* **14**, 29 (2013).
434. Vengalil, S. *et al.* Duchenne Muscular Dystrophy and Becker Muscular Dystrophy Confirmed by Multiplex Ligation-Dependent Probe Amplification: Genotype-Phenotype Correlation in a Large Cohort. *J. Clin. Neurol.* **13**, 91 (2017).
435. Aartsma-Rus, A., Janson, A. A. M., van Ommen, G.-J. B. & van Deutekom, J. C. T. Antisense-induced exon skipping for duplications in Duchenne muscular dystrophy. *BMC Med. Genet.* **8**, 43 (2007).

436. Greer, K. L., Lochmüller, H., Flanigan, K., Fletcher, S. & Wilton, S. D. Targeted Exon Skipping to Correct Exon Duplications in the Dystrophin Gene. *Mol. Ther. - Nucleic Acids* **3**, e155 (2014).
437. Wein, N. *et al.* Efficient Skipping of Single Exon Duplications in DMD Patient-Derived Cell Lines Using an Antisense Oligonucleotide Approach. *J. Neuromuscul. Dis.* **4**, 199–207 (2017).
438. Forrest, S. *et al.* Personalized exon skipping strategies to address clustered non-deletion dystrophin mutations. (2010). doi:10.1016/j.nmd.2010.07.276
439. Hong, A. CRISPR in personalized medicine: Industry perspectives in gene editing. *Semin. Perinatol.* **42**, 501–507 (2018).
440. Juan-Mateu, J. *et al.* DMD Mutations in 576 Dystrophinopathy Families: A Step Forward in Genotype-Phenotype Correlations. *PLoS One* **10**, e0135189 (2015).
441. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
442. Thomas, P. & Smart, T. G. HEK293 cell line: A vehicle for the expression of recombinant proteins. *J. Pharmacol. Toxicol. Methods* **51**, 187–200 (2005).
443. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
444. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–32 (2013).
445. Lin, Y. *et al.* CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res.* **42**, 7473–85 (2014).
446. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–6 (2013).
447. Cui, Y., Xu, J., Cheng, M., Liao, X. & Peng, S. Review of CRISPR/Cas9 sgRNA Design Tools. *Interdiscip. Sci. Comput. Life Sci.* **10**, 455–465 (2018).
448. Wang, T., Wei, J. J., Sabatini, D. M. & Lander, E. S. Genetic Screens in Human Cells Using the CRISPR-Cas9 System. *Science (80-.).* **343**, 80–84 (2014).
449. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* (2014). doi:10.1038/nbt.3026

450. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
451. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
452. Kuan, P. F. *et al.* A systematic evaluation of nucleotide properties for CRISPR sgRNA design. *BMC Bioinformatics* **18**, 297 (2017).
453. Xu, H. *et al.* Sequence determinants of improved CRISPR sgRNA design. *Genome Res.* **25**, 1147–57 (2015).
454. Moreno-Mateos, M. A. *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**, 982–988 (2015).
455. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).
456. Perez, A. R. *et al.* GuideScan software for improved single and paired CRISPR guide RNA design. *Nat. Biotechnol.* **35**, 347–349 (2017).
457. Qiu, P. *et al.* Mutation detection using Surveyor nuclease. *Biotechniques* **36**, 702–7 (2004).
458. Vouillot, L., Thélie, A. & Pollet, N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)*. **5**, 407–15 (2015).
459. Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P. & Pruett-Miller, S. M. A Survey of Validation Strategies for CRISPR-Cas9 Editing OPEN. *Sci. RePoRTS* | **8**, 888 (2018).
460. Liesche, C. *et al.* Death receptor-based enrichment of Cas9-expressing cells. *BMC Biotechnol.* **16**, 17 (2016).
461. Vouillot, L., Thélie, A. & Pollet, N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. *G3 (Bethesda)*. **5**, 407–15 (2015).
462. Bell, C. C., Magor, G. W., Gillinder, K. R. & Perkins, A. C. A high-throughput screening strategy for detecting CRISPR-Cas9 induced mutations using next-generation sequencing. *BMC Genomics* **15**, 1002 (2014).
463. Vaheri, A., Kurkinen, M., Lehto, V. P., Linder, E. & Timpl, R. Codistribution of

- pericellular matrix proteins in cultured fibroblasts and loss in transformation: fibronectin and procollagen. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 4944–8 (1978).
464. Duncan, M. R. & Berman, B. Stimulation of Collagen and Glycosaminoglycan Production in Cultured Human Adult Dermal Fibroblasts by Recombinant Human Interleukin 6. *J. Invest. Dermatol.* **97**, 686–692 (1991).
 465. Story, M. T. Cultured human foreskin fibroblasts produce a factor that stimulates their growth with properties similar to basic fibroblast growth factor. *Vitr. Cell. Dev. Biol.* **25**, 402–408 (1989).
 466. Tanaka, H., Okada, T., Konishi, H. & Tsuji, T. The effect of reactive oxygen species on the biosynthesis of collagen and glycosaminoglycans in cultured human dermal fibroblasts. *Arch. Dermatol. Res.* **285**, 352–355 (1993).
 467. Miller, J. D. & Schlaeger, T. M. in *Methods in molecular biology (Clifton, N.J.)* **767**, 55–65 (2011).
 468. Hugnot, J. P. *et al.* Expression of the Dystrophin Gene in Cultured Fibroblasts. *Biochem. Biophys. Res. Commun.* **192**, 69–74 (1993).
 469. Wood, W. M., Etemad, S., Yamamoto, M. & Goldhamer, D. J. MyoD-expressing progenitors are essential for skeletal myogenesis and satellite cell development. *Dev. Biol.* **384**, 114–27 (2013).
 470. Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
 471. Davis, R. L., Cheng, P.-F., Lassar, A. B. & Weintraub, H. The MyoD DNA binding domain contains a recognition code for muscle-specific gene activation. *Cell* **60**, 733–746 (1990).
 472. Zammit, P. S. Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. *Semin. Cell Dev. Biol.* **72**, 19–32 (2017).
 473. Halevy, O. *et al.* Correlation of terminal cell cycle arrest of skeletal muscle with induction of p21 by MyoD. *Science* **267**, 1018–21 (1995).
 474. Crescenzi, M., Fleming, T. P., Lassar, A. B., Weintraub, H. & Aaronson, S. A. MyoD induces growth arrest independent of differentiation in normal and transformed cells. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8442–6 (1990).
 475. Sorrentino, V., Pepperkok, R., Davis, R. L., Ansorge, W. & Philipson, L. Cell proliferation inhibited by MyoD1 independently of myogenic differentiation.

Nature **345**, 813–815 (1990).

- 476. Kitzmann, M. *et al.* The muscle regulatory factors MyoD and myf-5 undergo distinct cell cycle-specific expression in muscle cells. *J. Cell Biol.* **142**, 1447–59 (1998).
- 477. Cao, Y. *et al.* Genome-wide MyoD Binding in Skeletal Muscle Cells: A Potential for Broad Cellular Reprogramming. *Dev. Cell* **18**, 662–674 (2010).
- 478. Shintaku, J. *et al.* MyoD Regulates Skeletal Muscle Oxidative Metabolism Cooperatively with Alternative NF- κ B. *Cell Rep.* **17**, 514–526 (2016).
- 479. Zhang, Z. *et al.* Tandem duplications of two separate fragments of the dystrophin gene in a patient with Duchenne muscular dystrophy. doi:10.1007/s10038-007-0235-1
- 480. Rezakhanlou, A. M. *et al.* Highly efficient stable expression of indoleamine 2,3 dioxygenase gene in primary fibroblasts. *Biol. Proced. Online* **12**, 107–112 (2010).
- 481. Pariente, N., Mao, S.-H., Morizono, K. & Chen, I. S. Y. Efficient targeted transduction of primary human endothelial cells with dual-targeted lentiviral vectors. (2007). doi:10.1002/jgm.1151
- 482. Dull, T. *et al.* A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72**, 8463–71 (1998).
- 483. Kleinman, H. K. & Martin, G. R. Matrigel: Basement membrane matrix with biological activity. *Semin. Cancer Biol.* **15**, 378–386 (2005).
- 484. Liu, Z. *et al.* Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci. Rep.* **7**, 2193 (2017).
- 485. Davis, H. E., Morgan, J. R. & Yarmush, M. L. Polybrene increases retrovirus gene transfer efficiency by enhancing receptor-independent virus adsorption on target cell membranes. *Biophys. Chem.* **97**, 159–72 (2002).
- 486. Han, S. Y. *et al.* Nucleofection is a highly effective gene transfer technique for human melanoma cell lines. doi:10.1111/j.1600-0625.2007.00687.x
- 487. Maurisse, R. *et al.* Comparative transfection of DNA into primary and transformed mammalian cells from different lineages. *BMC Biotechnol.* **10**, 9 (2010).
- 488. Gresch, O. & Altrogge, L. in *Methods in molecular biology (Clifton, N.J.)* **801**, 65–74 (2012).

489. Lesueur, L. L., Mir, L. M. & André, F. M. Overcoming the Specific Toxicity of Large Plasmids Electrotransfer in Primary Cells In Vitro. *Mol. Ther. - Nucleic Acids* **5**, e291 (2016).
490. Prella, A. *et al.* APPEARANCE AND LOCALIZATION OF DYSTROPHIN IN NORMAL HUMAN FETAL MUSCLE. *Int. J. Devl. Neuroscience* **9**, (1991).
491. Lattanzi, L., Mavilio, F. & Cossu, G. High efficiency myogenic conversion of human fibroblasts by adenoviral vector-mediated MyoD gene transfer. An alternative strategy for ex vivo gene therapy of primary myopathies. *J Clin Invest* **101**, (1998).
492. Fletcher, S. *et al.* Targeted exon skipping to address “leaky” mutations in the dystrophin gene. *Mol. Ther. Nucleic Acids* **1**, e48 (2012).
493. Hidaka, C. *et al.* CAR-dependent and CAR-independent pathways of adenovirus vector-mediated gene transfer and expression in human fibroblasts. *J. Clin. Invest.* **103**, 579–587 (1999).
494. Bakker, A. *et al.* A tropism-modified adenoviral vector increased the effectiveness of gene therapy for arthritis. *Gene Ther.* **8**, 1785–1793 (2001).
495. Kreiss, P. *et al.* Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Res.* **27**, 3792–3798 (1999).
496. Robbins, P. D. & Ghivizzani, S. C. Viral vectors for gene therapy. *Pharmacol. Ther.* **80**, 35–47 (1998).
497. Durand, S. & Cimorelli, A. The inside out of lentiviral vectors. *Viruses* **3**, 132–59 (2011).
498. Zhang, G. & Wang, T. Efficient lentiviral transduction of different human and mouse cells. *bioRxiv* 1–34 (2018). doi:10.1101/253732
499. Delrue, I., Pan, Q., Baczmanska, A. K., Callens, B. W. & Verdoodt, L. L. M. Determination of the Selection Capacity of Antibiotics for Gene Selection. *Biotechnol. J.* **13**, 1700747 (2018).
500. Geraerts, M., Willems, S., Baekelandt, V., Debyser, Z. & Gijssbers, R. Comparison of lentiviral vector titration methods. *BMC Biotechnol.* **6**, 34 (2006).
501. Distler, J. H. W. *et al.* Nucleofection: a new, highly efficient transfection method for primary human keratinocytes*. *Exp. Dermatol.* **14**, 315–320 (2005).
502. Horlbeck, M. A. *et al.* Nucleosomes impede Cas9 access to DNA in vivo and in

- vitro. *Elife* **5**, (2016).
503. Davie, J. R. Inhibition of Histone Deacetylase Activity by Butyrate. *J. Nutr.* **133**, 2485S–2493S (2003).
 504. Lattanzi, L. *et al.* High efficiency myogenic conversion of human fibroblasts by adenoviral vector-mediated MyoD gene transfer. An alternative strategy for ex vivo gene therapy of primary myopathies. *J. Clin. Invest.* **101**, 2119–2128 (1998).
 505. Fujii, I. *et al.* Adenoviral mediated MyoD gene transfer into fibroblasts: Myogenic disease diagnosis. *Brain Dev.* **28**, 420–425 (2006).
 506. Murry, C. E., Kay, M. A., Bartosek, T., Hauschka, S. D. & Schwartz, S. M. Muscle differentiation during repair of myocardial necrosis in rats via gene transfer with MyoD. *J. Clin. Invest.* **98**, 2209–17 (1996).
 507. Manandhar, D. *et al.* Incomplete MyoD-induced transdifferentiation is associated with chromatin remodeling deficiencies. *Nucleic Acids Res.* **45**, 11684–11699 (2017).
 508. Tapscott, S. J. The circuitry of a master switch: MyoD and the regulation of skeletal muscle gene transcription. *Development* **132**, 2685–2695 (2005).
 509. de la Serna, I. L. *et al.* MyoD Targets Chromatin Remodeling Complexes to the Myogenin Locus Prior to Forming a Stable DNA-Bound Complex. *Mol. Cell. Biol.* **25**, 3997–4009 (2005).
 510. Fernandez-Fuente, M. *et al.* Assessment of the transformation of equine skin–derived fibroblasts to multinucleated skeletal myotubes following lentiviral-induced expression of equine myogenic differentiation 1. *Am. J. Vet. Res.* **69**, 1637–1645 (2008).
 511. Chaouch, S. *et al.* Immortalized Skin Fibroblasts Expressing Conditional MyoD as a Renewable and Reliable Source of Converted Human Muscle Cells to Assess Therapeutic Strategies for Muscular Dystrophies: Validation of an Exon-Skipping Approach to Restore Dystrophin in Duchenne Muscular Dystrophy Cells. *Hum. Gene Ther.* **20**, 784–790 (2009).
 512. Sabourin, L. A. & Rudnicki, M. A. The molecular regulation of myogenesis. *Clin. Genet.* **57**, 16–25 (2001).
 513. Pini, V., Morgan, J. E., Muntoni, F. & O'Neill, H. C. Genome Editing and Muscle Stem Cells as a Therapeutic Tool for Muscular Dystrophies. *Curr. stem cell reports* **3**, 137–148 (2017).

514. Linkhart, T. A., Clegg, C. H. & Hauschka, S. D. Myogenic differentiation in permanent clonal mouse myoblast cell lines: Regulation by macromolecular growth factors in the culture medium. *Dev. Biol.* **86**, 19–30 (1981).
515. Wedel, D. J. Malignant hyperthermia and neuromuscular disease. *Neuromuscul. Disord.* **2**, 157–64 (1992).
516. Chang, N. C., Chevalier, F. P. & Rudnicki, M. A. Satellite Cells in Muscular Dystrophy - Lost in Polarity. *Trends Mol. Med.* **22**, 479–496 (2016).
517. Soriano-Arroquia, A., Clegg, P. D., Molloy, A. P. & Goljanek-Whysall, K. Preparation and Culture of Myogenic Precursor Cells/Primary Myoblasts from Skeletal Muscle of Adult and Aged Humans. *J. Vis. Exp.* (2017).
doi:10.3791/55047
518. Costa, M. L., Escaleira, R., Cataldo, A., Oliveira, F. & Mermelstein, C. S. Desmin: molecular interactions and putative functions of the muscle intermediate filament protein. *Brazilian J. Med. Biol. Res.* **37**, 1819–1830 (2004).
519. Li, H. *et al.* Inhibition of desmin expression blocks myoblast fusion and interferes with the myogenic regulators myoD and myogenin. *J. Cell Biol.* **124**, 827–841 (1994).
520. Li, Z. *et al.* Desmin is essential for the tensile strength and integrity of myofibrils but not for myogenic commitment, differentiation, and fusion of skeletal muscle. *J. Cell Biol.* **139**, 129–44 (1997).
521. Charrier, S. *et al.* Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. *Gene Ther.* **18**, 479–87 (2011).
522. Hildyard, J. C. W. & Wells, D. J. Identification and validation of quantitative PCR reference genes suitable for normalizing expression in normal and dystrophic cell culture models of myogenesis. *PLoS Curr.* **6**, (2014).
523. Saga, S., Hamaguchi, M., Hoshino, M. & Kojima, K. Expression of meta-vinculin associated with differentiation of chicken embryonal muscle cells. *Exp. Cell Res.* **156**, 45–56 (1985).
524. Dalby, B. *et al.* Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications. *Methods* **33**, 95–103 (2004).
525. Lee, M., Chea, K., Pyda, R., Chua, M. & Dominguez, I. Comparative Analysis

- of Non-viral Transfection Methods in Mouse Embryonic Fibroblast Cells. *J. Biomol. Tech.* **28**, 67–74 (2017).
526. Rahimi, P. *et al.* Comparison of transfection efficiency of polymer-based and lipid-based transfection reagents. *Bratislava Med. J.* **119**, 701–705 (2018).
 527. Perroud, J., Bernheim, L., Frieden, M. & Koenig, S. Distinct roles of NFATc1 and NFATc4 in human primary myoblast differentiation and in the maintenance of reserve cells. (2017). doi:10.1242/jcs.198978
 528. Brees, C. & Fransen, M. A cost-effective approach to microporate mammalian cells with the Neon Transfection System. *Anal. Biochem.* **466**, 49–50 (2014).
 529. Mamchaoui, K. *et al.* Immortalized pathological human myoblasts: towards a universal tool for the study of neuromuscular disorders. *Skelet. Muscle* **1**, 34 (2011).
 530. Muses, S., Morgan, J. E. & Wells, D. J. A new extensively characterised conditionally immortal muscle cell-line for investigating therapeutic strategies in muscular dystrophies. *PLoS One* **6**, e24826 (2011).
 531. Yáñez-Muñoz, R. J. *et al.* Effective gene therapy with nonintegrating lentiviral vectors. *Nat. Med.* **12**, 348–353 (2006).
 532. Li, S. *et al.* Stable transduction of myogenic cells with lentiviral vectors expressing a minidystrophin. *Gene Ther.* **12**, 1099–1108 (2005).
 533. Kimura, E., Li, S., Gregorevic, P., Fall, B. M. & Chamberlain, J. S. Dystrophin delivery to muscles of mdx mice using lentiviral vectors leads to myogenic progenitor targeting and stable gene expression. *Mol. Ther.* **18**, 206–13 (2010).
 534. Schlimgen, R. *et al.* Risks Associated With Lentiviral Vector Exposures and Prevention Strategies. *J. Occup. Environ. Med.* **58**, 1159–1166 (2016).
 535. Roy-Chowdhury, J. & Horwitz, M. S. Evolution of Adenoviruses as Gene Therapy Vectors. *Mol. Ther.* **5**, 340–344 (2002).
 536. Penaud-Budloo, M. *et al.* Adeno-associated virus vector genomes persist as episomal chromatin in primate muscle. *J. Virol.* **82**, 7875–85 (2008).
 537. Balci, B. & Dinçer, P. Efficient transfection of mouse-derived C2C12 myoblasts using a matrigel basement membrane matrix. *Biotechnol. J.* **4**, 1042–1045 (2009).
 538. Quenneville, S. P. *et al.* Nucleofection of muscle-derived stem cells and myoblasts with ϕ C31 integrase: stable expression of a full-length-dystrophin

- fusion gene by human myoblasts. *Mol. Ther.* **10**, 679–687 (2004).
539. Ye, L. *et al.* Nonviral Vector-Based Gene Transfection of Primary Human Skeletal Myoblasts. *Exp. Biol. Med.* **232**, 1477–1487 (2007).
 540. Trivedi, R. A. & Dickson, G. Liposome-mediated gene transfer into normal and dystrophin-deficient mouse myoblasts. *J. Neurochem.* **64**, 2230–8 (1995).
 541. Yamano, S., Dai, J. & Moursi, A. M. Comparison of Transfection Efficiency of Nonviral Gene Transfer Reagents. *Mol. Biotechnol.* **46**, 287–300 (2010).
 542. Dodds, E., Dunckley, M., Naujoks, K., Michaelis, U. & Dickson, G. Lipofection of cultured mouse muscle cells: a direct comparison of Lipofectamine and DOSPER. *Gene Ther.* **5**, 542–551 (1998).
 543. Boukhnikachvili, T. *et al.* Structure of in-serum transfecting DNA-cationic lipid complexes. *FEBS Lett.* **409**, 188–194 (1997).
 544. Jackson, M. F. *et al.* Genetic manipulation of myoblasts and a novel primary myosatellite cell culture system: comparing and optimizing approaches. doi:10.1111/febs.12072
 545. Chelly', J. *et al.* Quantitative estimation of minor mRNAs by cDNA-polymerase chain reaction Application to dystrophin mRNA in cultured myogenic and brain cells. *Eur. J. Biochem* **187**, (1990).
 546. Verkuijl, S. A. & Rots, M. G. The influence of eukaryotic chromatin state on CRISPR–Cas9 editing efficiencies. *Curr. Opin. Biotechnol.* **55**, 68–73 (2019).
 547. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).
 548. Takahashi, K. *et al.* Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, 861–872 (2007).
 549. Aasen, T. *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* **26**, 1276–1284 (2008).
 550. Staerk, J. *et al.* Reprogramming of Human Peripheral Blood Cells to Induced Pluripotent Stem Cells. *Cell Stem Cell* **7**, 20–24 (2010).
 551. Zhou, T. *et al.* Generation of human induced pluripotent stem cells from urine samples. *Nat. Protoc.* **7**, 2080–2089 (2012).
 552. Yu, J. *et al.* Induced Pluripotent Stem Cell Lines Derived from Human Somatic

- Cells. *Science* (80-.). **318**, 1917–1920 (2007).
553. Huangfu, D. *et al.* Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat. Biotechnol.* **26**, 795–797 (2008).
 554. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. Generation of Mouse Induced Pluripotent Stem Cells Without Viral Vectors. *Science* (80-.). **322**, 949–953 (2008).
 555. Lin, T. *et al.* A chemical platform for improved induction of human iPSCs. *Nat. Methods* **6**, 805–808 (2009).
 556. Zhou, W. & Freed, C. R. Adenoviral Gene Delivery Can Reprogram Human Fibroblasts to Induced Pluripotent Stem Cells. *Stem Cells* **27**, 2667–2674 (2009).
 557. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
 558. Somers, A. *et al.* Generation of Transgene-Free Lung Disease-Specific Human Induced Pluripotent Stem Cells Using a Single Excisable Lentiviral Stem Cell Cassette. *Stem Cells* **28**, 1728–1740 (2010).
 559. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* **85**, 348–62 (2009).
 560. Ban, H. *et al.* Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proc. Natl. Acad. Sci.* **108**, 14234–14239 (2011).
 561. Warren, L. *et al.* Highly Efficient Reprogramming to Pluripotency and Directed Differentiation of Human Cells with Synthetic Modified mRNA. *Cell Stem Cell* **7**, 618–630 (2010).
 562. Knoepfler, P. S. Deconstructing stem cell tumorigenicity: a roadmap to safe regenerative medicine. *Stem Cells* **27**, 1050–6 (2009).
 563. Chang, Y.-L. *et al.* Docosahexaenoic Acid Promotes Dopaminergic Differentiation in Induced Pluripotent Stem Cells and Inhibits Teratoma Formation in Rats with Parkinson-Like Pathology. *Cell Transplant.* **21**, 313–332 (2012).

564. Fang, I.-M., Yang, C.-H., Chiou, S.-H. & Yang, C.-M. Induced pluripotent stem cells without c-Myc ameliorate retinal oxidative damage via paracrine effects and reduced oxidative stress in rats. *J. Ocul. Pharmacol. Ther.* **30**, 757–70 (2014).
565. Sheridan, S. D., Surampudi, V. & Rao, R. R. Analysis of Embryoid Bodies Derived from Human Induced Pluripotent Stem Cells as a Means to Assess Pluripotency. *Stem Cells Int.* **2012**, 1–9 (2012).
566. Laurent, L. C. *et al.* Dynamic Changes in the Copy Number of Pluripotency and Cell Proliferation Genes in Human ESCs and iPSCs during Reprogramming and Time in Culture. *Cell Stem Cell* **8**, 106–118 (2011).
567. Quaas, A., Pomeroy, J., Huang, G., Paulson, R. J. & Pera, M. Early Markers of Reprogramming in Induced Pluripotent Stem Cells (iPSCs): A Timeline of Key Steps in the Reprogramming Process. *Fertil. Steril.* **95**, S5 (2011).
568. Abujarour, R. *et al.* Optimized Surface Markers for the Prospective Isolation of High-Quality hiPSCs using Flow Cytometry Selection. *Sci. Rep.* **3**, 1179 (2013).
569. Bharathan, S. P. *et al.* Systematic evaluation of markers used for the identification of human induced pluripotent stem cells. *Biol. Open* **6**, 100–108 (2017).
570. Kodaka, Y., Rabu, G. & Asakura, A. Skeletal Muscle Cell Induction from Pluripotent Stem Cells. *Stem Cells Int.* **2017**, 1–16 (2017).
571. Chal, J. & Pourquié, O. Making muscle: skeletal myogenesis in vivo and in vitro. *Development* **144**, 2104–2122 (2017).
572. Reshef, R., Maroto, M. & Lassar, A. B. Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. *Genes Dev.* **12**, 290–303 (1998).
573. Kuroda, K., Kuang, S., Taketo, M. M. & Rudnicki, M. A. Canonical Wnt signaling induces BMP-4 to specify slow myofibrogenesis of fetal myoblasts. *Skelet. Muscle* **3**, 5 (2013).
574. Münsterberg, A. E., Kitajewski, J., Bumcrot, D. A., McMahon, A. P. & Lassar, A. B. Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. *Genes Dev.* **9**, 2911–22 (1995).
575. Marcelle, C., Stark, M. R. & Bronner-Fraser, M. Coordinate actions of BMPs, Wnts, Shh and noggin mediate patterning of the dorsal somite. *Development*

- 124**, 3955–63 (1997).
- 576. Rios, A. C., Serralbo, O., Salgado, D. & Marcelle, C. Neural crest regulates myogenesis through the transient activation of NOTCH. *Nature* **473**, 532–535 (2011).
 - 577. Wood, W. M., Etemad, S., Yamamoto, M. & Goldhamer, D. J. MyoD-expressing progenitors are essential for skeletal myogenesis and satellite cell development. *Dev. Biol.* **384**, 114–127 (2013).
 - 578. Maffioletti, S. M. *et al.* Efficient derivation and inducible differentiation of expandable skeletal myogenic cells from human ES and patient-specific iPS cells. *Nat. Protoc.* **10**, 941–958 (2015).
 - 579. Chal, J. *et al.* Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969 (2015).
 - 580. Shelton, M., Kocharyan, A., Liu, J. & Skerjanc, I. S. Robust generation and expansion of skeletal muscle progenitors and myocytes from human pluripotent stem cells. *Methods* **101**, 73–84 (2016).
 - 581. Akiyama, T. *et al.* Efficient differentiation of human pluripotent stem cells into skeletal muscle cells by combining RNA-based MYOD1-expression and POU5F1-silencing. *Sci. Rep.* **8**, 1189 (2018).
 - 582. Dekel, I., Magal, Y., Pearson-White, S., Emerson, C. P. & Shani, M. Conditional conversion of ES cells to skeletal muscle by an exogenous MyoD1 gene. *New Biol.* **4**, 217–24 (1992).
 - 583. Tanaka, A. *et al.* Efficient and Reproducible Myogenic Differentiation from Human iPS Cells: Prospects for Modeling Miyoshi Myopathy In Vitro. *PLoS One* **8**, e61540 (2013).
 - 584. Buckingham, M. Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr. Opin. Genet. Dev.* **16**, 525–32 (2006).
 - 585. Barr, F. G. *et al.* Predominant expression of alternative PAX3 and PAX7 forms in myogenic and neural tumor cell lines. *Cancer Res.* **59**, 5443–8 (1999).
 - 586. Sakurai, H. *et al.* Bidirectional induction toward paraxial mesodermal derivatives from mouse ES cells in chemically defined medium. *Stem Cell Res.* **3**, 157–169 (2009).
 - 587. Xu, C. *et al.* A zebrafish embryo culture system defines factors that promote vertebrate myogenesis across species. *Cell* **155**, 909–21 (2013).

588. Wu, D. & Pan, W. GSK3: a multifaceted kinase in Wnt signaling. *Trends Biochem. Sci.* **35**, 161 (2010).
589. Makino, T. *et al.* Basic fibroblast growth factor stimulates the proliferation of human dermal fibroblasts via the ERK1/2 and JNK pathways. *Br. J. Dermatol.* **162**, 717–723 (2009).
590. Chal, J. *et al.* Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat. Protoc.* **11**, 1833–1850 (2016).
591. Borchin, B., Chen, J. & Barberi, T. Derivation and FACS-Mediated Purification of PAX3+/PAX7+ Skeletal Muscle Precursors from Human Pluripotent Stem Cells. *Stem Cell Reports* **1**, 620–631 (2013).
592. Liang, G. & Zhang, Y. Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* **13**, 149–59 (2013).
593. Kawakami, A., Kimura-Kawakami, M., Nomura, T. & Fujisawa, H. *Distributions of PAX6 and PAX7 proteins suggest their involvement early and late phases of chick brain development in both. Mechanisms of Development* **66**, (1997).
594. Blake, J. A. & Ziman, M. R. Pax genes: regulators of lineage specification and progenitor cell maintenance. *Development* **141**, 737–51 (2014).
595. Zhou, Z.-J. *et al.* Selection of Suitable Reference Genes for Normalization of Quantitative Real-Time Polymerase Chain Reaction in Human Cartilage Endplate of the Lumbar Spine. *PLoS One* **9**, e88892 (2014).
596. Young, C. S. *et al.* 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**, 533–40 (2016).
597. Long, C. *et al.* Correction of diverse muscular dystrophy mutations in human engineered heart muscle by single-site genome editing. *Sci. Adv.* **4**, eaap9004 (2018).
598. Niemitz, E. Patient-derived isogenic iPSCs. *Nat. Genet.* **43**, 824–824 (2011).
599. Wang, Y. *et al.* Genome editing of human embryonic stem cells and induced pluripotent stem cells with zinc finger nucleases for cellular imaging. *Circ. Res.* **111**, 1494–503 (2012).
600. Howden, S. E., Vanslambrouck, J. M., Wilson, S. B., Sin Tan, K. & Little, M. H. Fate-mapping within human kidney organoids reveals conserved mammalian nephron progenitor lineage relationships. (2018). doi:10.1101/432161

601. Ihry, R. J. *et al.* p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* **24**, 939–946 (2018).
602. Yin, H., Price, F. & Rudnicki, M. a. Satellite cells and the muscle stem cell niche. *Physiol. Rev.* **93**, 23–67 (2013).
603. Beauchamp, J. R., Morgan, J. E., Pagel, C. N. & Partridge, T. A. Dynamics of myoblast transplantation reveal a discrete minority of precursors with stem cell-like properties as the myogenic source. *J. Cell Biol.* **144**, 1113–22 (1999).
604. Hosoyama, T., McGivern, J. V., Van Dyke, J. M., Ebert, A. D. & Suzuki, M. Derivation of Myogenic Progenitors Directly From Human Pluripotent Stem Cells Using a Sphere-Based Culture. *Stem Cells Transl. Med.* **3**, 564–574 (2014).
605. Hicks, M. R. *et al.* ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development and hPSCs. *Nat. Cell Biol.* **20**, 46–57 (2018).
606. Wu, J. *et al.* A Myogenic Double-Reporter Human Pluripotent Stem Cell Line Allows Prospective Isolation of Skeletal Muscle Progenitors. *CellReports* **25**, 1966–1981.e4 (2018).
607. Sherr, C. J. & DePinho, R. A. Cellular senescence: mitotic clock or culture shock? *Cell* **102**, 407–10 (2000).
608. Long, C. *et al.* Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* **345**, 1184–8 (2014).
609. Ousterout, D. G. *et al.* Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat. Commun.* **6**, 6244 (2015).
610. Brunet, E. *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc. Natl. Acad. Sci.* **106**, 10620–10625 (2009).
611. Soltanzadeh, P. *et al.* Clinical and genetic characterization of manifesting carriers of DMD mutations. *Neuromuscul. Disord.* **20**, 499–504 (2010).
612. Lin, M. *et al.* Effects of short indels on protein structure and function in human genomes. *Sci. Rep.* **7**, 9313 (2017).
613. Danisovic, L., Culenova, M. & Csobonyeiova, M. Induced Pluripotent Stem Cells for Duchenne Muscular Dystrophy Modeling and Therapy. *Cells* **7**, (2018).
614. Falzarano, M. S. *et al.* Duchenne Muscular Dystrophy Myogenic Cells from Urine-Derived Stem Cells Recapitulate the Dystrophin Genotype and

Phenotype. *Hum. Gene Ther.* **27**, 772–783 (2016).

615. Zhang, D., Wei, G., Li, P., Zhou, X. & Zhang, Y. Urine-derived stem cells: A novel and versatile progenitor source for cell-based therapy and regenerative medicine. *Genes Dis.* **1**, 8–17 (2014).
616. Lattanzi, A. *et al.* Correction of the Exon 2 Duplication in DMD Myoblasts by a Single CRISPR/Cas9 System. *Mol. Ther. Nucleic Acids* **7**, 11–19 (2017).
617. Arpke, R. W. *et al.* 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 (2013).
618. Karpati, G. *et al.* Age-related conversion of dystrophin-negative to -positive fiber segments of skeletal but not cardiac muscle fibers in heterozygote mdx mice. *J. Neuropathol. Exp. Neurol.* **49**, 96–105 (1990).
619. Brown, B. D. *et al.* In vivo administration of lentiviral vectors triggers a type I interferon response that restricts hepatocyte gene transfer and promotes vector clearance. *Blood* **109**, 2797–805 (2007).
620. Qin, J. Y. *et al.* Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* **5**, e10611 (2010).
621. Zhang, X.-H., Tee, L. Y., Wang, X.-G., Huang, Q.-S. & Yang, S.-H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol. Ther. - Nucleic Acids* **4**, e264 (2015).
622. Kleinjan, D.-J. & Coutinho, P. Cis-rupture mechanisms: disruption of cis-regulatory control as a cause of human genetic disease. *Briefings Funct. Genomics Proteomics* **8**, 317–332 (2009).
623. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
624. Frock, R. L. *et al.* Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.* **33**, 179–186 (2015).
625. Tsai, S. Q. *et al.* CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat. Methods* **14**, 607–614 (2017).
626. Friedland, A. E. *et al.* Characterization of *Staphylococcus aureus* Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol.* **16**, 257 (2015).
627. Lessard, S. *et al.* Human genetic variation alters CRISPR-Cas9 on- and off-

- targeting specificity at therapeutically implicated loci. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E11257–E11266 (2017).
628. Jensen, K. T. *et al.* Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.* **591**, 1892–1901 (2017).
 629. Uusi-Mäkelä, M. I. E. *et al.* Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (*Danio rerio*). *PLoS One* **13**, e0196238 (2018).
 630. Ran, F. A. *et al.* In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* **520**, 186–91 (2015).
 631. Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403–7 (2016).
 632. Long, C. *et al.* Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400–3 (2016).
 633. Bengtsson, N. E. *et al.* Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. *Nat. Commun.* **8**, 14454 (2017).
 634. Arnett, A. L. *et al.* Adeno-associated viral (AAV) vectors do not efficiently target muscle satellite cells. *Mol. Ther. Methods Clin. Dev.* **1**, 14038 (2014).
 635. Dumont, N. A., Bentzinger, C. F., Sincennes, M.-C. & Rudnicki, M. A. Satellite Cells and Skeletal Muscle Regeneration. *Compr. Physiol.* **5**, 1027–59 (2015).
 636. Newall, A. E. *et al.* Primary non-random X inactivation associated with disruption of Xist promoter regulation. *Hum. Mol. Genet.* **10**, 581–9 (2001).
 637. van Putten, M. *et al.* The effects of low levels of dystrophin on mouse muscle function and pathology. *PLoS One* **7**, e31937 (2012).
 638. van Putten, M. *et al.* Low dystrophin levels increase survival and improve muscle pathology and function in dystrophin/utrophin double-knockout mice. *FASEB J.* **27**, 2484–95 (2013).
 639. van Putten, M. *et al.* Low dystrophin levels in heart can delay heart failure in mdx mice. *J. Mol. Cell. Cardiol.* **69**, 17–23 (2014).
 640. Godfrey, C. *et al.* How much dystrophin is enough: the physiological consequences of different levels of dystrophin in the mdx mouse. *Hum. Mol. Genet.* **24**, 4225–4237 (2015).

641. Sharp, P. S., Bye-a-Jee, H. & Wells, D. J. Physiological Characterization of Muscle Strength With Variable Levels of Dystrophin Restoration in mdx Mice Following Local Antisense Therapy. *Mol. Ther.* **19**, 165–171 (2011).
642. Neri, M. *et al.* Dystrophin levels as low as 30% are sufficient to avoid muscular dystrophy in the human. *Neuromuscul. Disord.* **17**, 913–8 (2007).
643. Charlesworth, C. T. *et al.* Identification of Pre-Existing Adaptive Immunity to Cas9 Proteins in Humans. doi:10.1101/243345
644. Bhat, P., Leggatt, G., Waterhouse, N. & Frazer, I. H. Interferon- γ derived from cytotoxic lymphocytes directly enhances their motility and cytotoxicity. *Cell Death Dis.* **8**, e2836 (2017).
645. Rosenberg, A. S. *et al.* Immune-mediated pathology in Duchenne muscular dystrophy. *Sci. Transl. Med.* **7**, 299rv4 (2015).
646. Crudele, J. M. & Chamberlain, J. S. Cas9 immunity creates challenges for CRISPR gene editing therapies. doi:10.1038/s41467-018-05843-9
647. Salva, M. Z. *et al.* Design of Tissue-specific Regulatory Cassettes for High-level rAAV-mediated Expression in Skeletal and Cardiac Muscle. *Mol. Ther.* **15**, 320–329 (2007).

APPENDIX I: Formulation of culture medium used during iPSCs differentiation

The name of each medium recapitulates its main components, as indicated below:

D = Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12

C = CHIR99021

L = LDN193189

F = bFGF

K = knock-out serum (KSR)

I = IGF-1

H = HGF

DICL medium		
Reagent	Working conc.	Note
DF12	1x	Daily medium change
IST	1x	
NEAA	1x	
CHIR	3 μ M	
LDN	0.5 μ M	

DICLF medium		
Reagent	Working conc.	Note
DF12	1x	Daily medium change
IST	1x	
NEAA	1x	
CHIR	3 μ M	
LDN	0.5 μ M	
bFGF	10ng/ml	

DKHIFL medium		
Reagent	Working conc.	Note
DF12	1x	Daily medium change
NEAA	1x	
KSR	150 µl/ml	
β-ME	2µl/ml	
HGF	10ng/ml	
IGF-1	2ng/ml	
bFGF	20ng/ml	
LDN	0.5µM	

DKI proliferation medium		
Reagent	Working conc.	Note
DMEM	1x	Daily medium change
KSR	15%	
NEAA	1x	
β-ME	2µl /ml	
IGF-1	2ng/ml	

DKHI proliferation medium		
Reagent	Working conc.	Note
DF12	1x	change medium every other day
KSR	15%	
NEAA	1x	
β-ME	200 ul/100 ml	
HGF	10ng/ml	
IGF-1	2ng/ml	

Passage cells		
Reagent	per well of 6-well plate	Note
Collagenase	1ml	Split 1 to 3-5 ratio
Dispase	100ul	