GENE EDITING AND HUMAN INDUCED PLURIPOTENT STEM CELLS FOR ADVANCED MODELLING OF DUCHENNE MUSCULAR DYSTROPHY

Moustafa Omar Ali Omar Khedr, MRes.

University College London

Department of Cell and Developmental Biology

A thesis submitted for the degree of

Doctor of Philosophy

June 2022

DECLARATION OF AUTHORSHIP

I, Moustafa Omar Khedr, confirm that the work presented in this thesis is my own. Where information has been derived from other sources or as part of collaborative work, I confirm that this has been indicated in the thesis.

ABSTRACT

Duchenne muscular dystrophy (DMD) is a rare x-linked genetic disorder characterised by progressive muscle wasting and severe weakness. It is caused by absence or dysfunction of dystrophin, a protein encoded by the largest gene known in nature that protects muscle cell membrane from contraction-induced damage. To date DMD remains incurable.

Several experimental therapies are being developed for DMD, including exon-skipping using anti-sense oligonucleotides (AONs), read-through technologies using small molecules, and gene therapy and genome editing strategies. However, the development and optimisation of these therapies and determining their relative efficiency is limited by the lack of an overarching (temporally, economically, and ethically viable) platform to model DMD and human dystrophin dynamics.

Here CRISPR/Cas9 gene editing is utilised to generate human induced pluripotent stem (iPS) cells that can be used to detect dystrophin protein production qualitatively and quantitatively, by inserting a multifunctional reporter cassette capable of tracking dystrophin expression temporally and spatially, in real time and in fixed cells. Reporter functionality was assessed via transgene-free differentiation of iPS cells into skeletal muscle cells. Dystrophin-reporting iPS cells were also gene-edited to induce pathogenic DMD mutations in exon 52 which were targeted by mutation-specific AONs.

In parallel, iPS cells were also differentiated in 3D artificial muscle constructs to model DMD in comparison to traditional monolayer cultures, with the aim of combining the potential of both platforms (multifunctional reporter and 3D tissue architecture) into a novel advanced DMD modelling platform to accelerate development of personalised and precision therapies for DMD.

IMPACT STATEMENT

The work presented in this thesis has a significant translational and, on the longer run, societal impact. It demonstrates the superiority of the developed dystrophin reporting human iPS cell lines, their derivatives and 3D artificial skeletal muscle constructs to model human dystrophin expression dynamics in a more physiologically relevant *quasi vivo* environment. This in turn lays the foundations for screening DMD therapies using the platform. DMD in the most common muscular dystrophy of childhood and remains incurable to date: hence this work has significant potential to impact on optimisation of current therapeutics as well as development of new molecule and genetic treatment to tackle this devastating disorder. The need for a novel tool is further highlighted by the orphan nature of the disease, its severity and fatality, and the fact that the patient population is highly heterogenous, mutation-wise, which requires different therapeutic targets and approaches.

This work has also significant impact on enhancing knowledge on molecular and cellular dynamics of dystrophin in skeletal muscle and beyond, as the same iPS cells can be differentiated in other tissues involved in the disorder such as cardiomyocytes and neurons.

This work generated tangible tools such as new genetic reporters and cell lines, which are currently being assessed with our technology transfer and translational team to harness their commercialisation potential, which we believe

could be significant given the interest of several major companies in developing experimental therapies for DMD.

Finally, this work has a major impact on animal welfare, as it harnesses the potential of emerging technologies such as gene editing and human 3D cultures to transition the early stages of therapeutic discovery for DMD from animal models (mostly mice) to high resolution and fidelity humanised platforms. As a tangible result of this, over the past 5 years our laboratory has reduced the number of mice used in muscular dystrophy projects of approximately 90%.

ACKNOWLEDGEMENTS

I would like to extend my thanks to my supervisors Prof. Francesco Saverio Tedesco and Prof. Francesco Muntoni for the opportunity to pursue my PhD under their supervision, and Prof. Michael Duchen for his help and support in his role as my graduate tutor.

Thanks to Muscular Dystrophy UK for the funding towards my doctoral studies, the European Research Council for supporting this and other research projects in the laboratory, the Francis Crick Institute for the opportunity to work there most of my PhD, and The Company of Biologists and UCL for funding my conference and course participations abroad.

I am indeed very grateful to many people who have supported this work: Dr. Hidetoshi Hoshiya and Dr. Louise Moyle who laid the foundations of the project, Prof. Konstantinos Anastassiadis and Dr. Katrin Neumann who welcomed me in Dresden and gave me valuable insight and guidance, and members of the Tedesco lab for the help, suggestions and engaging discussions.

I truly enjoyed my PhD studies because of my great colleagues, whom I consider friends now. So, a huge thank you to Eleonora, Luca P., Luca R., SungWoo, Noreen, Valentina, Lorenza and Daniel. I would also like to particularly thank Luca, Shilipita, Valentina and Elisa for all the fruitful collaborative work, SungWoo for proof-reading my discussion, and students I helped supervise, MengMeng and Margaux.

I am blessed with many friends, scattered around the world, whom I shared lovely moments with and can truly depend on. Omar, Lorraine, Amr, Manon, Alexandra, Nour, Sara, Khalid, Nadim, Maria, Rowan, and Farah thank you for everything.

Growing up, I was surrounded by a loving family that gave me with an amazing childhood and continue to be my biggest supporters. My grandmas, Myasser and Sawsan, my aunts, Hana, Heba, Hoda and Magda, thank you for being in my life and always believing in me. I would also like to thank my cousins, especially Sara and Yousr whom I consider my sisters, for everything; the memories of the grand plans and dreams we had with Aya, Tarek and Aly still resonate with me and your love and support is what keeps me going.

Finally, and most importantly, I would like to dedicate this thesis to my parents and siblings. I never express my gratitude to you, but I am truly blessed to have you all in my life. I couldn't have achieved anything without your unwavering support, never-ending encouragement, and belief in me. Mama and Papa, you raised us with compassion, taught us to be independent and go after our dreams, and continue to be our rock in life. Aya, Aly and Tarek, I couldn't have asked for better siblings, I am lucky to have shared my childhood with you and look forward to always being together and there for each other. اسرتي السعيدة, I am forever indebted to you, I love you!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

3D	Three dimensional
AAV	Adeno Associated Virus
AChR	Acetylcholine receptor
AON	Antisense Oligonucleotide
Beta FGF	basic fibroblast growth factor
BMD	Becker muscular dystrophy
BMP	Bone morphogenic protein
BSD	Blasticidin
CD54	Cluster of differentiation 54
c-Met	cancer mesenchymal epithelial transition factor
cAMP	cyclic adenosine monophosphate
CK	creatine kinase
CRISPR	Clustered regularly interspaced short palindromic repeats
CXCR4	chemokine receptor
DGC	Dystrophin glycoprotein complex
DMD	Duchenne muscular dystrophy
DYS-HAC	dystrophin human artificial chromosome
EB	embryoid body
ERBB3	receptor tyrosine-protein kinase erbB-3
ESC	embryonic stem cell
Flp	Flippase
FRT	Flippase recognition target

GSK	glycogen synthase kinase
HAC	human artificial chromosome
hESCs	human embryonic stem cells
HGF	hepatocyte growth factor
HIDEM	Human iPS cell-derived mesoangioblast-like cell
hiPSC	human induced pluripotent stem cell
HNK-1	human natural killer-1
IGF	insulin-like growth factor
iPS	induced pluripotent stem
Klf4	krüppel-like factor 4
LMNA	lamin a/c
MDX	x-linked muscular dystrophy
MMCT	microcell mediated chromosome transfer
MN	motor neuron
Myf5	myogenic factor 5
MyHC	myosin heavy chain
NCAM	neural cell adhesion molecule
NGFR	nerve growth factor receptor
NMC	neuromuscular circuits
NMJ	neuromuscular junction
Oct4	octamer binding transcription factor 4
PAX	paired box
(D) PBS	(distilled) Phosphate Buffered Saline

PVDF	polyvinylidene difluoride
SDC2	syndecan 2
SC	satellite cell
Sox2	SRY (sex determining region Y)-box 2
ssDNA	single stranded DNA
TALEN	Transcription activator-like effector nucleases
TBS	Tris-buffered saline
TGFβ	transforming growth factor β
WNT	wingless-related integration site
UK	United Kingdom
ZFN	zinc finger protein

1 INTRODUCTION

1.1 SKELETAL MUSCLE

1.1.1 STRUCTURE AND FUNCTION

Skeletal muscle is the most abundant tissue in the human body, representing between 30% and 38% of its mass (Janssen *et al.*, 2000). It consists of large bundles of multinucleated cells, known as myofibers, which are derived from the fusion of myoblasts into a syncytial architecture which is detrimental to its function, contraction and relaxation for movement (Capers, 1960). A myofibre is a bundle of myofibrils that in turn consists of units of sarcomeres that are aligned. Sarcomeres are composed of alternating thin actin filaments and thick myosin, which results in the striated phenotype of skeletal muscle tissue (Exeter and Connell, 2010). Myofibres range in length and diameter (from 10-100 μm) depending on the muscle type and location (Hopkins, 2006) and are singly innervated. Connective tissues encase the different layers and components of the muscle fibre, giving the skeletal muscle tissue the final shape; the endomysium covers the muscle fibre, the perimysium encases bundles of myofibres into fascicles, and finally, the epimysium covers the entire muscle. All of these connective tissue types fuse into the tendon at both ends of each muscle (Frontera and Ochala, 2015).

Skeletal muscle main functions are locomotion, pulmonary ventilation, posture maintenance, metabolism and temperature control (Charge and Rudnicki, 2004).

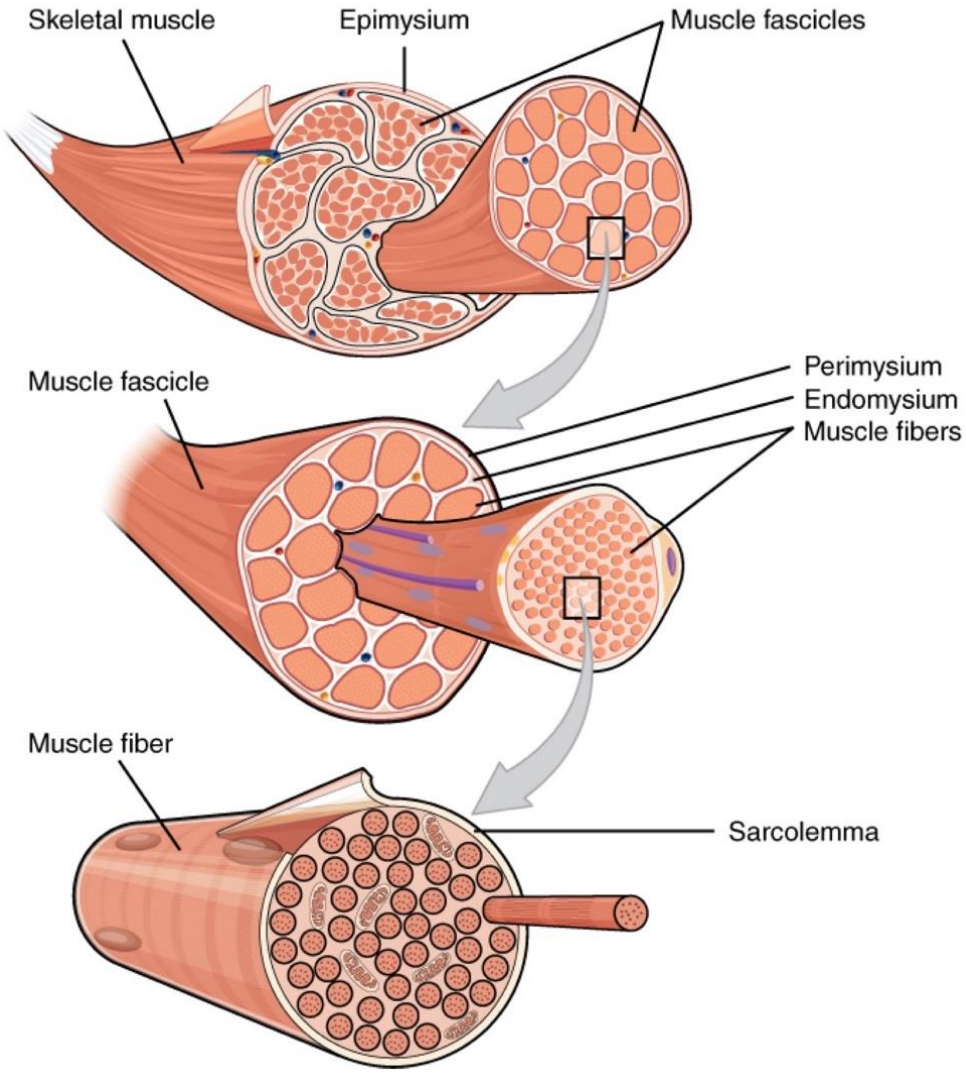

Figure 1-1. *Skeletal muscle structure, showing fibres and the three connective layers.*
Adapted from the book (Betts et al., 2017).

1.1.2 CONTRACTILE PROPERTIES

Contraction in skeletal muscles is controlled by actin-filament linked proteins, tropomyosin and troponin, that, upon calcium influx through calcium channels as a result of an action potential (through the motor neuron axon, delivered at the neuromuscular junction, NMJ) enable contraction through cross-bridge cycling between actin and the myosin heads (Exeter and Connell, 2010). According to Huxley (Huxley, 1969), the relative sliding of thick and thin microfilaments (actin and myosin, respectively) induces the muscle contraction, with the myosin heads pulling the actin filaments towards the centre of the sarcomere (Huxley, 1969). Abundantly lying parallel to the filaments is titin, an important protein that serves as scaffolding to stabilise the sarcomere structure through anchoring the central location of the myosin thick filaments, in addition to that, it acts as a spring that allows the muscle fibre to return to resting length following contraction, majorly contributing to the elasticity and extensibility of myofibrils (Sherwood, no date; Cooper, 2000).

1.1.3 DEVELOPMENT AND MYOGENESIS

Skeletal muscle early progenitors can be traced back to the transient embryonic structures known as somites, that are formed by the segmented paraxial mesoderm, on either side of the neural tube (Buckingham, 2006). The exception to this is the craniofacial group of muscles which have a different embryonic origin (Christ and Ordahl, 1995; Noden and Francis-West, 2006). In

general, amniotes' myogenesis consists of four phases, embryonic, foetal, neonatal, and adult myogenesis (Murphy and Kardon, 2011).

In primary/embryonic myogenesis, myogenic precursors specify as myoblasts then differentiate into post-mitotic mononuclear myocytes, which fuse into primary myofibres (Murphy and Kardon, 2011) and derive from Paired Box Protein (Pax)3^+, in mouse models, or $Pax3^+/Pax7^+$, in chicken models), dermomyotomal cells (Horst *et al.*, 2006; Otto, Schmidt and Patel, 2006). Initially, primary myofibres with few myonuclei act as a scaffold for more myogenic progenitors. Then during secondary/foetal myogenesis, the myofibres re-organise, as a result of the muscle innervation, and these become the bulk of post-natal muscle tissue (McClearn and Noden, 1988; Buckingham *et al.*, 2003; Murphy and Kardon, 2011).

After this stage, satellite cells (SCs), which are skeletal muscle stem cells, are present in post-natal tissue and are identified in their niche, beneath the basal lamina of muscle fibres (Mauro, 1961). SCs originally arise in the central domain of the dermomyotome and are marked by the expression of the transcription factors paired-box 3 and 7 (Relaix *et al.*, 2005; Schienda *et al.*, 2006). SCs migrate in the developing muscles and invade the tissue, partially fusing with secondary muscle fibres and contributing to the generation of multinucleated muscle fibres. It is common belief that the majority of nuclei in adult skeletal muscle fibres are derived from differentiated SCs (Zhang, Koishi and McLennan, 1998). A subset of these SCs returns to the quiescent state and is reserved as

the stem cell pool in post-natal skeletal muscle tissue, allowing muscle growth and regeneration upon their reactivation.

Elegant studies using a series of knockout mouse models helped elucidate the molecular circuitry of genes required for muscle formation. It is from these studies that *Pax3* and *Pax7* were identified as the main transcription factors associated with myogenesis (Buckingham and Relaix, 2015). Pax3 acts predominantly during the earlier stages of development, namely during embryonic development, and while Pax7 is not required for embryonic myogenesis (Seale *et al.*, 2000), double Pax3/Pax7 mutants generate early myotome but completely lose muscle formation due to lack of stem cells (Kassar-Duchossoy *et al.*, 2005; Relaix *et al.*, 2005, 2006).

Downstream of the Pax transcription factors are myogenic factor 5 (Myf5), myogenic regulatory factor 4 (Mrf4) and myogenic differentiation 1 (MyoD), which are critical for further tissue specification, validated through studies utilising triple knockout mice in which embryos demonstrated complete lack of muscle (Valdez *et al.*, 2000).

In adult organisms, skeletal muscle is characterised by its remarkable regenerative capacity and potential. Research using animal models allowed the elucidation of this process, including the signalling and cellular components involved, and classification of the major steps: 1. Skeletal muscle tissue injury, 2. Inflammatory response, 3. Activation, proliferation and differentiation of SCs, 4. Maturation and remodelling of newly formed tissue.

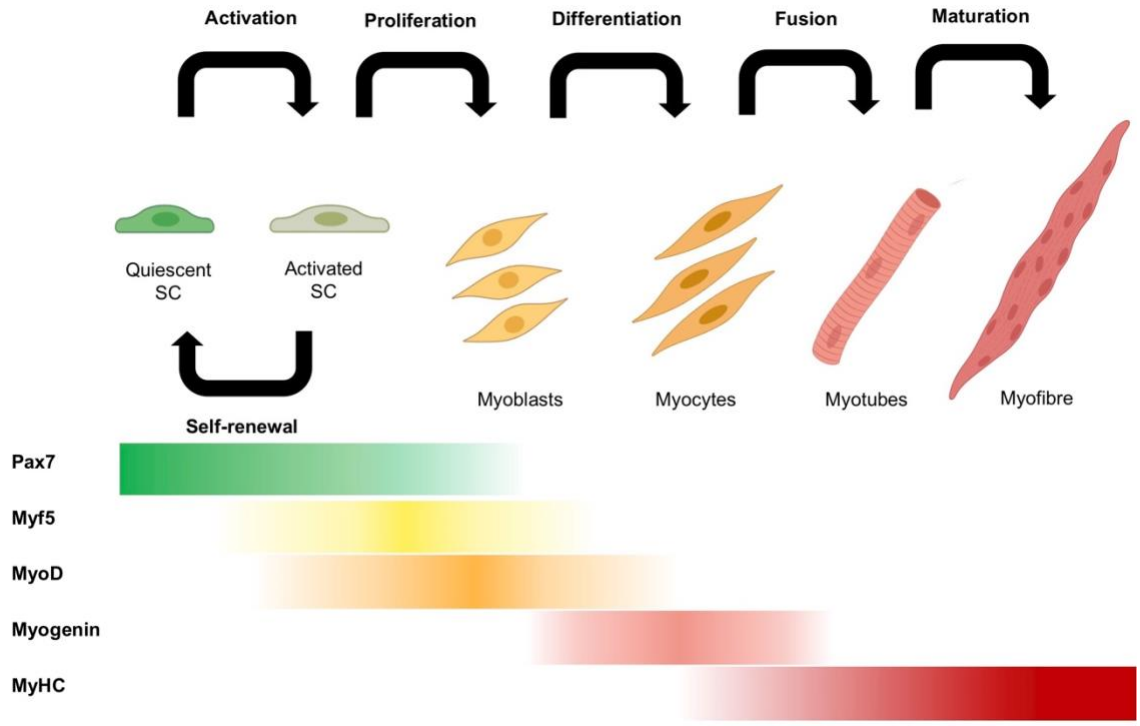

Figure 1-2. *Satellite cell activation and differentiation during muscle regeneration.*
Top: Major steps and key cell states from satellite cells to differentiated myofibers. Bottom: Key transcription factors and regulators expression correlating to cell states in top figure. Graphical illustration generated using BioRender.com and Microsoft Powerpoint.

Mature muscle fibres have a resident population of satellite cells (around 4% in humans (Schmalbruch and Hellhammer, 1976)) that lies beneath the basal lamina of adult muscle fibres (Mauro, 1961).

SCs are quiescent in normal adult muscle tissue states, at G_0 phase of cell cycle and express Pax7 (Zammit, Partridge and Yablonka-Reuveni, 2006), the only exception to this is a subset that expresses Pax3 located in tissues like the diaphragm (Relaix *et al.*, 2006).

Upon injury, SCs activate, asymmetrically proliferate into a quiescent pool of progenitors replenishing the stem cell niche, returning under the basal lamina, and a progeny of committed myoblasts that regenerates skeletal muscle tissue (Siegel, Kuhlmann and Cornelison, 2011). First, the progeny, known as transient amplifying progenitors, prime to myogenic differentiation, progressively losing Pax7 expression and acquiring Myf5 and MyoD expression (Cooper *et al.*, 1999). The activation of SCs is induced by the factors released in injured inflamed tissue such as transforming growth factor beta (TGF-β), fibroblast growth factor (FGF), hepatocyte growth factor (HGF), insulin-like growth factor (IGF) I/ II and interleukin 6 (IL-6) (Chargé and Rudnicki, 2004). The pro-inflammatory environment is gradually reversed and factors supporting myoblast proliferation are enriched (Chazaud *et al.*, 2003; Arnold *et al.*, 2007) promoting myoblast proliferation and fusion to each other and already existing fibres, repairing the tissue (Schwander *et al.*, 2003). The newly formed myofibres then undergo maturation, during which expression of embryonic and neonatal myosin heavy chain (MyHC) proteins are expressed, in a pattern that mirrors that of development (Schiaffino *et al.*, 2015).

1.2 MUSCULAR DYSTROPHIES AND DUCHENNE MUSCULAR DYSTROPHY (DMD)

1.2.1 *Muscular Dystrophies*

Inherited myopathies are genetic muscle disorders, which include metabolic myopathies, caused by muscle energy metabolism defects (Berardo, DiMauro and Hirano, 2010), congenital myopathies, which mostly present during childhood and have complex phenotype-genotype correlation (North *et al.*, 2014), and finally muscular dystrophies.

Muscular dystrophies are a group of inherited muscular disorders, and while they are heterogenous, all characterised by progressive degeneration of myfibres that leads to muscle weakness, often with cardiac and respiratory muscle involvement (Emery, 2002; Mercuri and Muntoni, 2013). These disorders are considered incurable and have grave prognoses, including disability and premature death.

1.2.2 *Clinical Features and Epidemiology of DMD*

Duchenne muscular dystrophy (DMD), a type of muscular dystrophy, is the most common muscle disease of childhood, affecting around 1 in 5,000 males in the UK (Muntoni, Torelli and Ferlini, 2003). The DMD disease phenotype onset is during early childhood, between 3 and 5 years of age, starting with muscle weakness (initially proximally). The disease progresses rapidly, and patients lose ambulation in the second decade of life, becoming wheelchair reliant in their teens. Some patients live into their 30s and the majority die of respiratory failure (Mercuri and Muntoni, 2013). DMD affects cardiac as well as skeletal muscles,

and also clinically manifests as dilated cardiomyopathy affecting the majority of patients by age 18 (Emery, 2002).

DMD disease manifestations also occur in other tissues, and mostly lead to cognitive impairment and effects, which is a feature in approximately one third of DMD patients (Ricotti *et al.*, 2016). In a recent meta-analysis report of more than 1000 patients with a dystrophin mutation, DMD and BMD patients, the mean of full-scale IQ scores was one standard deviation below population mean, with a range of scores from above average to severe impairment (Cotton, Voudouris and Greenwood, 2001). Delay in language development can be used as an early sign of the disease (Ciafaloni *et al.*, 2009) as a disproportionate effect on verbal working memory skills has been validated (Cotton, Voudouris and Greenwood, 2001; Ricotti, Roberts and Muntoni, 2011).

1.2.3 GENETICS AND PATHOPHYSIOLOGY OF DMD

DMD is caused by mutations in the *dystrophin* gene (Hoffman *et al.*, 1988). Dystrophin is an X-linked gene (Xq21) and is the largest human gene consisting of 79 exons and spanning around 2.5Mb (Muntoni, Torelli and Ferlini, 2003).

The dystrophin protein is a rod-shaped cytoplasmic protein which, in skeletal muscles, is localised at the sarcolemma as a part of the large Dystrophin-associated Glycoprotein Complex (DGC) (Muntoni, Torelli and Ferlini, 2003). The DCG is traditionally recognised as a shock absorber connecting the actin-myosin cytoskeleton, muscle membrane and the extracellular matrix, stabilising myofibres during contractions (Petrof *et al.*, 1993). But more recently, it has been

discovered that it also plays an important cell-signaling role in different processes, including: blood perfusion, vasoregulation, muscle performance and cell survival, and finally satellite cell asymmetric cell division (Rando, 2001; Dumont *et al.*, 2015; Allen, Whitehead and Froehner, 2016).

The majority (60-65%) of DMD cases arise due to deletions in the dystrophin genetic code. Additionally, duplications account for 5-15% of cases, and the remainder of cases are a result of point mutations, exonic insertions, intronic deletions, and multiple repetitive sequences. These mutations can either maintain or disrupt the reading frame, which is an important determinant of the severity/form of the disease. However, the size of the deletion does not correlate with the severity of the disease (Muntoni, Torelli and Ferlini, 2003). Patients with in-frame mutations are usually grouped under a milder allelic form of DMD: Becker's Muscular Dystrophy (BMD), in which the functionality of dystrophin is mostly preserved (England *et al.*, 1990). Interestingly, in 16% of the patients registered on the eDystrophin database carry in-frame mutations with severe DMD phenotype (Monaco *et al.*, 1988).

The dystrophin protein has multiple isoforms, and at least 7 different promoters. The three full-length isoforms (Dp427), derived from the expression of three tissue-specific promoters (however, not exclusively), are: muscle (M), brain (B), Purkinje (P). These contain an identical set of 78 exons, however, the first exon is spliced differently resulting in the unique tissue-specific isoforms (Górecki *et al.*, 1992). Some of the other isoform variants are also generated in a tissue-specific fashion such as Dp260 in the retina, Dp116 in peripheral nerves,

in contrast to Dp71 which is ubiquitously expressed apart from within skeletal muscles (Muntoni, Torelli and Ferlini, 2003).

The two other isoforms of Dp427 (C and P) are expressed in the human brain (Holder, Maeda and Bies, 1996; Doorenweerd *et al.*, 2017), but Dp140 and Dp71 are more abundantly found in the brain (Bar *et al.*, 1990; Muntoni, Torelli and Ferlini, 2003; Doorenweerd *et al.*, 2017). Several studies have demonstrated the role of both of these isoforms in the central nervous system (Felisari *et al.*, 2000; Moizard *et al.*, 2000; Daoud *et al.*, 2009; Ricotti *et al.*, 2016). This explains why patients with different mutations will have different cognitive impairment levels, depending on which dystrophin isoforms are affected.

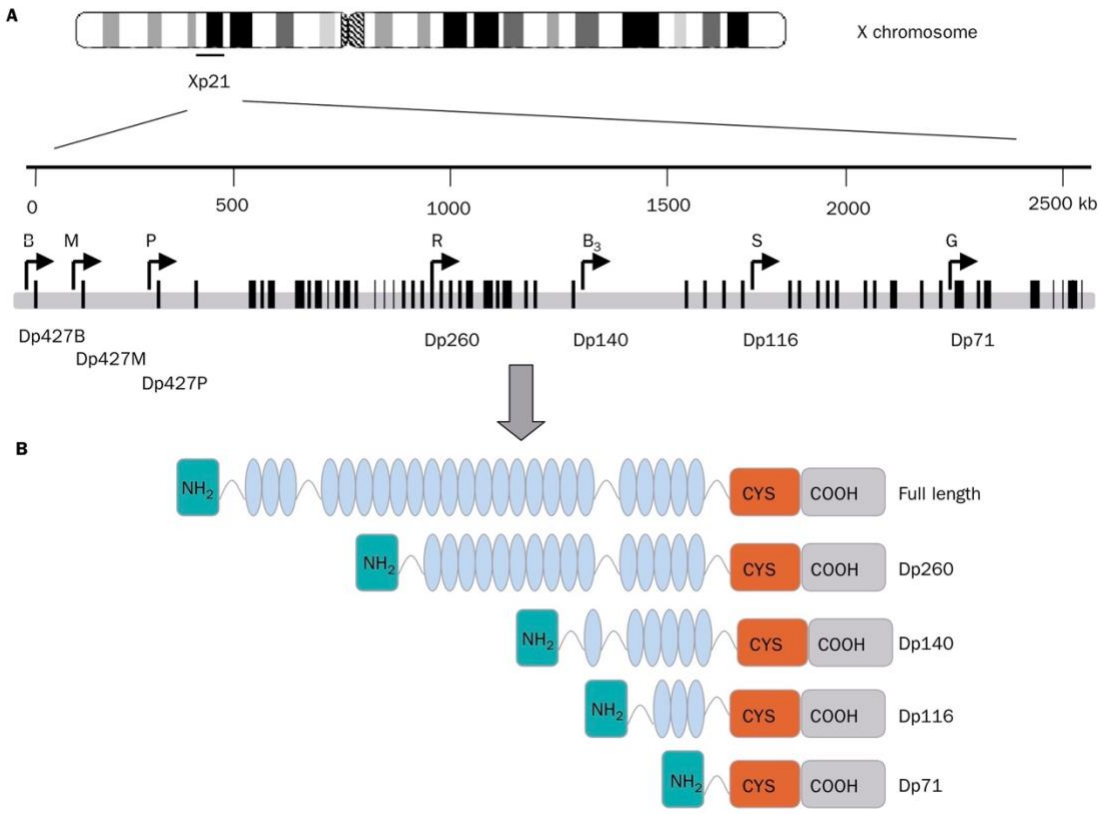

Figure 1-3. *A. Genomic sequence of the dystrophin gene, located in Xp21, black vertical lines represent the 79 exons of dystrophin.*
The arrows indicate the various promoters: are brain (B), muscle (M), and Purkinje (P) promoters; R, B3, S, and G represent the Dp260 (retinal), Dp140 (brain3), Dp116 (Schwann cells), and Dp71 (general) promoters. B: The domain composition of the various dystrophin proteins is indicated. The amino-terminal domain is followed by the spectrin like domain, the cysteine rich, and the carboxy-terminal domain. From (Muntoni, Torelli and Ferlini, 2003).

To summarise, the absence of a functional dystrophin disturbs the DGC, leading to the delocalisation of the other complex proteins and disruption of the cytoskeleton, resulting in increased susceptibility to mechanical stress and membrane instability (Deconinck and Dan, 2007). The loss of the DGC's main function (connecting the intracellular cytoskeleton and the endomysium) results in a highly pronounced susceptibility to contraction-induced injuries, leading to highly accelerated successive rounds of fibre tissue degeneration and regeneration cycles, compared to normal non-DMD tissues- where satellite cells are activated to replace lost tissue. This is exacerbated by the impaired regeneration of satellite cells, which eventually get exhausted (Blau, Webster and Pavlath, 1983; Fukada *et al.*, 2013). Over time, damaged skeletal muscle tissue in DMD patients can no longer be adequately regenerated, and progressively becomes substituted with fibrotic and adipose tissue (Mann *et al.*, 2011). Chronic inflammation of skeletal muscles in DMD patients is a hallmark of the disease that plays a crucial role in its progression. Necrotising myofibres attract pro-

inflammatory and pro-regenerative macrophages that constantly release cytokines that further foster muscle damage (De Paepe and De Bleecker, 2013).

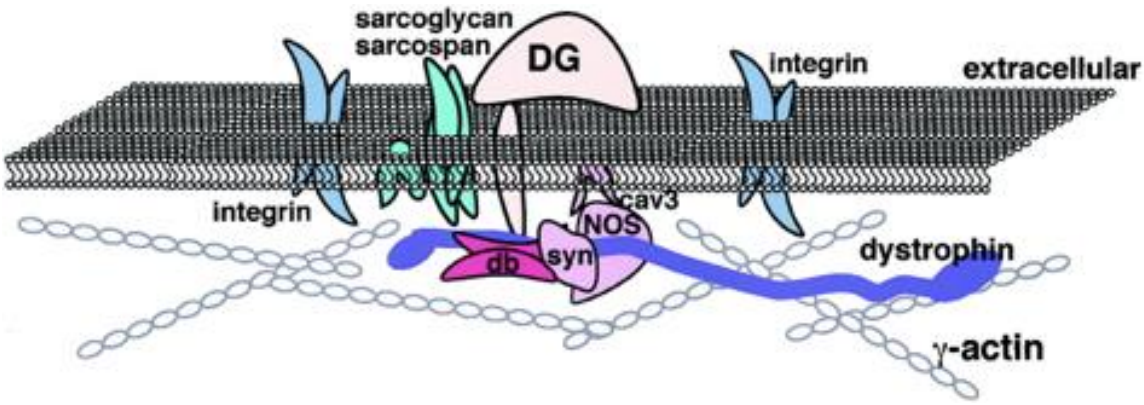

Figure 1-4. *Illustration showing the DGC at cell membrane.*
The DGC is composed of dystrophin (blue) that binds to the tail of γ-actin (white) and dystroglycan (orange). Dystroglycan binds to the extracellular matrix protein laminin. The sarcoglycan complex is in blue-green. Additional subcomplexes in the DGC in skeletal muscle include dystrobrevin, the syntrophins, nNOS, and caveolin 3 (pink). The major α/β integrin subunits (gray) are not part of the DGC. Adapted from (Lapidos, Kakkar and Mcnally, 2004).

1.3 OVERVIEW OF DMD THERAPIES

1.3.1 STANDARDS OF CARE AND APPROVED THERAPIES

1.3.1.1 Standards of Care

Currently, there are no cures for DMD. The standard of care includes physiotherapy, orthotics and corrective surgery to address some symptoms. In addition, the use of glucocorticoids is widely accepted as a treatment and has been shown to delay loss of ambulation for approximately 2 years (Manzur, Kinali

and Muntoni, 2008; Bushby *et al.*, 2010; Matthews *et al.*, 2016). However, corticosteroids are associated with adverse effects and current approaches do not address the underlying genetic cause of the disease (Matthews *et al.*, 2016).

1.3.1.2 Readthrough Technology

Ataluren is a drug approved in the European Union that induces selective ribosomal read-through of premature termination codons (not affecting normal stop codons) caused by non-sense mutations (Welch *et al.*, 2007). This approach is relevant to approximately 15% of DMD patients. Ataluren is the first treatment approved for the underlying cause of DMD by both the European Medicines Agency (EMA) and as an orphan medicinal product following clinical trials that attested to the efficacy and safety of the compound in slowing down the disease progression (Bushby *et al.*, 2014). Further trials are needed to assess the long-term benefits of the drug (McDonald *et al.*, 2017).

1.3.1.3 Antisense Oligonucleotide (AON) Mediated Exon Skipping

AONs encompass a group/class of molecules of different chemistries that are 15 to 30 nucleotides long which bind to mRNA or pre-mRNA to either promote degradation of targeted mRNA through the recruitment of enzymes such as RNAse H or Argonaute 2 (Shen and Corey, 2018), or through the exploitation of RNA splicing. It is through the exploitation of the second approach that researchers have developed AON-mediated exon skipping therapies. AONs bind to the target sequence on pre-mRNA, masking it, and disrupts the splicing process through sterically blocking the polymerase from binding to the pre-

mRNA, making the splicing sites to induce an exclusion, or an extra inclusion, of an exon (Levin, 2019). Both approaches, inclusion and exclusion of exons, are utilisable in therapeutic applications, the former has been exploited in the treatment of spinal muscular atrophy (SMA), by the drug Nusinersen (https://bnf.nice.org.uk/drugs/nusinersen/). SMA patients have a dysfunctional SMN1 protein, encoded by a mutated *SMN1* gene. The gene *SMN2* is almost identical to *SMN1* but codes for a slightly different protein, SMN2, which is shorter compared to SMN1, lacking amino acids encoded by exactly 1 exon and is of general low expression in the cell and normally degraded. Nusinersen binds to the SMN2 pre-mRNA, masking the exon from excision, and leading to an mRNA coding for a more functional SMN1-like SMN2 protein (Hoy, 2017).

In contrast, the exclusion approach leads to the skipping of the transcription of mutated exons. This is of particular relevance to DMD as most mutations are out of frame and lead to premature stop codons, giving rise to a truncated, dysfunctional protein lacking the c-terminus end that binds to the dystrophin-associated dystroglycan complex (Lapidos, Kakkar and Mcnally, 2004). This exclusion approach is known as non-sense mediated RNA degradation (Hug, Longman and Cáceres, 2016; Moon *et al.*, 2017). The AONs designed can target the masking one of three sites involved in splicing, the donor, acceptor, or enhancer splice sites, preventing the transcription of the target exon. This strategy is only viable if the exons that flank the skipped exon give rise to an in-frame transcript, leading to a centrally truncated but mostly functional protein (depending on the length and functional groups in the skipped sequence). For

DMD, multiple AON drugs are in clinical trials as therapies to skip various exons. This results in a shorter, yet still functional, dystrophin protein that may resemble those present in Becker muscular dystrophy (Cirak *et al.*, 2011). Exon skipping has been successful *in vitro* and *in vivo* (reviewed in (Benedetti, Hoshiya and Tedesco, 2013)) and could potentially treat 83% of DMD patients (Aartsma-Rus *et al.*, 2009). AONs have been developed in different chemistries, one being 2′-O methyl phosphorothioate (2′OMePS) and the other being phosphorodiamidate morpholino oligomers (PMO).

Two drugs targeting exon 51 skipping (applicable to approximately 13% of DMD patients) have also been developed: Drisapersen (2′OMePS) and Eteplirsen (PMO). Drisapersen has not been approved for use because of safety concerns that limit dosing and led to the failure in meeting the primary endpoint in phase III clinical trial. Eteplirsen by Sarepta Therapeutics, which is an FDA-approved drug, relies on skipping exon 51, which in patients with exon 50 or 52 deletion, would restore the reading frame by having exon 49 binding to exon 52 or 50 to 53, respectively, some other longer deletions are also amenable to exon 51 skipping therapies (Cirak *et al.*, 2011).

Golodirsen, another AON therapy by Sarepta Therapeutics, which has been more recently approved by the FDA, targets exon 53 for skipping. This approach leads to a shorter, truncated in the rod-domain of the protein, novel ‘isoform’ of dystrophin that is still functional and can eventually ameliorate some of the disease phenotype in patients (Cirak *et al.*, 2011).

Both Eteplirsen and Golodirsen are of the second-generation class of AONs, they are phosphorodiamidate morpholino oligomers (PMOs), that do not possess the natural phosphoribose backbone as a morpholine ring replaces the deoxyribose moiety, and an uncharged phosphorodiamidate linkage replaces the charged phosphodiester inter-subunit linkage (Summerton *et al.*, 1997). These modifications make them very resistant to nuclease and protease degradation (Hudziak *et al.*, 1996).

AONs are also promising because targeting of an additional nine exons could achieve restoration of dystrophin in roughly 70% of patients who carry out-of-frame deletions (Mercuri and Muntoni, 2013). However, some limitations need to be overcome to achieve the full potential of AON-mediated treatments, including: variable tissue uptake, transient effect requiring repeated injections, insufficient dystrophin rescue in cardiac muscle and central nervous system, and the fact that the restoration of dystrophin expression after 1 year treatment is low (0.22%-0.32% of normal levels) (Fairclough, Wood and Davies, 2013; Guncay and Yokota, 2015; Tabebordbar *et al.*, 2016; Amoasii *et al.*, 2018).

The two hurdles in AON-mediated therapies are the stability and delivery of the molecule, directly affecting the efficiency, bioavailability, cell uptake and target specificity of the molecules. As mentioned earlier, different chemistries of the backbone and chemical modifications can be used to address both issues. As for delivery and cellular uptake, commonly recognised as the major barrier, and studies have shown the nuclear uptake is not the rate limiting step for AON activities, but rather uptake into the cytoplasm and release from endosomes

(Hartig *et al.*, 1998; Lorenz *et al.*, 2000). In the past years, great progress has been made in employing polymer- and lipid-based nanocarriers to facilitate AON delivery (Falzarano, Passarelli and Ferlini, 2014). Neutrally charged backbones, including PMOs, conjugated to short cationic peptides known as cell-penetrating peptides (which have strong transmembrane capacity) have shown great potential to treat neurodegenerative disorders (Lehto *et al.*, 2016). A novel series of molecules, peptide nucleic acids, also known as PMO internalisation peptides (Pips) have been recently recognised for their increased efficacy and more favourable safety profile over other alternatives. They are structurally modified forms of the R6-Penetratin peptide derived from *Drosophila melanogaster* (Ivanova *et al.*, 2008). These peptide conjugated PMOs (PPMOs) have been shown to effectively restore therapeutic levels of dystrophin in skeletal and also cardiac tissue in animal models (Jearawiriyapaisarn *et al.*, 2008; Wu *et al.*, 2008; Yin *et al.*, 2011).

1.3.2 POTENTIAL THERAPIES

With recent advancements in biomedical research, multiple new approaches for treatment of neuromuscular genetic diseases like DMD including new drugs, gene therapies and editing, and cell therapy have been developed. The goal for all approaches is to restore the function of dystrophin, whether through modulating other proteins with a similar function, “fixing” the endogenous DMD gene, delivering a new copy of DMD or delivering cells bearing a functional copy of DMD (Benedetti, Hoshiya and Tedesco, 2013).

Regardless of the approach, there are two main hurdles any new therapy needs to overcome before being a viable treatment for DMD. First is the size of the dystrophin gene, which is relevant to all therapies with the exception of cell therapies (autologous and HLA matched). Secondly, the systemic delivery of the therapeutic product, as the vast majority of skeletal muscles are affected in DMD.

1.3.2.1 Modulation of Other Cellular Pathways

The modulation and pharmacological upregulation of cellular pathways/proteins that would compensate for the lack of dystrophin is an approach under investigation. One of the target proteins using this approach for DMD is utrophin which has an 80% amino acid sequence identity match to the dystrophin gene (Love *et al.*, 1989). Both proteins bind actin filaments in the cell membrane, albeit with different modes of contact (Rybakova *et al.*, 2006). Utrophin is more ubiquitously expressed than dystrophin in adults, and mainly found in the sarcolemma of developing muscle before birth (Tomé *et al.*, 1994). Utrophin has been discovered to be increased in the skeletal muscles of DMD patients, with higher rates being positively correlated to the severity of DMD manifestation in the patients (Kleopa *et al.*, 2006). Thus, it is postulated the utrophin is a potential surrogate that can compensate the lack of dystrophin in dystrophic muscles.

The drug Ezutromid, developed by Summit LLC., upregulates Utrophin through the Utrophin A promoter (Tinsley *et al.*, 2011), has underwent clinical trials, after initial promising results. However, in phase 2 of the trials, it failed to

meet any of the primary or secondary endpoints and therefore, investigations were discontinued (Summit Therapeutics, 2020).

Another target to compensate for the lack of dystrophin is Biglycan, a small extracellular matrix (ECM) protein, that regulates expression of some of the key cell surface proteins including Utrophin (Amenta *et al.*, 2011), (Mercado *et al.*, 2006). Ubiquitously expressed in the body, in muscle it is expressed in non-glycanated form, which is a member of the dystrophin/utrophin associated protein complex, TVN-102 by Tivorsan is derived from that (granted FDA orphan drug approval in 2016)- still in preclinical research stage.

1.3.2.2 Cell Therapies

Finally, cell-based approaches to treat DMD have been under investigation since the 1980s, with a transplantation study of healthy donor myoblasts in the muscles of *mdx* mouse model showing robust engraftment and rescue of dystrophin protein production (Partridge *et al.*, 1989). This was followed by a wave of research investigating the efficacy of myoblast intramuscular injections in DMD patients. However, the results were disappointing owing to multiple issues including host immune response, rapid cell death and poor cell migration (Gussoni *et al.*, 1992; Karpati *et al.*, 1993; Vilquin *et al.*, 1994). Even after attempts of delivery under immunosuppressive regiments and high-density injections (Morgan and Partridge, 1992; Vilquin *et al.*, 1994; Skuk *et al.*, 2007) but results were still suboptimal (Skuk and Tremblay, 2015). One of the most

recent clinical trials utilises mesoangioblasts from HLA-matched donors in patients with DMD, and while no major adverse effects have been observed, the results show very little engraftment and minimal efficacy in reducing disease phenotype (Cossu *et al.*, 2015).

Currently, research is ongoing for discovering an engraftable myogenic cell population for widespread muscular disorders and research into bottlenecks such as migration is underway (Gerli *et al.*, 2019; Ferrari *et al.*, 2021).

While systemic delivery has been a major barrier to adoption and use of cell therapies for muscular dystrophies, dystrophies with localised and/or limited affected muscles can benefit from intramuscular injections of myogenic cells. A recent clinical trial utilised myoblast transplantation in patients oculopharengyeal muscular dystrophy in a PhaseI/IIa study demonstrated safety and improved quality of life score for all 12 participating patients with no functional degradation observed in 10 of the patients (Périé *et al.*, 2014).

1.3.2.3 Gene therapies

Gene therapies aiming to restore expression of the endogenous dystrophin can be broadly divided into two groups. The first group of therapies relies on delivering a copy of the dystrophin gene, which often is a truncated version. The second group is therapies relying on the endogenous copy of the gene to produce the protein, such as stop-codon read-through technologies, acting at a translational level, antisense oligonucleotide (AON)-mediated exon

skipping, that acts a transcriptional level, discussed in 1.3.1.3, and gene-editing that acts at genomic level.

1.3.2.3.1 Gene Replacement Therapy

In gene replacement therapies researchers aim to deliver the entire dystrophin sequence (or a miniaturised form of it) through viral or non-viral vectors (Benedetti, Hoshiya and Tedesco, 2013). Due to the size of dystrophin cDNA (14Kb) and the limited capacity of viral vectors, truncated versions of the dystrophin gene containing the functional domains have been developed and delivered in adeno-associated viruses (AAVs) in preclinical models of DMD (Scott *et al.*, 2002). Other forms of micro- and mini- dystrophins have been developed and showed efficacy in restoring dystrophin expression in dogs and non-human primates (Yue *et al.*, 2008; Rodino-Klapac *et al.*, 2010; Koo *et al.*, 2011; Wang *et al.*, 2012). AAVs are very promising being non-integrative, having high infectivity that is independent of proliferative state, and their specificity infecting muscle cells (AAV9) (Kouprina *et al.*, 2014); however, they have some drawbacks, including: not efficiently targeting quiescent SCs (Arnett *et al.*, 2014) and severe immune response against the viral capsid and the mini-dystrophin (Mendell *et al.*, 2010; Bowles *et al.*, 2012).

Lentiviral vectors have been investigated for their pros over AAVs that include: infectivity being independent from tissue type so they can target SCs (Kobinger *et al.*, 2003) and larger cloning capacity (7.5 kb vs. 4.5 kb for AAVs). However, they have a high risk of insertional mutagenesis (Colella, Ronzitti and Mingozzi, 2018). Non-viral delivery vectors are mainly lipid, polymer or

polyethelyne glycol (PEG) containing nanoparticles (Rui, Wilson and Green, 2019). Therefore, they are not considered candidates for use in endogenous in vivo applications, however, they have gained popularity in *ex vivo* applications, such as CAR-T cell therapies (Morgan *et al.*, 2020). For DMD, their potential is limited to use in conjunction with a cell therapy approach.

Preliminary results from the most recent clinical trials employing AAV mediated delivery of micro-dystrophins have shown confirmed expression in over 80% of muscle fibres in muscle biopsies at levels of over 60% (Verhaart and Aartsma-Rus, 2019; Mendell *et al.*, 2020). With a further trial using different versions of micro-dystrophin cDNA utilising multiple AAV serotypes at very high doses currently ongoing (Duan, 2018).

1.3.2.3.2 Targeted Gene Editing

Targeted gene editing is an attractive option for DMD therapy as it permanently corrects the dystrophin gene. Many gene editing approaches exist under the umbrella of endonucleases such as meganucleases, zinc finger endonucleases (ZFN) and transcription-activator like effector nucleases (TALENs) and the most recent (and most exciting) clustered regularly interspersed short palindromic repeats (CRISPR/Cas9). All of these approaches are capable of targeting specific mutations to restore the gene sequence utilising homologous directed repair (HDR) and non-homologous end joining repair and to deliver systemically, viruses such as adeno-associated viruses and lentivirus are used (Benedetti, Hoshiya and Tedesco, 2013).

Multiple aspects of DMD make gene editing a viable therapeutic approach (Lim, Yoon and Yokota, 2018; Zhang *et al.*, 2018; Min, Bassel-Duby and Olson, 2019; Young, Pyle and Spencer, 2019; Chemello, Bassel-Duby and Olson, 2020). The first being the modular structure of the protein in which the central rod domain has redundancies, allowing deletion of segments of the gene that include loss-off-function causing mutations and restoring the reading frame, generating BMD-like dystrophin forms. The second being the location of dystrophin on the X chromosome, meaning there is a single allele to be edited. Finally, only a small percentage of the normal level of dystrophin is required to confer significant therapeutic benefit (Neri *et al.*, 2007; Godfrey *et al.*, 2015; Gentil *et al.*, 2016). The MDX mouse model has been used to demonstrate that only 15% of gene editing is required for dystrophin expression to reach wild type levels (Long *et al.*, 2014; H. L. Li *et al.*, 2015; Zhang *et al.*, 2017).

Many other groups reported using CRISPR-Cas 9 mediated restoration of dystrophin expression in mice (Nelson *et al.*, 2016; Tabebordbar *et al.*, 2016; Bengtsson *et al.*, 2017). The canine model has also been used in such studies. A notable example is the DMD dog model that closely resembles human DMD muscle pathology, fibrosis, atrophy and loss of ambulation, harbouring a missense mutation in the upstream splice site of exon 50 causing a premature termination of dystrophin translation as a result of a nonsense codon arising upon splicing of exon 49 to exon 51 (Hildyard *et al.*, 2018). The dogs were subjected to two intravenous infusions at 1 month of age of AAV9 containing the Cas9 and another containing the sgRNA targeting permanent skipping of exon 51 and their

muscles were analysed 8 weeks post-treatment (Amoasii *et al.*, 2018). Results of western blot analyses detected dystrophin levels between 3 and 90% of normal dystrophin levels in different skeletal muscles and up to 92% of normal levels in cardiac tissue (Amoasii *et al.*, 2018). Initial studies in a porcine DMD model, lacking exon 52, showed restoration of dystrophin expression preventing muscle pathology and cardiac arrhythmias within 4 months of treatment with AAV9 delivered Cas9 and a pair of sgRNA flanking exon 51 (Moretti *et al.*, 2020).

1.4 CURRENT MODELS OF DMD

1.4.1 ANIMAL MODELS

Current DMD animal models have contributed to massive developments in understanding the disease pathophysiology and in therapy development, however, they are limited and have a several disadvantages. The most widely used mouse model is the *mdx* mouse; it has a nonsense mutation in exon 23 leading to an early stop codon (Sicinski *et al.*, 1989). The *mdx* mouse is very popular because of its fertility, near-normal lifespan and availability, which meant that protocols, tools, standard operating procedures and outcome measures are well established and mostly apply to other murine models. However, the major drawback is that most exhibit a mild phenotype of the disease. Other dystrophin-deficient mice have been generated, they also do not reproduce the full clinical

spectrum of the human disease (Chamberlain *et al.*, 2007; McGreevy *et al.*, 2015). Approaches have been developed to improve these models such as knocking-out genes coding for proteins such as myoD and cytoskeleton-ECM interaction genes generating double knock-out symptomatic mice, however, they are often hard to breed, not commercially available and more importantly how the extra mutations these models possess affect other aspects of their physiology and thus influencing data interpretation is still unknown (McGreevy *et al.*, 2015). In efforts to better elucidate dystrophin expression *in vivo* and especially in longitudinal studies, Petkova and colleagues created DMD-EGFP reporter *mdx* mice and demonstrated their use in visualising AAV-mediated exon skipping (CRISPR-Cas9 DMD exon 23 excision) restoration of dystrophin expression (Petkova *et al.*, 2016).

More recently, researchers started using rat models of DMD. Unlike the *mdx* mouse, these rats are engineered using genome modification techniques. They are, compared to larger animals, easy to breed, and less expensive to maintain. A rat model with an exon 23 mutation was developed by Larcher and colleagues using TALENs, and in contrast to its equivalent, exhibits a more pronounced phenotype of the disease ranging from severe fibrosis and adipose tissue infiltration in muscles to cardiac impaction and altered diastolic function, and finally, brain abnormalities (Larcher *et al.*, 2014; Caudal *et al.*, 2020; Xu *et al.*, 2020). Nakamura and colleagues used CRISPR Cas9 technology to produce a range of mutations in exons 3 and 16 of the DMD gene (Nakamura *et al.*, 2015). Dystrophic rats exhibited a similar disease phenotype as exon 23 deletion rats;

they displayed increased fibrosis in cardiac tissue and at an earlier age than the *mdx* mouse. A DMD exon 52 deleted rat model, R-DMDdel52 rats, has been recently developed using CRISPR/Cas9 technology and was used in modelling human disease pathophysiology more faithfully than the *mdx* mouse model, as they exhibited progressive and severe skeletal muscle loss associated with fibrotic deposition, fat infiltration and fibre type switch (Taglietti *et al.*, 2022). The group used the model to study the histological modifications that lead to premature death, and used RNAseq analysis to elucidate the cellular and molecular mechanisms underlying them to propose a novel biomarker to be used in human clinical trials (Taglietti *et al.*, 2022).

While they are better than mouse models in recapitulating the disease phenotype, and are more suitable for use in behavioural studies, no studies so far have investigated the use of DMD rats in development of therapies.

Another widely animal is the canine model. Dogs are better than rodent models in addressing issues such as immune response and dosage of therapies. They have been described in the literature for over 50 years with dystrophin deficiency being confirmed in over 20 dog breeds, whose clinical phenotype is more severe than *mdx* mice and , therefore, regarded as a better model of human DMD (McGreevy *et al.*, 2015). Canine models have not been studied in depth, mostly just case reports. Experimental colonies have been established for a few breeds, the most common of which is the golden retriever muscular dystrophy dogs with a DMD intron 6 mutation (McGreevy *et al.*, 2015). More recently a beagle model with an exon 50 DMD deletion, which is in the human mutation

hotspot of DMD, and is amenable to exon 51 skipping, was developed and used in a proof of concept AAV-mediated CRISPR-Cas9 gene-editing therapy (Amoasii *et al.*, 2018). DMD dogs are remarkably similar to human DMD patients, possessing a similar clinical course (Shimatsu *et al.*, 2005; Smith, 2011). Like humans, canine DMD models show variations in their symptoms (Zatz *et al.*, 2014) and similar histological lesions (Yang *et al.*, 2012). However, there are disparities between the clinical progression of canine models and humans. Firstly, 20/30% of puppies die within 2 weeks of birth, most likely due to diaphragm failure which is not observed in DMD boys (Nakamura *et al.*, 2015). Growth retardation is only observed in canine models (West *et al.*, 2013). Lastly, unlike DMD patients, ambulation loss is not a clinical feature of young DMD dogs (Duan *et al.*, 2015).

Animal models have proven essential in early research into DMD; however, they cannot meet all the needs of DMD drug discovery due to their limitations. They don't accurately replicate human disease progression or severity and do not reflect the impact of genetic diversity in patients. It is also worth mentioning that almost 90% of tested drugs fail to gain approval due to differential toxicity responses between animals and humans (Hay *et al.*, 2014).

1.4.2 IN VITRO MODELS

For the reasons mentioned above, and the need for models and platforms for high-throughput drug screens, developing patient-derived, specific, human

DMD models was necessary. These fall under two categories, primary cell culture from muscle biopsies and differentiated iPS cell-derived muscle cell cultures.

1.4.2.1 Monolayer Patient-Derived Cell Cultures

One source of cells for such models are muscle biopsy-derived myoblasts; however, evidence shows that endogenous myogenic progenitors are depleted and dysfunctional in DMD patients, due to the constant regenerative demands placed on these stem cell pools as a result of the degeneration-regeneration cycles that occur in dystrophic muscles and defects their self-renewal potential (Blau, Webster and Pavlath, 1983; Sacco *et al.*, 2010; Tedesco *et al.*, 2012; Dumont *et al.*, 2015).

Therefore, patient-derived satellite cells (SCs) would be unable to meet the proliferative demands that would be required of them to deliver macroscopic skeletal muscle organoid (Fatehullah, Tan and Barker, 2016).

This hurdle, in addition to the invasiveness of sampling to harvest SCs, can be overcome through the use of patient-derived immortalised myoblasts, generated through transduction with cyclin-dependent-kinase 4 (CDK4) and human telomerase reverse transcriptase (hTERT)-expressing vectors to generate myoblast cell lines with enhanced proliferative capacity (Nguyen and Yokota, 2017). These cell lines have been demonstrated to reliably differentiate and fuse into skeletal muscle cells and have been used in various settings, in particular, in preclinical screening of AONs for exon 51 and exon 53 skipping (Mamchaoui *et al.*, 2011).

Despite their demonstrated potential, immortalised myoblast cell lines are limited by their limited differentiation potential to skeletal muscle cells only, in contrast to iPS cells that can differentiated into any cell lineage. This is especially relevant to DMD research given the involvement of various cell types and tissues in the disease.

1.4.2.2 Monolayer iPS Cell-derived Cultures

1.4.2.2.1 Induced Pluripotent Stem (iPS) Cells

After the work in 1981 obtaining pluripotent stem cells from the inner cell mass of mouse blastocysts (Evans and Kaufman, 1981; Martin, 1981) and the success in obtaining embryonic stem (ES) cells from human embryos in 1998 (Thomson *et al.*, 1998) and establishing their potential, scientists were interested in recapitulating/obtaining cells with similar properties from adult cells, especially given the ethical considerations. The breakthrough came in 2006, when Takahashi and Yamanaka first demonstrated that the expression of the four transcription factors octamer-binding factor 4 (*OCT4*), Kruppel-like factor 4 (*KLF4*), sex determining region Y-box 2 (*SOX2*) and c-Myc (*OSKM*) stimulates murine and human fibroblasts to reacquire an embryonic stem cell-like pluripotent state; these cells are what are now called induced pluripotent stem cells, iPS cells (Takahashi and Yamanaka, 2006; Takahashi *et al.*, 2007).

iPS cells have a theoretically unlimited expansion and self-renewal capacity, in addition to their potential of differentiating into any cell type (Yamanaka, 2009). This is particularly beneficial as some cell types, such as

neurons are particularly hard to obtain yet this is possible through iPS cell differentiation (Russo *et al.*, 2015). For the aforementioned reasons, iPS cells have great potential for disease modeling, cell therapy, regenerative medicine and high-throughput drug screening (Bellin *et al.*, 2012).

1.4.2.2.2 Derivation of Myogenic Cells from iPS Cells

There are multiple methods and protocols to derive skeletal muscle cells from iPS cells, which differ in length, efficiency, purity of final cell population, and intermediate steps (sorting-mediated enrichment, transgene expression, cell culture technique, etc.).

1.4.2.2.2.1 Transgene-based Approaches

The first approach is transgene-based, which encompasses methods that rely on the expression of muscle-specific factors such as Paired box 7 (PAX7), Myogenic Factor 5 (MYF5) and Myoblast Determination protein 1 (MYOD). These methods, have been developed following extensive work deriving myogenic cells from murine and human Embryonic Stem Cell (hESC) (Loperfido *et al.*, 2015).

Expression of *MYOD* in human iPS cells that successfully led to myotube formation was first demonstrated in 2012 by Tedesco and colleagues followed by Goudenege and colleagues (Goudenege *et al.*, 2012; Tedesco *et al.*, 2012), with further studies published in 2013/2014 (Tanaka *et al.*, 2013; Abujarour *et al.*, 2014; Yasuno *et al.*, 2014). In both (Tanaka *et al.*, 2013) and (Yasuno *et al.*, 2014) a doxycycline-inducible *MYOD* was delivered into iPS cells by means of

piggyback transposition. Over expression of *MYOD* directly in iPS cells induced myogenic differentiation *in vitro;* with data form Tanaka and colleagues suggesting that upon transplantation of iPS cell-derived myogenic cells, *in vivo* engraftment occurs. Abujarour and colleagues demonstrated that direct myogenic fate induction via exogenous expression of *MYOD* in iPS cells is feasible, they specifically used a lentiviral system expressing *MYOD* driven by a tet-inducible promoter (Abujarour *et al.*, 2014).

Work in our laboratory led to the development of a protocol that was able to convert both hESCs and iPS cells into highly expandable myogenic progenitors that terminally differentiate upon tamoxifen-mediated induction of *MYOD* expression (lentiviral system), the four-stage protocol of only monolayer cultures with different conditions (cell densities and media composition) yields the highly myogenic progenitors in approximately 22 days and fully differentiated myotubes in a further 5-8 days (Maffioletti *et al.*, 2015).

1.4.2.2.2.2 Transgene-free Approaches

The next aim in the field of skeletal muscle derivation from iPS cells was to use methods that rely on small molecules and transcription factors that simulate developmental cues orchestrating the transition through presomitic mesoderm, somite and dermomyotome stages to form the myotome. Small molecule-based differentiation protocols are non-mutagenic by nature and can act as a clinically relevant standardised platform as it facilitates the use of animal component-free media relevant standardised platform.

Initial strategies for small molecule-based differentiation of myogenic progenitors from iPS cells involved initial formation of embryoid bodies (EB). One of the first published methods developed a 63-day stepwise culture protocol that entailed an initial embryoid body (EB) formation step, followed by spontaneous differentiation into the mesenchymal lineage and then serial media changes until terminal differentiation (Awaya *et al.*, 2012). Following that, two papers by (Xu, Du and Deng, 2015) and (Borchin, Chen and Barberi, 2013) reported results with another method/protocol that was also EB mediated. They showed that the inhibition of GSK3β and activation of cAMP and βFGF pathways during EB formation are sufficient to promote iPS cells commitment towards the skeletal muscle lineage, then a purification step is needed to get rid of the non-myogenic population by selecting acetylcholine receptor (AChR)-positive, C-X-C chemokine receptor type 4 (CXCR4)-positive and hepatocyte growth factor receptor (HGF/C-MET) and human natural killer-1 (HNK-1)-negative [AChR +/ CXCR4 +/ C-MET +/ HNK-1 -) cells as these represent highly myogenic *(in vitro*) Pax3+/Pax7+ satellite-like cells (Borchin, Chen and Barberi, 2013). The step of creating and maintaining EBs is often very sensitive and therefore other methods that rely entirely on monolayer cultures are becoming more common.

One of those monolayer based studies reported generation of SC-like cells by recapitulating molecular signals that match the stages of development (based on a transcriptomic analysis of mouse muscle development) (Chal *et al.*, 2015, 2016). Human iPS cells are pushed towards commitment to the paraxial mesoderm fate via GSK3 and bone morphogenic proteins (BMP)-inhibition and

growth factors (insulin-like growth factor (IGF), fibroblast growth factors (FGF), hepatocyte growth factor (HGF)) are used to expand myogenic cells in a 40-day serum-free protocol. This expandable population is then sub-cultured stepwise to enrich for the myogenic fraction. The final obtained population is not homogenous, it has been later found to be a heterogenous population containing myogenic (some PAX7+ cells) and ectoderm contaminants (Kim *et al.*, 2017).

Following that, in 2016, Choi and colleagues (Choi *et al.*, 2016) used a GSK-3β inhibitor followed by a Notch-inhibitor (γ-secretase inhibitor: DAPT) and a final differentiation step in N2 media obtaining multinucleated contracting myotubes from human iPS cells (notably this was derived from healthy and DMD lines and tested *in vivo* and *in vitro*). An optional step of further sorting (for the NCAM+/ HNK1- fraction) is also included (Choi *et al.*, 2016).

Shortly after, a similar small molecule-based protocol was published that entails, first, deriving PAX3 and PAX7 expressing cell population, then committing this population into the myogenic lineage through high MyoD and Myf5 expression and finally, terminally differentiating into multinucleated skeletal muscle myotubes (without any sorting steps) (Caron *et al.*, 2016). Obtained cells have not been assessed *in vivo* yet; however, the media required for this protocol has been developed and produced in a good manufacturing practice (GMP) manner and is commercially available (*Skeletal muscle differentiation kit, SKM-KIT* | *AMSBIO*) owing to its ease and simplicity.

In efforts to reduce heterogeneity of the myogenic population in the previous methods, many groups have worked on developing isolation and

purification strategies. Myogenic progenitor population specific surface markers integrin α9β1, CD54 and SDC2 have been identified through whole transcriptome analysis of myogenic populations derived from human iPS cell using doxycycline inducible PAX7-mediated derivation by Magli and colleagues (Magli *et al.*, 2017).

Xi and colleagues performed transcriptomic analysis of presomitic and somitic mesoderm development in human embryos and identified WNT-β-catenin, BMP and transforming growth factor β (TGF β) as crucial regulators of this process (Xi *et al.*, 2017). By modulation of these signalling pathways, they were able to derive somitic cells that were found to be multipotent from iPS cells (able to differentiate into skeletal myocytes, osteocytes and chondrocytes); and their expression profile faithfully matched that of cells undergoing somitogenesis *in vivo* (Xi *et al.*, 2017).

More recently, (Hicks *et al.*, 2018) demonstrated that receptor tyrosine-protein kinase erbB-3, encoded by the ERBB3 (a muscle specific receptor) and nerve growth factor receptor NGFR/CD271 as markers of Pax7-enriched foetal muscle progenitor cells through RNA-sequencing of foetal and iPS cell-derived skeletal myogenic progenitors. Hicks et al. utilised these markers to sort iPS cell-derived myogenic cells, which showed increased myogenic potential *in vitro* and consistent *in vivo* engraftment comparable to foetal myoblasts. In this study, they employed the two protocols published by (Shelton *et al.*, 2014) and (Chal *et al.*, 2016) and pointed out that differentiation is not affected by the absence of dystrophin, contradictory to what has been published by (Choi *et al.*, 2016).

1.4.2.2.3 Advantages of iPS Cell Derived Myogenic Cells

Cultures derived from iPS cells have been used to investigate different therapeutic approaches for DMD. As a proof-of-concept Kazuki and colleagues showed that *mdx* mouse and patient derived iPS cells that are subjected to microcell-mediated chromosome transfer (MMCT) with a human artificial chromosome (HAC) containing a complete human genomic dystrophin sequence (DYS-HAC) can differentiate into different germ layers and human dystrophin was detected in muscle like tissue (Kazuki *et al.*, 2010). Zhao and colleagues derived iPS cells from *mdx* mouse fibroblasts and inserted a full-length dystrophin cDNA with bxb1 integrase in a site-specific manner. They demonstrated that the iPS cell-derived myogenic progenitors maintained their myogenic differentiation potential (Zhao *et al.*, 2014). Shoji and colleagues suggest that DMD patient derived iPS cells, with deletion in exon 44 and 46-47, can be used to evaluate antisense oligonucleotide therapies with robust outcome measures. They first showed that electric stimulation of the iPS cell-derived myotubes caused pronounced calcium ion (Ca2+) influx in comparison to control conditions. They tested ASO AO88, which causes the skipping of exon 45, and restoration of dystrophin suppressed calcium overflow and reduced creatine kinase (CK) in DMD iPS cell-derived myotubes (Shoji *et al.*, 2015). Li and colleagues compared 3 different approaches to genetically correct DMD patient derived iPS cells with exon 44 deletion using gene-editing therapies (TALENs and CRISPR Cas9); first the disruption of splicing acceptor site of exon 45 to restore reading frame, secondly introducing tiny indels to cause a frameshift and finally knocking in the

entire length of exon 44 to restore the full genetic code of the gene. All approaches lead to fully restored dystrophin in differentiated iPS cell-derived cells (H. L. Li *et al.*, 2015). These studies demonstrate potential for autologous cell therapy using corrected patient iPS cells if derivation and delivery hurdles are overcome.

The rise of iPS cells and advancements in their derivation, expansion and differentiation has provided a platform for studying patient derived cells without the expansion constraints of primary cultures. They also facilitate research on DMD in early stages of differentiation and different tissues and cell types that would otherwise be difficult to obtain from patients using a biopsy such as brain and cardiac tissue. Most importantly, patient-derived iPS cells enable research on all different mutations and are faster and cheaper to derive and maintain than creating humanised DMD animal models with different mutations. However, as with primary monolayer cultures, detachment of myotubes during maturation prevents long-term drug investigations and thus decreases physiological relevance (Funanage, Smith and Minnich, 1992).

1.4.3 Novel Models

1.4.3.1 Artificial Muscles and Organoid-Like 3D Cultures

The development of artificial skeletal muscles, 3D cultures of skeletal muscle cells and other types of cells normally present in the muscle tissue, would provide an invaluable tool that will allow us to more faithfully test potential therapeutics, develop tissue replacement protocols and study pathological

mechanisms (Maffioletti *et al.*, 2018). The use of similar approaches (3D cultures) in other tissues has been transformational on the drug development and regenerative medicine research front (Lancaster and Knoblich, 2014; Fatehullah, Tan and Barker, 2016).

Several groups have attempted developing skeletal muscle tissues, however, most of these were conducted on rodent cells (Maffioletti et al., 2018). Some other groups/studies used human cells, however, most of them employed primary human cells from muscle biopsies (Maffioletti et al., 2018), which have great drawbacks that were listed earlier in 1.4.2.2.

1.4.3.1.1 Self-Organising Organoids

The recent advancements in stem cell technologies have led researchers to investigate more physiologically faithful models, moving from monolayer cultures to three-dimensional cultures where different cell types differentiate from a common precursor (Brassard and Lutolf, 2019). Self-organising *in vitro* tissue models known as organoids have been developed for a wide range of tissues using adult stem cells and pluripotent stems. Models using pluripotent stem cells as a starting population include the intestine, pancreas, kidney, liver, spinal cord, brain and retina (Spence *et al.*, 2011; Eiraku and Sasai, 2012; Lancaster *et al.*, 2013; Huch *et al.*, 2015; Morizane *et al.*, 2015; Pasca *et al.*, 2015; Broutier *et al.*, 2016).

Martins and colleagues have developed neuromuscular organoids from human PS cells that contain self-organising spinal cord neurons and skeletal

muscle cells that can be maintained in culture for several months (Faustino Martins *et al.*, 2020). Remarkably, they contain functional neuromuscular junctions supported by Schwann cells, in addition to being contractile and develop central pattern generator-like neuronal circuits. However, their potential to model diseases using patient derived iPS cells and as a platform for screening therapeutic has not been investigated yet.

1.4.3.1.2 Engineered Platforms

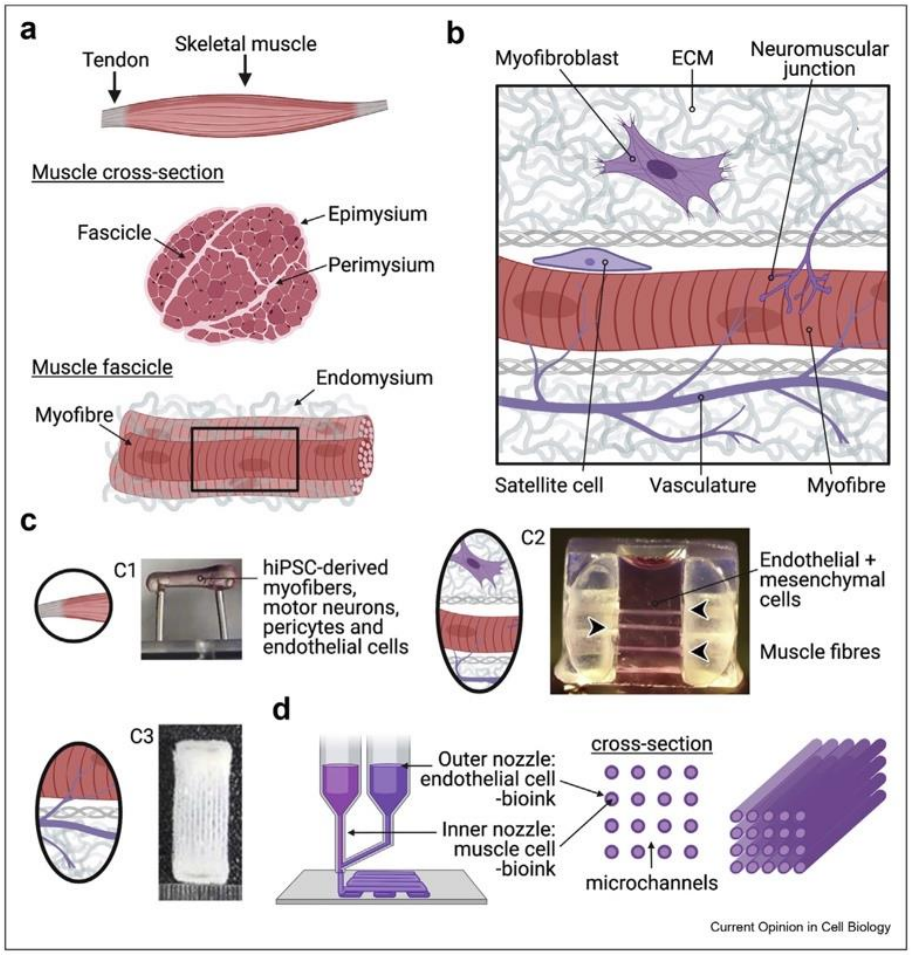

Figure 1-5. *Skeletal muscle architecture in vivo and engineered attempts at recreation.*

A. *Structure of human skeletal muscle tissue.* ***B.*** *Schematic of skeletal muscle tissue niche.* ***C.*** *Examples of 3D skeletal muscle engineered constructs C1: human iPS cell derived construct in a fibrin hydrogel. C2: Myoblast cells differentiated in fibrin hydrogel then surrounded by fibrin gel containing endothelial cells and fibroblasts. C3: Myogenic progenitors in 3D bioprinted cylindrical tubes with or without endothelial cell layer encasement.* ***D.*** *Coaxial 3D bioprinting setup illustration. Created using BioRender.com and adapted from (Jalal, Dastidar and Tedesco, 2021),*

1.4.3.1.2.1 Organ-On-Chip and Microfluidic Platforms

Research into organ-on-a-chip and microfluidic platforms is generally directed towards the study of muscle and neural interactions, neuromuscular junctions (NMJ) and neuromuscular circuits (NMC). This is of particular relevance to neuromuscular disorders, including DMD, as they are complex disorders in which both tissues are impacted and contribute to disease phenotype. Modelling these interactions *in vitro* is challenging due to the early detachment of differentiating myotubes in monolayer cultures, hence the need for other systems.

In 2018, Osaki and colleagues report the development of a microfluidics platform, to model ALS, using iPS cell-derived skeletal micro 3D muscle in a collagen-I Matrigel® matrix between 2 pillars in one compartment and iPS cell-derived motor neuron cells in a collagen-I Matrix in another compartment (Osaki, Uzel and Kamm, 2018). They observed axonal outgrowth and formation of neuromuscular junctions. Motor neuron progenitor cells from wild type and ALS patients were transfected with an optogenetic light-sensitive ion channel, channelrhodopsin-2, plasmid that enabled light activated muscle contractions;

the results show that ALS MNs generated fewer muscle contractions compared to control group (measured through pillar deflections), had MN degradation, and increased apoptosis in the micro muscle. Muscle contractions were recovered by single treatments and cotreatment with rapamycin and bosutinib as a proof-of-concept that the platform can be used for therapy screening.

In 2019, a couple further reports were published, both of which use iPS cell- derived muscle and neural cells, with incorporation of optogenetic stimulation, in a microfluidics platform. Vila and colleagues developed a system where there are different chambers for muscle cells, in a collagen mix matrix which form a bundle between two posts, and optogenetic motor neurons (Vila *et al.*, 2019). They validate the system in measuring NMJ function through a series of experiments that first validate their outcome measures (displacement, which can be used to calculate contractions), then demonstrate disruptions to NMJ function following neurotoxin treatment and finally they show that using myasthenia gravis (MG) patients' sera in culture medium of their model, in concentrations as low as 0.1% in one case, can be detected by the platform. Bellmann and colleagues have a very similar approach in their microfluidics platform (MFP), however, they rather than rely on a reactive polymer poly(ethylene-alt-maleic anhydride) (PEMA) coating to ensure that the skeletal muscle differentiating cells adhere for longer and do not delaminate before functional circuits can be formed; motor neurons also adhere to PEMA (Bellmann *et al.*, 2019). Following that they developed an assay platform for monosynaptic tracing, based on a recombinant rabies virus (RABV) conjugated to mCherry in

starter myoblasts that would only activate and infect motor neurons only in the presence TVA (only expressed in myoblasts), therefore the virus can only infect MNs at NMJ sites. They demonstrated the first proof-of-concept that monosynaptic RABV tracing can be used on NMCs *in vitro*. In contrast to patch clamping, calcium imaging and muscle contraction assays, the RABV tracing assay can be automated relatively easily, making it suitable for drug discovery (Bellmann *et al.*, 2019). Their MFP has 100 experimental units, which enables medium-throughput screening.

In conclusion, research into microfluidic platforms has been increasing in the past years specially that iPS cell technology is becoming more common. Some of these platforms incorporate 3D myotube bundles in efforts to prolong their time in culture, as monolayer cultures tend to delaminate, which comes with the added benefit of being more physiologically relevant. However, the use of different compartments to house different cell types mean less physiologically and architecturally relevant models, however, they are still very useful for specific purposes, such as studying the NMJ. Perhaps that explains why out of the different platforms have recently been used to develop models for neuromuscular diseases, microfluidics platforms are mostly focusing on diseases with a major NMJ involvement such as ALS. So far, the use of these platforms to model DMD, to our knowledge, has not been explored.

1.4.3.1.2.2 Myobundles and 3D cultures

The more popular approach to creating 3D muscle organoids is using a mixture of biomaterials and progenitor cells, often with anchors, to facilitate

creating aligned, more physiologically relevant organoids. Most of these studies utilise primary cultures (myogenic populations) isolated from biopsies and human tissues.

In 2015, Madden and colleagues reported the development of a 3D muscle platform, referred to as “myobundles”, where they use primary cells in a fibrin and Matrigel® hydrogel. These myobundles exhibit hallmarks of muscle tissue including, aligned architecture, multinucleated and striated myofibers, a Pax7+ cell pool, spontaneous contractions, and response to electric stimuli (Madden *et al.*, 2015). They also evaluated the potential of using them for disease modelling by testing 3 classes of commonly used drugs and the results showed that myobundles closely mimic functional responses of native human muscle through multiple signalling pathways and can therefore be used as a model for drug screening.

Following that, the group characterised some metabolic functions and evaluated muscle injury biomarkers of the myobundles in efforts of standardising their use as models. Kondash and colleagues demonstrated that despite tissue-engineered human myobundles exhibiting minor increase in glucose uptake in response to insulin, they recapitulate key features of *in vivo* insulin sensitivity and exhibit similar drug-mediated perturbations in contractile function and glucose metabolism, glucose uptake and insulin response (Kondash *et al.*, 2020). Khodabukus and colleagues investigated a recently identified four-factor muscle injury panel (MIP) validated in DMD patients (Burch *et al.*, 2015; Hathout *et al.*, 2016; Goldstein, 2017). They found that 3 of the MIP biomarkers’ concentrations

slightly increase in culture media of myobundles that have been subjected to snake venom (notexin) (Khodabukus *et al.*, 2020). They also showed that myobundle contractile function was significantly reduced, along with myotube atrophy when subjected to dexamethasone, but also noticed decreased the release of both traditional and novel biomarkers (Khodabukus *et al.*, 2020).

In 2019, Mills and colleagues reported utilising the use of a collagen-I based hydrogel along with a serum-free differentiation protocol to create their human skeletal micro-muscle platform (hμMs) using primary human cells (Mills *et al.*, 2019). These micro-muscles are organised bundles of striated and functional myofibres that respond to electrical stimulation and, significantly increased expression of maturation markers in micro-muscles that are ontogenetically stimulated.

Despite the promising potential of the aforementioned studies on artificial 3D muscle platforms, none of them investigated creating organoids from neuromuscular patient cells, let alone the testing of therapeutics to observe whether there would be a restoration of function and how closely does it recapitulate what happens *in vivo*. It is likely that access to these cells is difficult given the rarity of most of these diseases and the need to use patient biopsies by their physicians. This confirms the need for iPS cell-derived platform, with their virtually unlimited expansion potential, they can be used to investigate platforms and high through-put set-ups for drug and therapeutics' screening.

One recent study by Rao and colleagues, has investigated the use of human PS cells for muscle tissue engineering, based on the methods published

by Madden and colleagues (Madden *et al.*, 2015). They used induced myogenic progenitor cells with transient overexpression of Pax7 in human iPS cell-derived paraxial mesoderm-like cells in a monolayer culture (Rao *et al.*, 2018). These cells were then suspended in a hydrogel (Matrigel®-based) in a nylon frame and differentiated in 3D. The study demonstrates that 3D muscle tissue obtained contains aligned multi-nucleated myotubes, exhibits a positive force–frequency relationship and robust calcium transients (in response to electrical or acetylcholine stimulation). The produced 3D skeletal muscle tissues were successfully engrafted and retained functionality upon implantation in immunocompromised mice, however, this was done subcutaneously, and the constructs were not tested in disease-modelling setups, in contrast to what is reported below by Maffioletti and colleagues (Maffioletti *et al.*, 2018).

It is worth noting that all the aforementioned artificial muscle organoids are not complex, as they are only composed of myogenic cells and no other cell lineages despite the need for clinically-relevant complex patient-specific models (Giacomelli *et al.*, 2017). Development of such platforms has only been explored, so far, by Maffioletti and colleagues (Maffioletti *et al.*, 2018). In this study, hESCs and human iPS cells have been used to produce 3D muscle constructs using both transgene-based (Tedesco *et al.*, 2012; Maffioletti *et al.*, 2015) and transgene-free (Caron *et al.*, 2016) protocols. This study reports the first use of cellularly isogenic multi-lineage muscle constructs containing essential muscle cell types derived from the same iPS cell source (skeletal myogenic cells, endothelial cells, pericytes and neural progenitors, that further differentiate within

the 3D culture). The resulting 3D skeletal muscle tissues were successfully engrafted and vascularised upon implantation in immunocompromised mice. The study also demonstrates that the method is reproducible using diseased (limb girdle muscular dystrophy, DMD and LAMIN A/C (LMNA)-related muscular dystrophy) and healthy lines. This platform has been used by Steele-Stallard and colleagues to model laminopathies, a group of disorders caused by mutations in the LMNA gene, and reports improved detection of disease-associated phenotypes in 3D skeletal muscle cultures compared to a monolayer culture (Steele-Stallard *et al.*, 2018).

Bakooshli and colleagues report creating mini-muscle organoids with two cell lineages, using iPS cell-derived motor neurons and primary human muscle progenitors (Bakooshli *et al.*, 2019). They show that these organoid co-cultures have multinucleated, aligned and striated muscle fibres and the MN clusters were positioned at the periphery of muscle bundles and recapitulate early neuromuscular junction (NMJ) synaptogenesis. Functional connectivity between motor neuron endplates and myofibres is confirmed with calcium imaging and electrophysiological recordings. In efforts to show the potential of using these organoids in modelling neuromuscular diseases, they treated them with myasthenia gravis patient sera; analysis showed a marked decrease in the area of stimuli responsive muscle fibres and a decline in the area of the tissue responsive to acetylcholine (Ach) stimulation.

The option of 3D printing skeletal muscle tissue has been explored by Capel and colleagues (Capel *et al.*, 2019) and Kim and colleagues (Ji Hyun Kim

et al., 2018). However, the Capel and colleagues report utilised murine cells and the Kim et al. study utilised primary human cells. Furthermore, none of these studies have reported robust myotube alignment, functional outputs and multilineage applications that are comparable to studies published in 2018 (Maffioletti *et al.*, 2018; Rao *et al.*, 2018). For a more in depth review on methods and approaches of bioprinting, refer to (Ostrovidov *et al.*, 2019).

The use of iPS cells to derive 3D models (engineered muscle platforms) for neuromuscular diseases has great potential for drug screening and testing therapies in medium and high throughput settings, and in patient-specific applications. This is of particular relevance and importance for DMD research which can be caused by over 4000 different mutations in the gene. While monolayer (2D) cultures of myogenic cells (whether primary myoblasts or derived from iPS cells) are used as models, recapitulating *in vivo* increased biomarker release, impaired myoblast fusion, and impaired calcium-handling; engineered muscle models offer physiological relevance and greater muscle maturation; in contrast to traditional monolayer cultures, localised mature membrane-bound dystrophin is present in both human primary and iPS cell-derived 3D engineered muscles (Rao *et al.*, 2018; Khodabukus *et al.*, 2020). However, these platforms need further validation and research into the more advanced DMD clinical features, including fibrosis, lipid accumulation, myofibre branching and membrane permeability; in addition to validation using commercially available DMD drugs and therapies.

1.5 GENE EDITING FOR DISEASE MODELLING

1.5.1 OVERVIEW OF GENE EDITING TOOLS - ENGINEERED NUCLEASES

Targeted nucleases have been used by researchers to manipulate various genomic sequences and in many ways, as early as the discovery of restriction enzymes in the 1970s, followed by more precise techniques, known as engineered nucleases, from the 1980s onwards. These tools enabled the creation of animal models and genetically isogenic cell lines, in addition to their potential in gene therapies as described earlier in this introduction.

There are four different types of engineered nucleases, defined by their distinct ways of identifying their target sequence and cleaving the DNA. These are, meganucleases, zinc finger endonucleases (ZFN), transcription-activator like effector nucleases (TALENs) and the most recently discovered clustered regularly interspersed short palindromic repeats (CRISPR/Cas9). These approaches rely on creating a double-strand break (DSB) in the genome, which then stimulates the intrinsic cell molecular repair pathways that happen during the cell cycle (Vilenchik and Knudson, 2003). There are two main mechanisms: non-homologous end joining (NHEJ), which is error-prone and produces random indel mutations mostly causing gene knockout and homology-directed repair (HDR) which is high-fidelity as it uses an additional oligonucleotide sequence

(naturally, this would be the other copy of the gene) to repair the damage. The prioritisation of one mechanism over the other is determined by cell-cycle specific molecular partners (Ira *et al.*, 2004; Shibata *et al.*, 2011; Tomimatsu *et al.*, 2014; Lord and Ashworth, 2017).

The NHEJ repair mechanism starts with the Ku70-Ku80 heterodimeric protein holding together the two cleaved ends, which subsequently recruits the DNA-dependent protein Kinase's catalytic subunit (DNA-PKcs) which is a signaling threonine/serine kinases activated upon cell stress (M. Li *et al.*, 2014). In cases where the DSB results in single stranded DNA (ssDNA) ends, they are first trimmed by the nuclease Artemis' exonuclease activity (Gu *et al.*, 2010). Finally, the DNA ligation complex composed of DNA ligase 4, co-factor X-ray cross-completion 4 (XRCC4) and XRCC4 like factor (XLF)/Cernunnos is recruited It then ligates the blunt-ends of DNA, repairing the break site (Williams *et al.*, 2014; Radhakrishnan and Lees-Miller, 2017; Li *et al.*, 2018).

The HDR repair mechanism is initiated through the resection of the 5' strand at the break site by nucleases and helicases (Limbo *et al.*, 2007; Stracker and Petrini, 2011). The nucleases generate 3' ssDNA overhangs (Sartori *et al.*, 2007). The recombinase RAD51 is then recruited and associates with the complex, creating a nucleoprotein filament (Wassing *et al.*, 2021). This enables the ssDNA to invade sister double stranded DNA which possess homologous region and use them as a repair template resulting in a scar-less repair (Shivji *et al.*, 2009). Depending on the desired genetic alteration, researchers utilise one of either pathways (Pawelczak *et al.*, 2018). NHEJ mainly occurs during the G1

phase of the cell cycle but can occur at other phases albeit much less frequently, where simple ligation of the two ends of the DNA (at the cut site) results in unfaithful repair products, these are either, small (between 1 bp and 50 bp) insertions and/or insertions (known as “indels”), substitutions, or translocations (Lieber, 2010). It is therefore mainly used to disrupt the reading frame of the gene of interest. On the other hand, HDR occurs during defined stages of the cell cycle (late S phase and/or G2 phase) (M. Li *et al.*, 2014). For HDR, the presence of a repair template sequence (an identical, or almost identical sequence) is essential and therefore this pathway is mainly exploited for more precise applications such as correction or insertion of specific mutations and substitution or insertion of specific sequences at a specific site.

1.5.1.1 Meganucleases

The term meganucleases, also commonly referred to as homing endonucleases, covers are a group of restriction enzymes endogenous to various prokaryotic and eukaryotic organisms (Belfort and Roberts, 1997; Jurica and Stoddard, 1999). They are classified into five major families, differentiated by specific sequences and structural motifs (Silva *et al.*, 2011). The first group, which is the largest and best characterised, is the LAGLIDADG homing endonucleases (LHEs). The name refers to the functional polypeptide motif that is characteristic of this group and is present as a single chain protein or as a tandemly repeated pair (Duan, Gimble and Quiocho, 1997; Heath *et al.*, 1997). All LHEs, regardless of their molecular structure, arrange in a αββαββα fold in which the LAGLIDADG

motif is central to the α-helices, surrounded by β-sheets (these interact with the target DNA).

Meganucleases target recognition sites are statistically uncommon in a genome and range from 12 bp to 40 bp in length (Thierry and Dujon, 1992; Chevalier and Stoddard, 2001). It is due to this feature that meganucleases are considered the most precise amongst all engineered nucleases. This precision comes with a major drawback, which is that the likelihood that a genomic target includes a recognition site is low, and for certain genes and applications it must be at a certain locus/exon. This is further complicated by the fact that the repertoire of naturally existing meganucleases is limited (Pâques and Duchateau, 2007).

In efforts to improving cleaving efficiency and enriching target options, scientists produced chimeric meganucleases by either fusing/associating them with enzymatic domains of different meganucleases or by introduction of small modifications in the sequence-recognition motif (Chevalier *et al.*, 2002; Seligman *et al.*, 2002; Rosen *et al.*, 2006; Smith *et al.*, 2006; Muñoz *et al.*, 2011).

1.5.1.2 Zinc-finger nucleases (ZFN)

ZFNs is a protein-DNA association based man-made tool composed of two parts. First part is the DNA cutting domain (that lacks specificity) of the natural type IIS restriction enzyme called *FokI*- which is easily separable from the *FokI* DNA recognition domain (Li, Wu and Chandrasegaran, 1992). The second part is the customisable DNA recognition domain that replaces the *FokI's* natural one.

Chandrasegaran *et al.* observed that replacing the DNA recognition/binding domain affects and directs genomic target specificity of the enzyme (Kim and Chandrasegaran, 1994; Kim, Cha and Chandrasegaran, 1996; Kim *et al.*, 1998).

Amongst the tested DNA binding domains, were a set of three Cys2His2 zinc-fingers, each of which is able to recognise a DNA 3-base pair site and insert an α-helix into the major groove of the DNA double helix and bind to it (Pavletich and Pabo, 1991). The combination of multiple domains enables the recognition of longer sequences, and therefore, these man-made chimeric proteins were named zinc finger nucleases (ZFNs).

As most ZFNs are composed of 3-4 zinc finger motifs (fused to a monomeric *FokI* DNA cutting domain), their recognisable target sequences are therefore 9 bp to 12 bp (Porteus, 2010). The *FokI* endonucleases are catalytically active as a dimer, and therefore, DSB DNA cleavage only occurs at sites where ZFNs are juxtaposed and bind on opposite strands- ideally with a 5 bp to 6 bp spacer sequence (Mani *et al.*, 2005). In comparison to MNs, the tailoring of ZFNs allows the identification of a wider set of DNA targets. These features inherently increase the cleavage specificity and reduce the off-targets effects of this technology, however, make it quite challenging to design and use.

The technology has been successfully applied in gene editing various organisms with different end results (gene disruption, insertion, and correction of mutations, utilising NHEJ or HDR) (Beumer *et al.*, 2008; Perez *et al.*, 2008; Santiago *et al.*, 2008; Foley *et al.*, 2009; Geurts *et al.*, 2009; Cost *et al.*, 2010; Laoharawee *et al.*, 2018; Ran *et al.*, 2018). However, it is worth noting that a

major hurdle to the wide adoption of ZFNs is the high cost of assembling ZFNs. In addition to that, the each of the of zinc-finger modules in the array has an affinity to the target DNA sequence is dependent and affected by the surrounding modules and the chromatin accessibility of the target sequence. This context dependence interferes with successful assembly and DNA cleavage (Ramirez *et al.*, 2008).

1.5.1.3 Transcription Activator-Like Effector Nucleases (TALENs)

The discovery of naturally synthesised virulence factors of the phytopathogenic bacteria in the *Xhantomonas* genus, known as transcription activator-like effectors (TALE) led to the development of TALENs as a gene editing tool (XU *et al.*, 2017). TALE's resemblance to eukaryotic transcription factors enables them to affect plants' gene expression and the development of disease symptoms (XU *et al.*, 2017). They are characterised by centrally located domain of tandem repeats that is able to bind to single DNA nucleotides (Boch *et al.*, 2009). Scientists recognised the potential use of TALEs as a customisable DNA sequence targeting tool (Mussolino and Cathomen, 2012).

TALENs technology is, similar to ZFNs, a chimeric protein-DNA association-based technology that relies on *FokI* endonuclease domain fused to an array of up 20 modules (and therefore a recognition sequence of 15 bp to 20 bp) of 33 to 35 amino acids each in tandem (rather than the zinc-finger modules in ZFNs) This makes the engineering of TALENs significantly easier than ZFNs. Within each of the modules of the array, the two central amino acids at positions

12 and 13 (commonly known as repeat-variable di-residues (RVDs)) are the amino acids that recognise and bind to the single DNA nucleotide (Deng *et al.*, 2012), this ability to multiplex modules at a 1 module per nucleotide target makes this nuclease the one with the highest resolution compared to all other nucleases developed up to then.

However, other challenges that prior technologies face still stand. Similar to ZFNs, dimerisation is required for cleavage to occur, therefore, two different juxtaposed TALENs binding on opposite strands close to the target DNA with a spacer length of 12 bp to 21 bp is required. The length of the spacer and the TALEN recognition site are longer than that of ZFNs and therefore have a reduced off-target effects (Gupta and Musunuru, 2014). Another challenge facing TALENs is that the cloning methods tested to produce functional arrays proved to be time consuming and laborious, however the advantages far outweigh all the disadvantages, and this led to the quick adoption of TALENs by scientists (Cermak *et al.*, 2011; Mussolino and Cathomen, 2012).

1.5.1.4 Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas9

The first reported instance and discovery of the clustered regularly interspersed short palindromic repeats was in the late 1980s. The Atsuo Nakata research group at Osaka University was trying to clone the *iap* gene from *Escherichia coli* and discovered a series of genomic repeats with a short “spacer” of DNA sequence in the genome, with a then-unknown function (Ishino *et al.*, 1987; Nakata, Amemura and Makino, 1989).

Soon after, a study by Van Soolignen and colleagues reported finding and the presence of these clustered genomic sequences of uninterrupted repeats in multiple strains of *Mycobacterium tuberculosis*, however, each strain had a different interrupting sequence (Groenen *et al.*, 1993; Van Soolingen *et al.*, 1993). Concurrently and independently, sequences of around 30 bp with regular repeats within the archaeal genomes of *Haloferax mediterranei* and *Haloferax volcanii* was identified by reported by Mojica and colleagues (Mojica, Juez and Rodriguez-Valera, 1993; Mojica *et al.*, 1995).

A few years later (late 1990s) the introduction and more widespread use of sequencing technology for various applications allowed the discovery of similar sequences in numerous microbial organisms (Mojica *et al.*, 2000; Rund Jansen *et al.*, 2002). Mojica and Jansen coined the acronym CRISPR to avoid the ambiguity in nomenclature, and it became widely used from then on (Ruud Jansen *et al.*, 2002). Since then and up to this date, CRISPR DNA sequences have been identified in approximately 90% of archaeal DNA and 40% of sequenced eubacterial genomes (Mojica *et al.*, 2000).

The function of these loci remained unknown until the year 2005, three independent research group reported their findings suggesting that the sequences come from plasmid DNA and DNA fragments from bacteriophages that have previously infected the prokaryotes (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel, Salvignol and Vergnaud, 2005). The work of Barrangou and colleagues demonstrated that upon insertion of an invader organism viral sequences in the CRISPR locus of the bacteria *Streptococcus thermophiles* it

became resistant to further bacteriophage infection (Barrangou *et al.*, 2007). A year later, Marraffini and colleagues reported CRISPR-mediated inhibition of horizontal plasmid transfer (Marraffini and Sontheimer, 2008). Together, these reports suggested that CRISPR is part of the prokaryotic adaptive immune response system against invading plasmids and phages.

The elucidation of the molecular mechanisms by which the CRISPR-based adaptive immune response system came shortly after (late 2000s). Brouns and colleagues (Brouns *et al.*, 2008) and Garneau and colleagues (Garneau *et al.*, 2010) demonstrated that the CRISPR-based system performs targeted cleavage of invading nucleic acids using a specific CRISPR-associated nuclease (Cas) and guided by RNA molecules transcribed from the CRISPR locus. The target DNA sequence is primarily dependent on its proximity to a 2-5 nucleotide motif, known as the protospacer adjacent motif (PAM) (Shah *et al.*, 2013; Cencic *et al.*, 2014).

Once the invading DNA fragment is cleaved, it is incorporated into the CRISPR locus. The fact that sequences from each invading genetic element is incorporated in the CRISPR locus of the host genome, means that the locus possess a chronological memory of all infections that occurred in the past.

The system is characterised by the proximity and arrangement of its components in the host genome. First, a set of homologous genes coding for the CRISPR-associate (Cas) genes, followed by an A-T rich sequence (known as the leader sequence) and finally the CRISPR array (composed of the repeats and spacers) (Ruud Jansen *et al.*, 2002).

Cas genes are a set of genes that code for helicases, nucleases, RNA-binding proteins, and polymerases that are involved in the recognition and cleavage of the target DNA (Ruud Jansen *et al.*, 2002). The leader sequence following the Cas genes is a promoter for the downstream CRISPR array (Pul *et al.*, 2010). The CRISPR array is composed of a variable number of repeated and semi-palindromic sequences that are separated by short sequences of DNA (protospacers) derived from invader genomes. The partially palindromic sequences range from 23 bp to 47 bp in length while the protospacers' length varies from 20 bp to 70 bp.

The CRISPR systems known so far are traditionally classed into two classes (I and II), this is depending on the number and biochemical roles of the Cas proteins; and each of those classes is divided into three types (Makarova *et al.*, 2015). Each of these types has been further divided into subtypes by some researchers (Makarova *et al.*, 2015; Koonin, Makarova and Zhang, 2017). CRISPR systems that belong to class I (including CRISPR type I, III and IV) rely on multiple Cas proteins to perform various molecular steps that finally lead to DNA cleavage (Cooper, Stringer and Wade, 2018). On the other hand, CRISPR systems that belong to class II (including CRISPR type II, V and VI) use a single Cas protein for the recognition and cleavage of target invading DNA.

Class	Type	# of sub-types	Cas endonuclease	Target	Requires tracrRNA?
Class I	Type I	7	Cas3	DNA	No
	Type III	4	Cas10	DNA/RNA	No
	Type IV	1	_____	_____	_____
Class II	Type II	3	Cas9	DNA	Yes
	Type V	3	Cas12	DNA	Yes (1 subtype)
	Type VI	3	Cas13	RNA	No

Table 1-1. *Classification of CRISPR-systems*

The most extensively studied and used of these systems is the type II CRISPR system (Cas9 endonuclease based) derived from *Streptococcus pyogenes*, which belongs to subtype IIA. This system the commonly referred to as CRISPR/Cas9. It is one of the subtypes that are characterised by the presence of a non-coding trans-acting RNA (tracrRNA) located upstream of the Cas genes (in this subtype, these are four in total) but on the opposite DNA strand (Chylinski, Le Rhun and Charpentier, 2013).

The four Cas genes that are present in this subtype are the Cas1, Cas2, Cas9 and Csn2. Csn2 is the differentiating gene specific for the subtype IIA CRISPR systems, it binds to the double stranded DNA (Koo, Jung and Bae, 2012). Heler and colleagues demonstrated that the target DNA is labelled by the Cas9 protein, which then recruits the integrases Cas1 and Cas2 proteins (Heler *et al.*, 2015).

1.5.2 CRISPR/CAS9 IN DEPTH

1.5.2.1 Molecular Mechanisms in Bacteria

When a *Streptococcus pyogenes* cell is invaded for a second time by a pathogen, the type II CRISPR system is activated and after a series of defined sequential steps, summarised below, induces a targeted DNA double stranded break.

Firstly, as soon a pathogen invades the cell, the leader sequence of the CRISPR system locus acts as a promoter of transcription of the CRISPR array into a premature crRNA (pre-crRNA transcript). Simultaneously, the non-coding tracrRNA is co-transcribed from the opposite strand of the CRISPR system locus. Subsequently, it hybridises with the pre-crRNA's short palindromic sequences (Deltcheva *et al.*, 2011). This hybridised RNA complex recruits Cas9, which binds to it and stabilises it, and RNAse III, to form a protein-RNA complex that cleaves the premature crRNA into a mature crRNA (Deltcheva *et al.*, 2011). This process happens with all pre-crRNAs transcribed from the CRISPR array. The cut induced by the Cas9-RNAse III complex is in the same site in all hybridised regions, therefore, all crRNAs that arise have a common sequence in addition to a unique sequence corresponding to their specific spacer DNA sequence (Brouns *et al.*, 2008).

Finally, the tracrRNA-crRNA complex probes the invading DNA and binds to the complementary sequence that is adjacent to the protospacer adjacent motif (PAM) sequence. In *Streptococcus pyogenes* this sequence is a 5'-NGG-3'

trinucleotide (Sternberg *et al.*, 2014). the Cas9 nuclease is then recruited, reconstituting the interference complex, binds transiently to the DNA and initiates its unwinding- starting with the 10 to 12 nucleotides that are proximal to the PAM and moving away from it (Szczelkun *et al.*, 2014; Gong *et al.*, 2018). The unwound DNA sequence is known as the seed region/sequence, and once the unwinding is completed, the Cas9 pairs it with the rest of the crRNA spacer sequence generating: 1- an RNA-DNA hybrid (known as an R-loop) and 2-a displaced single stranded DNA (ssDNA) which contains the PAM and is the non-target strand) (Jiang *et al.*, 2016). In cases of incomplete or inefficient pairing, the Cas9 is released and continues sampling other potential DNA targets (Yang *et al.*, 2018). However, in cases of successful pairing, the Cas9 cleaves both strands of the DNA precisely at 3 bp upstream of the PAM site, resulting in a blunt end double stranded break (DSB) (Jinek *et al.*, 2012).

The *Streptococcus pyogenes* Cas9 protein has a dual role recognition and cleavage. Studies have shown that it is composed of two lobes linked by two linker segments. The lobes are commonly referred to as the recognition (REC) lobe and the nuclease/cleavage (NUC) lobe. The quaternary structure of the Cas9 and the interaction between these two lobes creates a positively charged groove in between that accommodates the DNA-RNA hybrid. The REC lobe mostly interacts with the RNA, it is alpha helical and is the most variable region of the Cas9 protein (Jinek *et al.*, 2012, 2014; Nishimasu *et al.*, 2014; Palermo *et al.*, 2018). The NUC lobe's function is to cleave target DNA. It is composed of two defined domains that mediate cleavage (Zuo and Liu, 2017). The first domain,

the RuvC-like nuclease domain, cleaves the complementary strand and directs the entire nuclease to the target genomic region as it contains a c-terminal PAM-interacting region, while the other domain, HNH-like nuclease domain, cleaves the target DNA strand.

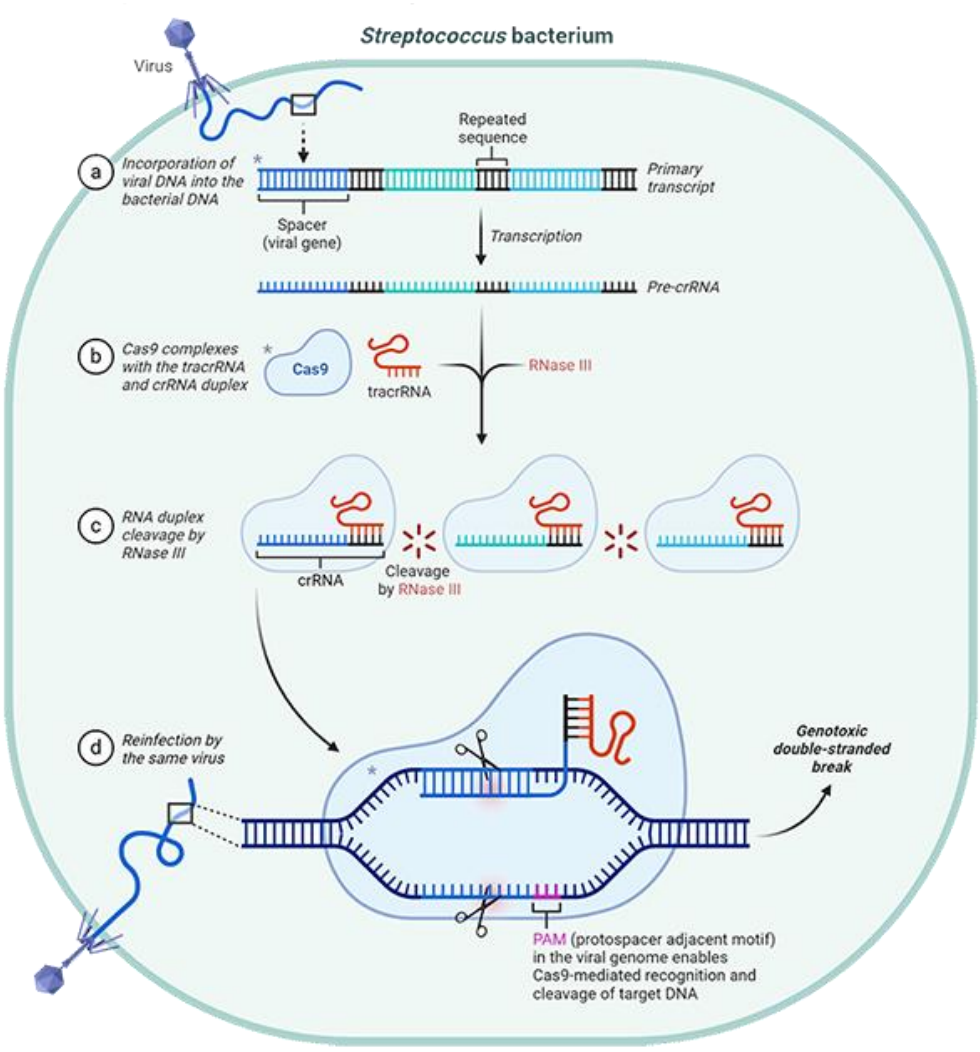

Figure 1-6. *Streptococcus' CRISPR/Cas9 endogenous molecular mechanism and immune response against viruses.*
Adapted from "2020 Nobel Prize in Chemistry: A Tool for Genome Editing (CRISPR-Cas9)". Made with BioRender.com.

1.5.2.2 CISPPR/Cas9 as a Gene Editing Tool

The fact that CRISPR Cas type II system of *Streptococcus pyogenes* is functional with the presence of just three components (mature crRNA, tracrRNA and the Cas9) allowed scientists to easily repurpose and exploit it for the development of RNA-guided genome editing (Mali *et al.*, 2013). In efforts to improve and simplify the system, scientists fused the crRNA (containing the target DNA sequence) and the tracrRNA, reducing the number of components required for targeted DNA cleavage. This fused structure is known as small guide RNA (sgRNA, or sometimes gRNA) (Mali *et al.*, 2013).

Small guide RNAs are composed of around 80 nucleotides, the first 20 nucleotides (from the 5' end) are the crRNA equivalent (complementary to the target DNA sequence) and the remaining 60 nucleotides fold into a hairpin structure replicating the tracrRNA- crRNA complex structure. This structure has been proven to allow a better orientation for sgRNA binding and facilitates Cas9 recruitment (Wiles *et al.*, 2015). The main requirement, without which the system would not work, is the presence of a PAM site starting at the 5' end of the sgRNA, and these sites are quite common in most genomes. Other efforts focused on the scaling of the system, such as introducing multiple guides to allow for multiple target cuts, and on reducing off-target effects by multiplexing sgRNA cutting in the same region but using modified Cas that possesses a nickase activity cutting only one strand rather than causing a double stranded break (Cong *et al.*, 2013). To optimise the system even further, scientists used codon optimisation and

added a nuclear localisation signal to direct the Cas9 to the nucleus where the genome is (Mali *et al.*, 2013).

As detailed earlier, the ability to exploit either HDR or NHEJ repair pathways following a DNA DSB means that any type of DNA mutation can be achieved using CRISPR/Cas9 gene editing system. The ease and standardisation of sgRNA customisation and production, in comparison to the targeting components of other engineered nucleases, makes CRISPR/Cas9-mediated gene editing the preferred system for various (*in vitro*, *ex vivo* and *in vivo*) applications, including producing *in vitro* and animal models and potentially as a gene therapy for the treatment of diseases, such as DMD (Waddington *et al.*, 2016).

1.5.2.3 CRISPR-Cas9 Gene-Editing Limitations

The off-target effects remain the most significant issue for gene-editing applications, as nuclease specificity in recognising the target sequence is well document to not be 100% accurate. Sites that share a high percentage of homology with the desired target site are found in both coding and non-coding regions of the genome- in the middle of a gene or in regions such as enhancers and repressors (Hsu *et al.*, 2013). The result of nonspecific nuclease activity can sometimes lead to significant effects, as other genes can be inactivated, mutated or up/downregulated through the introduction of random indels or acquisition of the cassette (Zhang *et al.*, 2015). This can lead to a cascade of events disrupting

the homeostasis of different cell pathways and mechanisms (Kleinjan and Coutinho, 2009).

In silico tools developed to predict off-target potential effects are not very accurate. Potential off target sites for a single sgRNA are different in predictions using different algorithms (Hsu *et al.*, 2013; Tsai *et al.*, 2014; Frock *et al.*, 2015; Doench *et al.*, 2016). This is due to multiple causes including the fact that *in silico* tools rely on reference genomes and do not consider the individual genetic background variability in a number of sites (estimated to be up to 5 million) (Lessard *et al.*, 2017). These variabilities and SNPs can also lead to the creation of *de novo* PAM sites that means potentially new targets (Lessard *et al.*, 2017). In addition, the variability in the epigenetic status of the target region can also play a role in the inaccuracy of *in silico* tools, which do not account for chromatin accessibility (Jensen *et al.*, 2017; Uusi-Mäkelä *et al.*, 2018). For a 100% guarantee that the designed sgRNA specifically targets the desired site, whole genome sequencing prior to guide design and use would be ideal; however, this is currently impractical due to cost and time limitations (Tsai *et al.*, 2017).

Another approach for increasing specificity and reducing off-target effects of Cas9 is engineering the protein itself, to be discriminatory towards on-target activity and disfavour any off-target activity as a result of partial complementarity between the sgRNA and genomic DNA; an approach that multiple research groups have investigated and developed higher precision engineered Cas9s (Slaymaker and Gaudelli, 2021). Two strategies are commonly being used, rational-structure guided designs and directed evolution (Adli, 2018). These

efforts have led to upwards of 20 spCas9 higher fidelity variants with one or more mutations each, generally concentrated around active and catalytic domains (Slaymaker and Gaudelli, 2021). These mutations confer differences in the trade-offs between overall gene editing efficiencies and specificities (Kim *et al.*, 2020). A single point mutation, R691A, that leads to reduced off-target editing while maintaining robust gene disruption was identified using an unbiased bacterial screen, leading to the high fidelity (HiFi) Cas9 variant (Lee *et al.*, 2018; Vakulskas *et al.*, 2018)l.

1.5.3 CRISPR-Cas9 Gene Edited Disease Models

The inherent simplicity and flexibility to introduce specific genetic modifications with CRISPR/Cas9 genome editing technology has so far proven to be a very useful tool in a variety of applications, from basic research to drug discovery, leading to widespread and fast adoptions by scientists.

The technology has been used in creating genetically modified animals and model organisms, simply through injecting sgRNAs and Cas9 mRNA into single cell embryos, decreasing the time and complexity of generating transgenic animal models compared to traditional methods, especially in polygenic diseases (Shen *et al.*, 2013; Wang *et al.*, 2013; Shao *et al.*, 2014).

Polygenic complex diseases and mechanisms that involve multiallelic mutants have particularly benefitted from the emergence of CRISPR/Cas9 technologies (Dow, 2015). This has been extensively used in mouse and porcine models (Wang *et al.*, 2013; Heckl *et al.*, 2014; P. Li *et al.*, 2015), DMD specific

models include (Nakamura *et al.*, 2015; Yu *et al.*, 2016; Birling *et al.*, 2017). Using CRISPR/Cas9 to introduce multiple double stranded breaks can be used to induce structural variations (SVs) which are large genomic structural differences such as chromosome deletions, duplications, insertions and inversions. Multiple human disorders are a result of one or more genes with such SVs and many models, previously very difficult to create, have been created and validated using CRISPR/Cas9 (Kraft *et al.*, 2015; Zhang *et al.*, 2015; Boroviak *et al.*, 2016; Khan *et al.*, 2016; Arend, Pereira and Markoski, 2017; Birling *et al.*, 2017; Millette and Georgia, 2017).

Another approach is using AAVs (using the serotype with the desired tissue target) to deliver CRISPR/Cas9 gene editing components in model embryos to create transgenic Cas9 expressing organisms under the drive of a tissue specific promoter. These organisms robustly express Cas9 in a tissue-specific manner, and when the organism is injected with AAVs delivering sgRNAs postnatally, this results in robust editing of the sgRNA target locus and only in the target tissue. This approach allows the study of adult tissue-specific functions of genes that are widely expressed in various tissues or that cause lethality in embryos when mutated, and has been demonstrated in mouse cardiac tissue to model cardiomyopathies (Carroll *et al.*, 2016).

CRISPR/Cas9 gene editing has also been used in *ex vivo* and *in vitro* disease modelling applications and platforms. In diseases where genotype phenotype correlation in humans does not match that of the model organism, for example colorectal cancer where in mice high tumour burden limits metastasis,

strategies such as inducing mutations and growing tumours *ex vivo* is necessary to model the adenoma-carcinoma-metastasis sequence of tumour progression (Roper *et al.*, 2018). *Ex vivo* studies are particularly relevant to cancer and metastasis research as genome sequencing studies have shown that human malignancies often have mutations in four or more key driver genes (TJ *et al.*, 2013). Heckl and colleagues demonstrated the use of CRISPR/Cas9 editing of 5 genes in producing haematopoietic stem cell lines with myeloid malignancies recapitulating the combinations of mutations observed in human disease (Heckl *et al.*, 2014). They used these cells in *in vivo* studies in mice and confirmed the robustness of the system to model acute myeloid leukaemia.

1.5.4 CRISPR/Cas9 in iPS Cell-derived Models

Perhaps one of the most promising areas of research that CRISPR/Cas9 technology has been utilised in, is editing human iPS cells. As extensively described above, the ethical and practical advantages of human iPS cells over alternatives (primary and immortalised cell lines, and animal models) have made them the platform of choice for multiple applications and prompted extensive research in differentiating them into various lineages and derivatives, in monolayers and three dimensions (which often includes multiple cell types).

One example of using CRISPR/Cas9 gene edited iPS cells in modelling disease is for immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF), a rare autosomal recessive disorder caused by mutations in DNA methyltransferase 3B which leads to hypomethylation of satellite 2 repeats

ultimately leading to genetic instability due to the decondensation of these regions. Researchers created ICF iPS cell lines by knocking out the gene in WT cells, and observed similar hypomethylation patterns to ICF patient cells (Horii *et al.*, 2013).

A main advantage of using iPS cells to create other disease models is the fact that this generates genetically isogenic pairs on which more robust analysis can be carried out. The two methods to generate isogenic iPS cell lines from parent lines is by either starting with a diseased line and correcting the mutation or a wild type line and inducing the mutation(s) desired. Work by Xu and colleagues highlighted the significance of isogenic pairs in disease modelling and research. To model Huntington's disease, they used CRISPR/Cas9 to generate corrected iPS cell lines from diseased iPS cells, after differentiation into forebrain neural progenitor cells then GABAergic neurons and functional confirmation of cell identity, they showed rescue of function in corrected iPS cells compared to isogenic diseased iPS cells, similar to wild type control cells. However, when they used global principal component analysis, the genetically isogenic (corrected and diseased) cells clustered together and apart from the non-isogenic wild type control (Xu *et al.*, 2017). Many studies relied on CRISPR/Cas9 generated isogenic pairs to derive more robust models and disease analyses (Flynn *et al.*, 2015; Gupta *et al.*, 2016; Ishikawa *et al.*, 2016; Marczenke *et al.*, 2017; Murakami *et al.*, 2017; Wang *et al.*, 2017).

Using CRISPR/Cas9 gene edited stem cells to generate self-organising organoids and tissue constructs is the logical next step to create more reliable

models and platforms. These human cell-based systems provide multiple niche interactions that more faithfully recapitulate structural and functional properties of organs (Lancaster and Knoblich, 2014) and the use of isogenic pairs in such systems allows for more robust analyses and findings.

Studies demonstrated that the use of CRISPR/Cas9 edited organoid models provides morphological and structural insights that are otherwise impossible to characterise as a phenotype in monolayer cultures. Latour and colleagues demonstrated that CRISPR/Cas9 edited iPS cell-derived cerebral organoids to induce GLB1 gene mutations that lead to GM1 gangliosidosis exhibited GM1 ganglioside storage disease phenotypes (Latour *et al.*, 2019). Freedman and colleagues created kidney organoids from hPS cells that have been edited to introduce disease mutations of polycystic kidney disease (PKD) in the PKD1 and PKD2 genes. Their results show that these organoids recapitulated glomerulopathy-like phenotypes *in vitro*. They formed cysts that are characteristic of the disease in kidney tubules *in vitro* and displayed junctional organisation defects in podocyte-like cells (Freedman *et al.*, 2015). Li and colleagues demonstrated that cerebral organoids generated from cells with a phosphate and tensin homolog (PTEN) gene knockout display a macrocephalic phenotype (Li *et al.*, 2017). The key findings elucidated many of the interspecies differences in the PTEN-related pathways, including stark differences in the folding mechanism of human cerebral organoids compared to mice (Li *et al.*, 2017).

2 AIM OF THE WORK

Despite developments in modelling DMD, the need for an overarching temporally-, economically-, and ethically-viable platform still stands. Several reasons make achieving this particularly challenging using traditional models: 1) The molecular and cellular dynamics of dystrophin, complexity of its expression and various isoforms, complicating its detection, especially in live cells. 2) The involvement of various tissues in the disease. 3) The difference between disease phenotype genotype in animal models compared to humans. 4) The lack of knowledge regarding the newly emerged humanised animal models and the time needed for in depth studies.

Therefore, this project aims to develop a novel human disease-modelling platform for DMD, on which dystrophin protein can be studied and experimental therapies can be tested. There are two components in this platform. The first is iPS cell lines that possess a novel multiple-reporter cassette inserted just before the dystrophin stop codon through CRISPR/Cas9 technology that allows for quantitative and qualitative tracking of dystrophin protein production levels temporally and spatially, in real-time and in fixed cells. The second is the three-dimensional skeletal muscle constructs made with iPS cell-derived myogenic progenitors

The key question driving this work was: is it possible to develop a robust human cell-derived platform to model DMD and study dystrophin dynamics in a

physiologically relevant environment without relying on animal models? The project was hence divided into the following aims:

1- Generate dystrophin reporting wild type human iPS cells using CRISPR/Cas9 gene editing technology and create isogenic dystrophic sister pair.
2- Validate the functionality and robustness of the various reporter sequences, NanoLuciferase luminescence, mCherry fluorescence, and HaloTag interchangeable labelling.
3- Assess the maturation and advantages of using 3D skeletal muscle constructs over monolayer cultures, in conjunction with setting up imaging modalities for the platform.
4- Provide proof-of-concept experiments utilising the dystrophin reporter lines in screening for next generation AON therapies.

3 MATERIALS AND METHODS:

3.1 BIOLOGICAL MATERIALS

All cells utilised for this project are from human origins; iPS cells and iPS-cell-derived myogenic and neurogenic cells. iPS cell lines used are:

UCLi007	https://hpscreg.eu/cell-line/UCLi007-A
NCRM1	https://hpscreg.eu/cell-line/ CRMi003-A
NCRM5	https://hpscreg.eu/cell-line/CRMi001-A
SBIi006-A	https://hpscreg.eu/cell-line/SBIi006-A
UCLi011-A	https://hpscreg.eu/cell-line/UCLi011-A

NCRM human iPS cell lines are CD34+ cord-blood derived progenitors that were reprogrammed to pluripotency using episomal expression of *OCT4, SOX2, KLF4, MYC, Nanog, Lin28* and *SV40T*. UCLI007 is an iPS cell line, also known as control number 4 and reported in (Tedesco *et al.*, 2012) that had been reprogrammed via retroviral vectors delivering the reprogramming factors *OCT4, KLF4,* and *SOX2.*

Other lines used in pilot trials for CRISPR/Cas9 editing are dystrophic lines from the MRC Centre for Neuromuscular Disorders Biobank London, with codes: CENSOi001, CENSOi007, CENSOi003.

CENSOi001	Exon 45-50 deletion	https://hpscreg.eu/cell-line/CENSOi001-B
CENSOi007	Exon 52 deletion	https://hpscreg.eu/cell-line/CENSOi007-A
CENSOi003	Exon 46-52 deletion	https://hpscreg.eu/cell-line/CENSOi003-B

We also used an immortalised myoblast line, AB1190, as a positive control for multiple experiments. The line was kindly provided by the Institut Myologie in Paris and is a line derived from paravertebral muscle of a 16-year-old male.

Work with human cells was conducted under the approval of the NHS Health Research Authority Research Ethics Committee reference no. 13/LO/1826; IRAS project ID no. 141100, study title: The use of cells as a model system to study pathogenesis and therapeutic strategies for Neuromuscular Disorders.

3.2 HUMAN IPS CELL CULTURES

3.2.1 HUMAN IPS CELL MONOLAYER CULTURES

All iPS cell lines used in this project were cultured in feeder-free conditions, maintained on Matrigel® (Corning) or Geltrex™ (Thermofisher) coated cell-culture plates in iPS cell maintenance media (TeSR™-E8, mTeSR™1 or mTeSR™Plus media- StemCell Technologies™) and incubated at 37 ℃ with 5% CO_2 and 5% O_2. When the colonies reach around 80% confluence, or their centre becomes bright under the microscope (indicating that cells are no longer a monolayer in the centre), they were passaged. First, to ensure a pure cell population any differentiating cells/sections of the colonies- identified by their morphology, were scraped off using a plastic pipette tip under a stereo microscope. Following that, media was aspirated, the cells were washed once with DPBS (Dulbecco's phosphate-buffered saline), incubated at 37℃ in 1 ml/10 cm^2 Gentle Cell

Dissociation Reagent (GCDR, StemCell Technologies™) for 4-6 minutes -until the edges of the colonies detach and round up, and then the GCDR was aspirated. Immediately, 1 ml/10 cm^2 of maintenance media was added, the colonies were detached using a cell scraper and collected using a 2 ml serological pipette after being pipetted up and down 3-4 time to break up the colonies to an appropriate size, and finally decanted in new wells (Matrigel™ coated and containing warmed-up maintenance media). Typical splitting ratio (area: area) was between 1:3 and 1:10, depending on the initial confluence, cell line growth rate and downstream application. To cryostore iPS cells, they were collected post-detachment in a 15 ml centrifuge tube and spun down at 800 RPM (approximately 135 RCF) for 5 minutes at room temperature. The supernatant was aspirated, then cells were resuspended in maintenance media supplemented with 10% dimethyl sulfoxide (DMSO) in a cryovial and stored for 24 hours in -80℃ in a freezing container, and finally stored in cryogenic liquid nitrogen storage vessels.

In the case of excessive spontaneous differentiation of an iPS cell culture, intact colonies or areas of colonies were identified- based on morphology- and scored using a sterile scalpel under a stereo microscope. A pipette with a P1000 tip was used to detach the desired sections, collected, and then immediately decanted in new wells.

For applications such as electroporation or nucleofection for CRISPR/Cas9 gene editing, differentiation into various lineages or single cell sorting, single cell iPS cell splitting was performed using TrypLE™ Express

Enzyme (Gibco™) or Accutase™ (StemCell Technologies™) dissociation solution. Briefly, media was aspirated, cells were washed using DPBS, incubated in 0.5 ml/ 10 cm^2 dissociation solution for 4-6 mins at 37℃, then media was added (2.5 ml/ 10 cm^2) and the colonies were dissociated into single cells by pipetting up and down up to 10 times using a 2 ml serological pipette. Cells were then collected in a 15ml centrifuge tube and centrifuged at 1200 RPM (approximately 300 RCF). Finally, cells are resuspended in respective buffer or media using a P1000 and single cell state is assessed using a counting chamber on an inverted microscope.

3.3 Derivation and Culturing of Myogenic Cells from iPS Cells

Muscle cells used were all iPS cell-derived cultures using one of two methods. The first is transgene based which utilises a series of media driving the commitment of the cells towards the myogenic lineage and the transduction lentiviral vector encoding *Myod1* cDNA fused with oestrogen receptor (MyoD-ER(T)) which allows inducible translocation of the protein from the cytoplasm to the nucleus upon tamoxifen administration. The second approach is transgene free, which utilises small molecules and transcription factors based on recapitulating skeletal muscle embryonic development is commercially available through Myocea (formerly, Genea BioCells).

3.3.1 TRANSGENE-BASED MYOGENIC DIFFERENTIATION OF HUMAN IPS CELLS

Human iPS cells were differentiated into inducible myogenic cells (HIDEMs) as previously described (Maffioletti *et al.*, 2015). In summary, HIDEMs were cultured in HIDEM proliferation medium which consists of MegaCell™ DMEM (Sigma-Aldrich™) containing 5% vol./vol. Foetal Bovine Serum (FBS), 2 mM GlutaMAX™ (Gibco™), 100 IU penicillin, 0.1 mg streptomycin, 1% vol/vol non-essential amino acids (Gibco™), 0.05 mM 2- mercaptoethanol (Gibco™), and 5 ng/ml ß-FGF (Peprotech)) at 37℃ with 5% CO_2 and 5% O_2. For terminal differentiation, HIDEM proliferation media was aspirated, cells were washed once with DPBS, covered with Trypsin-EDTA (Sigma-Aldrich™) at 37℃ until cells detach, for up to 5 minutes, and then the enzyme was deactivated using HIDEM proliferation medium. After the cells were centrifuged at 1200 RPM (approximately 300 RCF) and resuspended in HIDEM proliferation media, 25,000 cells /cm^2 were seeded onto Matrigel® coated cell-culture plate and kept in culture until at least 90% confluence is achieved, usually within two days. To initiate differentiation, replace medium with fresh HIDEM proliferation media supplemented with 1 µM 4-OH tamoxifen (Sigma-Aldrich™), 24 hours later, the medium was replaced with HIDEM differentiation media (composed of Dulbecco's Modified Eagle's Medium (Gibco™) with 2% vol./vol. Horse Serum (Thermo Fisher™), 2 mM GlutaMAX™ (Gibco™), 100 IU penicillin and 0.1 mg streptomycin) supplemented with 1 µM 4-OH tamoxifen (Sigma-Aldrich™). Cells

were inspected daily and HIDEM differentiation medium was replaced every 48 hours. Cultures typically differentiated in 4-12 days.

3.3.2 *Transgene-Free Myogenic Differentiation of Human iPS Cells*

Human iPS cells were differentiated into myogenic progenitor cells using the method reported in Caron and colleagues (Caron *et al.*, 2016), however, a homemade version was used briefly but ultimately media was switched to the commercially available media for better standardisation, given the complex and extensive number of components in the media. Differentiation occurs over 3 stages, all of which the cells were cultured at 37℃ with 5% CO_2 and 5% O_2. The first stage is 6-10 days long, during which iPS cells were split into single cells and seeded at 5,000 cells/cm^2 on Matrigel® coated cell-culture plate with induction media (containing 5% vol./vol. horse serum, 50 µg/ml fetuin, 10 ng/ml hR-EGF, 1 ng/ml hR- βFGF, 10 µg/ml insulin, 0.4 µg/µl dexamethasone, 10 µM ROCK inhibitor (Y-27632), 3 µM CHIR99021, 2 µM TGFβ- inhibitor (SB 431542), 200 µM ascorbic acid, 100 IU penicillin and 0.1 mg streptomycin) and kept in culture until 100% confluence was reached. The second stage cells are split again using the same method and seeded at 5,000 cells/cm^2 on Matrigel® coated cell-culture plate with skeletal myoblast media (containing 5% vol./vol. horse serum, 50 µg/ml fetuin, 10 ng/ml hR-EGF, 10 µg/ml insulin, 0.4 µg/µl dexamethasone, 10 µM ROCK inhibitor (Y-27632), 200 µM ascorbic acid, 2 µM TGFβ- inhibitor (SB 431542), 20 ng/ml HGF, 20 ng/ml β-FGF, 10 ng/ml IGF1, 10 ng/ml oncostatin, 10

ng/ml PDGF, 100 IU penicillin and 0.1 mg streptomycin), after 6-10 days, once cells reached 80% confluence, they were ready for terminal differentiation, the cells were also expanded (without Matrigel® coating and cryostored at this stage. For terminal differentiation, cells were seeded at 80%-100% on Matrigel® coated cell-culture plate and maintained in skeletal muscle differentiation medium (composed of Dulbecco's Modified Eagle's Medium (Gibco™) with 2% vol./vol. Horse Serum (Thermo Fisher™), 2 mM GlutaMAX™ (Gibco™), 100 IU penicillin and 0.1 mg streptomycin and 2 µM TGFβ- inhibitor (SB 431542)) for 4-12 days.

3.4 Developing Dystrophin Reporting Cell Lines

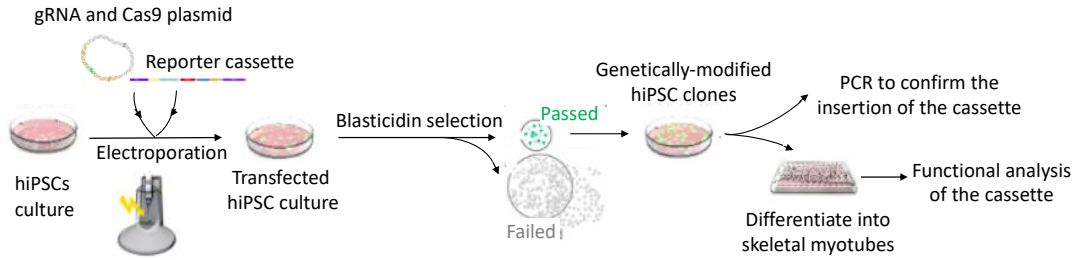

Figure 3-1. *CRISPR/Cas9 gene-editing workflow.*
Human iPS cell clones were disaggregated into single cells and transfected with the reporter cassette, Cas9-encoding plasmid, and the guide RNA (gRNA) by electroporation. Transfected human iPS cells were treated with blasticidin to select the genetically modified human iPS cells. Next, PCR was used to confirm the insertion of the cassette, while its functionality was investigated upon differentiation of selected human iPS cells into myotubes. Figure generated by Mengmeng Zhu, MRes syudent under my and Prof. Tedesco's supervision.

3.4.1 *Design of the Dystrophin Reporter Cassette*

To allow for the precise detection of dystrophin, a multiple reporter cassette that is inserted using Clustered Regularly Interspersed Short Palindromic Repeats (CRISPR)/Cas9 gene-editing technology just before the stop codon of the dystrophin gene is used. This approach allows for the detection of almost all dystrophin isoforms (naturally occurring and the resulting forms of therapies such as exon skipping) overcoming the inefficiencies of antibody detection of dystrophin. The presence of mCherry, a fluorescence marker, enables real-time live detection of dystrophin, NanoLuciferase enables live quantitative detection and finally, the interchangeable labeling sequence, HaloTag, can be used for multiple applications, including flagging different dystrophin isoforms in western blotting. The cassette is flanked by two 500bp homologous recombination arms, totalling 1000bp, that are needed for the Homology Directed Repair (HDR) mechanism utilised. Moreover, there is an excisable blasticidin resistance sequence, driven by a Phosphoglycerate kinase 1 (PGK) promoter, that is required for clone selection following gene editing. This sequence is flanked by two cis-placed Flippase Recognition Targets (FRT) with the same direction that allow for subsequent excision using a Flippase recombinase (FLP) enzyme.

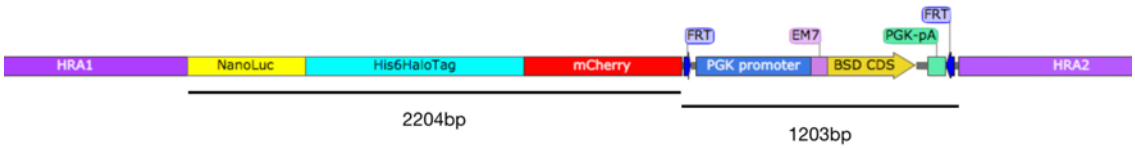

Figure 3-2. *The reporter cassette layout.*
Total size is 4407 bp; 5437 bp including homologous recombination arms.

Initial conception, design and generation of the first cassette was done by Prof. F. S. Tedesco and Dr. H. Hoshiya with input from Dr L. A. Moyle in collaboration with Prof. K.s Anastassiadis and Dr. K. Neumann at the University of Dresden (Germany).

3.4.2 Gene Editing and Genomic Analyses

3.4.2.1 Gene Editing

3.4.2.1.1 Original CRISPR/Cas9 Editing Approach

Initially, a “triple electroporation” approach to carry out CRISPR/Cas9 gene-editing was used. The components delivered were:

1. A pCAGGs plasmid containing the Cas9 enzyme sequence, driven by a cytomegalovirus (CMV) enhancer. The sequence contains a puromycin resistance sequence for selection of bacterial clones.
2. A pBR322 plasmid containing the guide RNA sequence. The plasmid includes an ampicillin resistance for selection of bacterial clones.
3. A pUC57 plasmid containing the hDMD reporter cassette.

These plasmids are first expanded in transformed XL gold *E. coli*, through a Maxiprep, harvested and purified using the phenol chloroform method. Prior to the purification, and for the hDMD reporter cassette only, it is linearised out of the plasmid backbone through a restriction digestion using MluI & NheI. A mixture of 25 µg hDMD reporter cassette, 15 µg gRNA and 5 µg Cas9 sequence is prepared in a total of 800 µl of PBS.

Induced pluripotent stem cells (iPS cells) were cultured in feeder-free conditions on Matrigel® coated plates in mTeSR™1 media were then harvested as single cells and 5 million are spun down, washed and gently resuspended in the prepared PBS mixture. The mixture was then quickly transferred to a Bio-Rad 4 mm cuvette, inserted in a Bio-Rad Pulser Xcell electroporator and pulsed with an exponential protocol: 300 V, 250 µF cap, and infinite resistance.

The solution was removed from the cuvette with a plastic pipette and transferred dropwise into a pre-prepared 15 ml falcon tube containing 2 ml of mTesR1 medium and Rho-associated protein kinase inhibitor Y-27632 (ROCKi) 10 µM. The solution is then divided onto three Matrigel®-coated 10cm dishes (each containing 7 ml of the same media) ending up with three different concentrations (2 million cells/dish, 1 million cells/dish and 500,000 cells/dish).

Cells were left to grow for 24 hours then media is changed daily with mTeSR™1 supplemented with 10 µM ROCKi and 2 ug/ml blasticidin, which will result in sudden death of multiple colonies in 3-5 days. Once clusters of 5-10 cells are seen, start reducing ROCKi inhibitor in a progressive manner over the course of 3 days. Approximately 2 weeks post electroporation, clones are big enough to pick and isolate in wells of 48-well plates. Clones are then expanded, frozen and a pellet is taken for analysis.

3.4.2.1.2 New CRISPR/Cas9 Editing Approach

After trial of the initial approach, a new approach was adopted in efforts to increase efficiency of gene editing cell lines.

Instead of using a DNA triple electroporation approach, in this approach a small guide ribonucleoprotein (sgRNP) complex is delivered. It is composed of a modified small guide ribonucleic acid (sgRNA) and a Cas9 protein (EnGen Spy Cas9 with a Nuclear Localisation Signal from New England Biosciences), that was pre-incubated for 10 minutes in TE buffer to achieve the binding and complex formation. The template cassette was still delivered in the same format as the initial method, in a linearised form.

To set-up this system, new sgRNAs were designed using the Benchling design tool that identified target sites within 20 bp of target insertion site (with an NGG protospacer adjacent motif, which is required for *S. Pyogenes* Cas9 that is being used in this experiment). Three 20 bp binding sequences were selected for sgRNA. The choice was based on two more criteria, in addition to target location, the on-target efficiency score and off-target efficiency score. These scores are calculated based on algorithms developed by (Hsu *et al.*, 2013; Doench *et al.*, 2014). The modified 100mer sgRNAs were ordered from Synthego.

An optimisation experiment for the cutting efficiency of the different sgRNAs and combinations of the three of them was conducted to determine the optimal conditions for use in the actual knock-in experiment. It is worth mentioning that for this new approach the AMAXA 4D nucleofector was used as it has been extensively used and proven effective for iPS cell electroporation and subsequent viability.

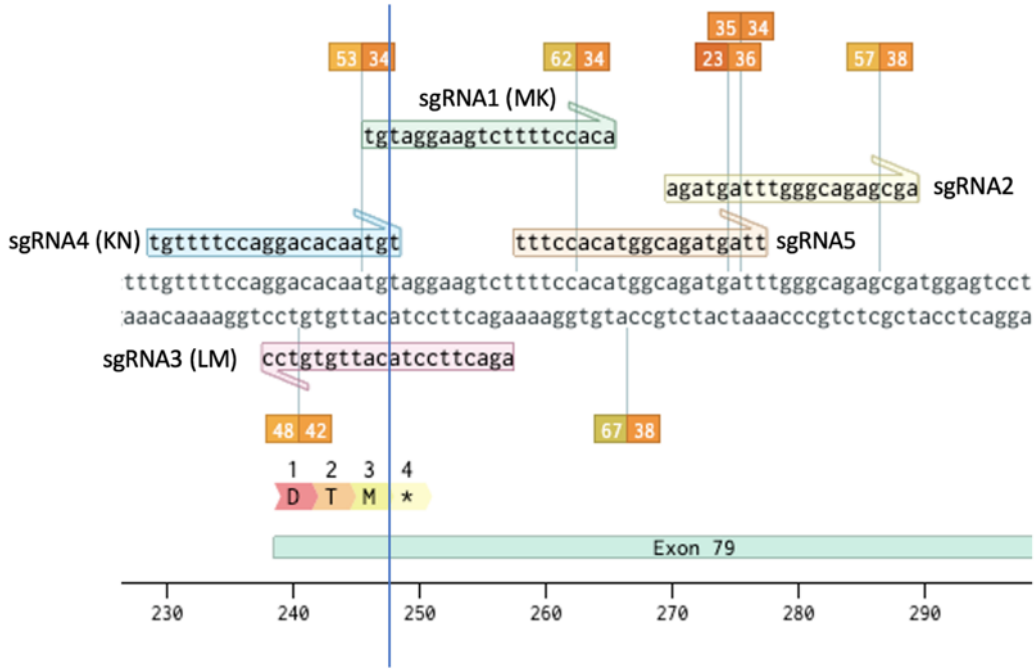

Figure 3-3. *Top 5 sgRNAs for CRISPR/Cas9 gene-editing aligned on Exon79. The blue line and asterisk denote the stop codon. sgRNA 2, 3 and 5 were used for the initial trial experiments using the triple electroporation method 4.1.1. sgRNA 1, 2 and 3 were used in optimisation and subsequent experiments from 4.1.2 onwards. The two boxes on every gRNA denote the on-target efficiency score and the off-target efficiency score, respectively.*

iPS cells were cultured without any antibiotics in the media at least 48 hours prior to the procedure and 500,000 cells per reaction were harvested as single cells (procedure in 3.2), then after a washing step in Phosphate Buffered Saline (PBS), were resuspended in 20 µl of Lonza's P3 buffer and mixed with 5 µl of the gRNP complex mix (containing 300 pmol of sgRNA and 40 pmol of Cas9). Using the sgRNA cocktail at 300 pmol with double concentration of Cas9 80 pmol was also tested. The final mixture was then transferred to a well of a

Nucleocuvette™ strip, inserted into the nucleofector and pulsed with programme CA137. 70 µl of pre-warmed media (mTeSR1 supplemented with 1:5 CloneR™ (StemCell Technologies™) which is a ROCK inhibitor, without any antibiotics) was added to each well, left for 10 mins at room temperature and then finally the total 95 µl volume was seeded onto a Matrigel® coated well of a 6-well plate, with 2 ml of pre-warmed mTeSR1 medium supplemented with 1:5 CloneR™ (StemCell Technologies™). Cells were then cultured for 2 weeks with the addition of 10 µM ROCKi that was progressively reduced over the course of 3 days as colonies for. 100 IU penicillin, 0.1 mg streptomycin were added to media starting from first medium change 24 hours after nucleofection. Finally, cells were pelleted for DNA extraction. A PCR reaction using primers spanning 2500 bp upstream and 2500 bp downstream the target region was carried out, the PCR reaction product was then purified using the QIAquick® PCR purification kit (Qiagen®) and outsourced for sanger sequencing using the same primers used for the PCR reaction- Genewiz, Azenta Life Sciences company. Finally, the results of the sequencing were uploaded onto Synthego's Inference of CRISPR Efficiency (ICE) tool to determine the cutting efficiency and optimal combination and concentration of sgRNAs.

After this optimisation experiment, for the actual knock in experiment, the same procedure and steps listed in the paragraph above were used, but 4 µg the linearised form of the cassette in the reaction mix (approximately 0.08 pmol) was also included. Guide RNAs 1, 2 and 4 were used for these experiments (gRNA 3 was excluded for low efficiency) in an individual fashion rather than in cocktails,

to be able to assess the efficiency rate at HDR following cutting reflected through the number of clones obtained from each condition. Experiments were conducted using both the original cassette and a shortened version of the reporter sequence (**Figure 3-4**) including just NanoLuc (and the blasticidin resistance sequence) to further increase the chances of success as the size of the original cassette is very big and is likely to have lower efficiency in HDR.

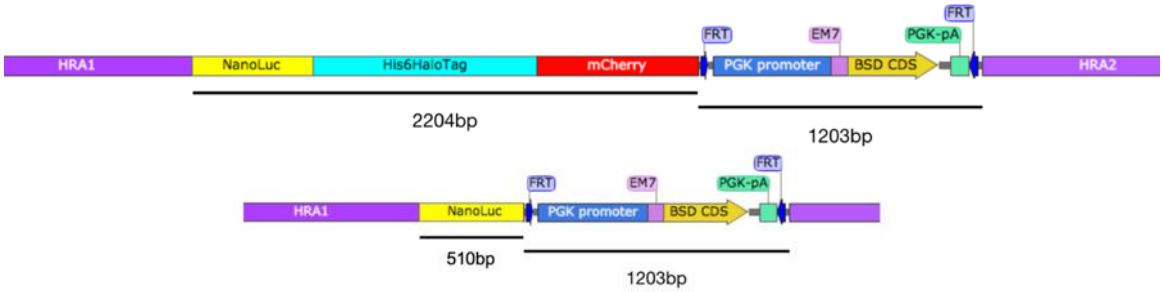

***Figure 3-4**. Layout of original and shortened version of the cassette.*
Top: Layout of the original cassette; flanked by 2 homologous recombination arms (HRA1 & HRA2) 100 0bp each, and including Nanolucifease, HaloTag, mCherry and Blasticidin resistance sequences. Total length of sequence to be inserted is 3407 bp. Bottom: Layout of the new, shortened, version of the cassette; flanked by 2 homologous recombination arms (HRA1 & HRA2) 1000 bp each, and including Nanolucifease and Blasticidin resistance sequences. Total length of sequence to be inserted is 1713 bp.

Finally, to select for colonies that have the cassette successfully integrated into their genome, blasticidin selection starting on approximately day 3 (after the colonies are 2-5 cells in size) was used at a concentration of 2 µM added with every media change. Ten days post nucleofection (1 week post blasticidin selection) colonies were ready to be manually picked using a P200 pipette under

a stereo microscope and isolated each in a different Matrigel® coated wells containing prewarmed mTeSR1 and expanded as normal monolayer iPS cell culture (procedures listed in section 3.2).

3.4.2.1.3 Flp-FRT Excision of BSD Resistance Sequence

Clones that have been edited and include the reporter cassette sequence, verified using PCR and sanger sequencing (detailed in section), were then subjected a flip recombination procedure to excise the blasticidin resistance sequence. 24 hours prior to the procedure, media was aspirated from an 80% confluent 3.5 cm dish of iPS cells and were washed once with DPBS, and 2 ml of pre-warmed fresh mTesR1 media (not containing any antibiotics) is added. The next day, media was aspirated, cells were washed twice with DPBS, and then covered with 500 µl of TrplE express for 5 minutes at 37°C. Then, 2.5 ml of mTeSR1 media (not containing any antibiotics) was added and cells were dissociated into single cells by pipetting up and down 5 times using a P1000 pipette, centrifuged at 1200 RPM (approximately 300 RCF) for 5 minutes, in the meantime, 800 µl of DPBS containing 5 µg of pCAG-Flpo-IRES-puro pDNA was prepared, and then cell pellet was resuspended in the solution. Solution was then transferred to a 4mm gap cuvette and electroporated using the Gene Pulser Xcell from Bio-Rad with exponential protocol, 320 V, 250 µF, 4mm cuvette gap, ∞. The cells suspension was then transferred into 3.2 ml mTeSR1 and distributed into Matrigel® coated 10 cm dishes containing mTeSR1 supplemented with 10 µM Rocki at different densities (1:40 (100 µl of 4 ml total cell suspension seeded to a 10 cm dish), 1:100 and 1:10), discard the rest. 24 hours later, medium is

replaced with fresh mTeSR plus medium containing 100 IU penicillin, 0.1 mg streptomycin and 10 µM ROCK inhibitor (Y-27632). Cells were cultured as standard until used for single cell sorting to isolate clonal cell populations.

3.4.2.1.4 Generating Isogenic Dystrophic Line

To generate the isogenic dystrophic cell line, dystrophin wild type clones that have been edited and acquired the reporter cassette, which have also undergone the flip recombination step to excise the blasticidin resistance sequence were used. An exon 52 deletion was chosen for multiple reasons. First, it is one of the most common deleted exons in DMD patients, secondly, it is amenable to both exon 51 and exon 53 skipping using antisense oligonucleotide therapies (such as Eteplirsen and Golodirsen, both of which are FDA approved) which can theoretically treat up to 24% of patients combined, thirdly, we have a patient iPS cell line, from the MRC Centre for Neuromuscular Diseases Biobank London, with an exon 52 deletion in the lab, which be useful in cross comparison experiments confirming *in vitro* disease phenotype.

Name	5'-Sequence-3'	Target site (from start/end of exon)
sgRNAEx52Upstream1	GCTGAAGAACCCTGATACTA	23 bp 5' start
sgRNAEx52Downstream1	GACCAACAGCCAAGGATATG	243 bp 3' end
sgRNAEx52Downstream2	GTTGCAAAGCATGCATTGAT	347 bp 3' end

Table 3-1. *sgRNA sequences used for CRISPR/Cas9 gene editing experiment to excise exon 52 from the endogenous dystrophin sequence.*

The strategy followed was based on introducing a cocktail of sgRNAs with targets upstream and downstream of the exon, within the intronic sequences, and relying on non-homologous end joining repair mechanism to repair the cut sites without the exon. The method was based on the sgRNP CRISPR/Cas9 gene-editing method detailed in section 3.4.2.1.2. For the sgRNA design, after using the Benchling guide design tool to shortlist several guides within 400 bp upstream and downstream of the exon, I relied on the literature to select 3 guides (one with an upstream and two with downstream targets). The selected guides are listed in **Table 3-1**; sgRNAEx52Upstream1 is quoted in Soblechero-Martin and colleagues (Soblechero-Martín *et al.*, 2021) (also validated in (Al Tanoury *et al.*, 2021)). For the downstream targets sgRNAEx52Downstream1 was used, it is also used in both publications but reported to have a very low efficiency by Al Tanoury and colleagues (Al Tanoury *et al.*, 2021), and sgRNAEx52Downstream2, which was validated by Soblechero-Martin and colleagues.

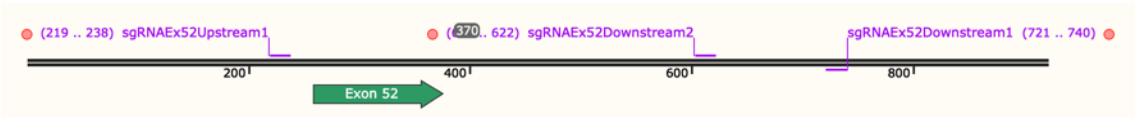

Figure 3-5. *Map of sgRNAs used in exon 52 CRISPR/Cas9 deletion experiments aligned to genomic target locus.*

The target cutting sites of the guides are listed in **Figure 3-5**. The breakpoints are close to the exon sequence, given the combined size of the introns upstream and downstream of the exon totalling over 90 kb. It is worth

noting that the intronic breakpoints vary from patient to patient and are not known in most patients.

3.4.2.2 Single Cell Sorting to Generate Clonal Populations

At different stages of the gene editing procedures, clonal populations were derived through single cell sorting that were then screened and some of which were selected for expansion and downstream applications. For this, edited iPS cell cultures were left to recover for a minimum of two days following the gene editing experiment, culturing media was switched to mTeSR1 media containing 100 IU penicillin, 0.1 mg streptomycin, and on the third day, 10 µM ROCK inhibitor (Y-27632) was added 2 hours before single cell sorting. Immediately prior to sorting, the cells were detached in a single cell manner using TrypLE express as reported in section 3.2, resuspended in mTeSR1 media containing 100 IU penicillin, 0.1 mg streptomycin and CloneR (at manufacturer's recommended concentration). For sorting, the cellenONE® (Cellenion) cell sorter was used, which deposited single cells in multiple 96-well plates coated with Matrigel® and containing mTeSR1 media containing 100 IU penicillin, 0.1 mg streptomycin and CloneR (at manufacturer's recommended concentration). Finally, cells were cultured at 37°C at 5% CO_2 and 5% O_2 using the same culture media for a week, then split and expanded as regular iPS cell cultures reported in section 3.2.

3.4.2.3 Genomic Analyses

3.4.2.3.1 Confirmation of Cell Identity

To confirm the identity of the cell line used to generate gene-edited hDMD report lines and their isogenic Exon 52 deleted progeny, low-pass whole genome sequencing was utilised, provided by the Francis Crick Institute's Advanced Sequencing Science Technology Platform, as an alternative to conventional karyotyping. The samples were submitted as a DNA extract eluted in Tris-HCl (pH 8.5) at 5 ng/µl and was processed using the Nextera Flex for Enrichment kit (currently known as Illumina® DNA prep with enrichment) according to manufacturer's protocol and sequenced at 0.1X coverage, the data was then analysed using a pipeline developed in-house for karyotyping.

To further confirm the identity of the cells, we used a TaqMan™ Copy Number Variation assay for the SRY gene (probe SRY-408), using TERT as an endogenous control according to manufacturer's protocol and the data was visualised using the CopyColour software.

3.4.2.3.2 PCR Confirmation and Sanger Sequencing

DNA is extracted from cell pellets using the Qiagen® Blood and Tissue® kit (eluted in nuclease-treated water) as described earlier in section 3.6.1and used for PCR screening of cassette insertion and BSD sequence excision. Primers are used to check for insertion and screening the entire sequence (including the Homologous Recombination Arms (HRA)).

A set of primers covering the entire sequence from upstream and downstream the sequences homologous to the recombination arms of the

template cassette was generated using the online tool, primer3plus. The parameters used for selecting these primers were: a length between 18 and 22 base pairs, melting temperatures around between 57 °C and 59 °C, spacing of 500 bp between primers on the same strand and an interval of 200 bp between 3' end of a primer to the 3' end of the following primer on the reverse strand. This is to allow using various combination of pairs together in a PCR reaction to amplify a certain area and satisfies the requirements for sanger sequencing.

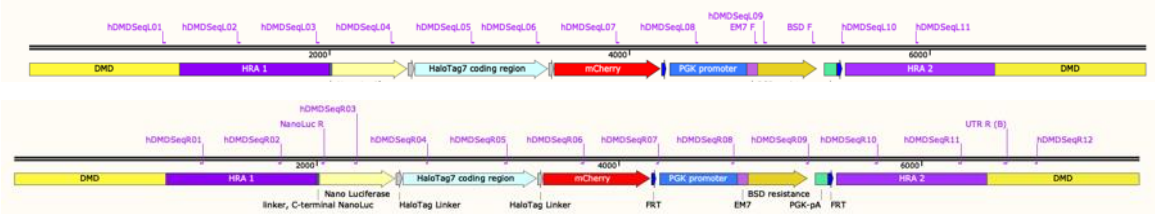

Figure 3-6. *Map of the reporter cassette, after insertion in target locus using CRISPR/Cas9.*

Yellow regions labelled DMD are endogenous genomic sequences, purple regions labelled HRA are endogenous genomic sequences that are homologous to the homology recombination arms of the cassette. The sequence is labelled with binding locations of the primers used for screening using PCR and sanger sequencing. Top panel, forward primers, bottom panel, reverse primers.

Primer name	Primer sequence	Primer binding area
hDMDSeqL01	aaagttgatacatggcaagg	Endogenous sequence 3' (FWD)
hDMDSeqL02	aaatagccacctccaccata	Homologous sequence 3' (FWD)

hDMDSeqL03	caaatggcaaagaaacagag	Homologous sequence 3’ (FWD)
hDMDSeqL04	ggaacggcaacaaaattatc	NanoLuciferase (FWD)
hDMDSeqL05	tgcatttatggagttcatcc	HaloTag (FWD)
hDMDSeqL06	gaatctgctgcaagaagaca	HaloTag (FWD)
hDMDSeqL07	ccgtaatgcagaagaagacc	mCherry (FWD)
hDMDSeqL08	cctcccctagtcaggaagtt	PGK promoter (FWD)
EM7 F	gacaaggtgaggaactaaacc	EM7 sequence (FWD)
hDMDSeqL09	aaagagcaacggctacaatc	BSD resistance sequence (FWD)
BSD F	ttgctgccctctggttatgt	BSD resistance sequence (FWD)
hDMDSeqL10	ttcgaattcaagcttgtgag	FRT & Linker 5’ (FWD)
hDMDSeqL11	cacacacatacacacacacaca	Homologous sequence 5’ (FWD)
hDMDSeqR01	tgacccacatgtattgtatgc	Homologous sequence 3’ (REV)
hDMDSeqR02	ggatcaggctactggtgttt	Homologous sequence 3’ (REV)
NanoLuc R	ccagtccccaacgaaatcttc	NanoLuciferase (REV)
hDMDSeqR03	cagggtacaccaccttaaaaa	NanoLuciferase (REV)
hDMDSeqR04	aacatgcgggatgatgtt	HaloTag (REV)
hDMDSeqR05	gcagccagtccatgtattct	HaloTag (REV)
hDMDSeqR06	cttcaagtagtcggggatgt	mCherry (REV)
hDMDSeqR07	cgcgtttaaactgaagttcc	FRT & Linker 3’ (REV)
hDMDSeqR08	atgaggaagaggagaacagc	PGK promoter (REV)
hDMDSeqR09	ttagccctcccacacataac	BSD resistance sequence (REV)
hDMDSeqR10	gacgtgtaaaacctgccatt	Homologous sequence 5’ (REV)
hDMDSeqR11	ggtaagcctggatgactgac	Homologous sequence 5’ (REV)

UTR R (B)	ccttctgattgatttccactgaagc	Endogenous sequence 5’ (REV)
hDMDSeqR12	gactgtgagaagagggcataa	Endogenous sequence 5’ (REV)

Table 3-2. *List of primers used in PCR reactions and sanger sequencing*.

The qRT-PCR reactions are carried out using LA Taq enzyme kit by Takara bio. LA Taq is a long-range high fidelity (with a DNA-proofreading polymerase with 3’ to 5’ endonuclease activity), which produces PCR products suitable for sequencing, if needed. The parameters used for the PCR reactions are listed in **Table 3-3** and **Table 3-4** according to manufacturer’s guide. Bio-Rad T100 thermal cycler was used for the reactions. Afterwards, an aliquot of each reaction is loaded using a 6X loading dye (from NEB) along with a ladder (Hyperladder 1kb, Bioline ladder 100bp or Bio-Rad 1kb or 100bp ladder), then run on an agarose gel (between 0.5% and 2%) with CyberSafe dye) using a gel electrophoresis set-up and powerpack (consort E844) delivering between 100V and 140 V. Finally visualised under UV light using a Bio-Rad Chemidoc MP.

Component	Volume
10X Buffer	5 µl
dNTPs (2.5mM)	8 µl

F primer (20uM)	0.5 µl
R primer (20uM)	0.5 µl
LA Taq enzyme	0.5 µl
DNA (5-150 ng) + Nuclease-free H_2O	35.5 µl
Total	50 µl

Table 3-3. *Component mix for PCR using LA Taq.*

Step	Temperature (°C)	Duration (mins)
Initial denaturation	94	5:00
Denaturation	98	0:20
Annealing	58	0:30
Extension	68	(0.5-1min/ Kb)
Cycles	Back to denaturation step	34X- 38X (if product to be sequenced)
Final elongation	72	10:00
Storage	4	∞

Table 3-4. *Thermocycler steps for PCR reactions.*

PCR products which were then selected for sanger sequencing underwent a purification procedure using the Monarch® PCR & DNA Cleanup Kit (NEB) according to manufacturer's instructions and the DNA is eluted in nuclease-free water at a concentration between 10 µM and 50 µM. The sanger sequencing procedures was outsourced to Genewiz (Azenta life sciences). Obtained reads

were then aligned to the expected sequence using the SnapGene software to confirm gene editing and ascertain whether there were any mutations.

3.4.2.3.3 Southern Blotting

Selected clones are screened using southern blotting to check for any off target inserts and any truncations that might have arisen during the homology directed repair (HDR) process. Southern blotting is performed primarily by our collaborator, Professor Konstantinos Anastassiadis in TU Dresden, Germany. In brief, extracted total genomic DNA was fragmented using a single or combination of restriction enzymes (in this case, the restriction enzymes used were NcoI, BamHI and NdeI), run on an agarose gel using electrophoresis and then transferred to a membrane. The membrane is then exposed to radioactively labelled unique nucleic acid probes which hybridise to complementary DNA sequence and visualised using an x-ray film.

3.4.3 FUNCTIONAL ASSESSMENT OF THE DIFFERENT REPORTER SEQUENCES

3.4.3.1 NanoLuciferase Expression Analysis

Cells were cultured according to protocol described above on 96-well plates with clear bottoms and imaged with a Tecan Spark Microplate reader every day from day 5 to day 8 of differentiation (cultured in skeletal muscle differentiation medium (composed of Dulbecco's Modified Eagle's Medium (Gibco™) with 2% vol./vol. Horse Serum (Thermo Fisher™), 2 mM GlutaMAX™ (Gibco™), 100 IU penicillin and 0.1 mg streptomycin and 2 μM TGFβ- inhibitor

(SB 431542)), 1 plate for every day of measuring was used (as plates must be trashed after being exposed to non-sterile conditions) that had 3 technical repeat wells. To induce luminescence, Nano-Glo® Luciferase Assay System was used/added according to manufacturer's recommendations.

To correlate the levels protein production levels of dystrophin and the luminescence level measured by the plate reader. Cultures were stained and imaged (following the procedures in sections 3.7 and 3.8). The antibody used to visualise dystrophin is Mandys106 (Sigma Aldrich). Finally, the number of nuclei in dystrophin positive cells was manually quantified using FIJI software.

3.4.3.2 mCherry Expression Analysis

The obtained clones were differentiated using the transgene-free Myocea Skeletal Muscle Kit as described in section 3.3.2. Briefly, single cell iPS cells are cultured on Matrigel® coated wells in skeletal muscle induction media for 10 days achieving satellite cell-like cells, then media is switched to myoblast medium for 1 week, after which point, cells can be expanded, frozen and differentiated if needed. Finally, to terminally differentiate into myotubes, cells (at 90-100% confluence) are cultured for 5-10 days in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Foetal Bovine Serum (FBS) on Matrigel® coated wells.

The cultures were checked under a fluorescence microscope for mCherry (red emission signal, under a 546 nm filter cube) to visualise Dystrophin protein. In addition to that, immunofluorescence staining for mCherry protein in these

cultured cells is carried out to determine whether the protein is present. Cell cultures are washed once with PBS, fixed with 4% paraformaldehyde (PFA) for 5 minutes, then washed again with PBS. Cultures are then permeabilised for 30 mins in PBS with 0.2% Triton and 1% Bovine Serum Albumin (BSA), blocked in appropriate serum stock (added to permeabilising buffer to 10% final conc.) for 30 minutes and then left overnight at 4 °C in the primary antibody mixture (antibody diluted according to manufacturer's suggestion). The wells are then washed in PBS for 5 mins, incubated in secondary antibody mixture (including Hoechst) for 1 hour, washed in PBS with 2% triton and finally covered in PBS to be imaged under an inverted microscope.

3.4.3.3 HaloTag Expression Analysis

The differentiated cultures used for the mCherry expression analysis were also used to assess the expression and function of the HaloTag sequence in a live-culture setting. Janelia Fluor® 646 HaloTag® Ligand was used according to manufacturer's instructions. In brief, cells were incubated with warm culture media containing 200 nm of the ligand for a period between 2 and 12 hours at 37 °C. Cells are then washed with DPBS once and then warmed-up media without phenol red was added back to the wells. Cultures were then inspected and imaged under an inverted (EVOS M5000 Imaging System) or a confocal microscope.

3.5 Production of 3D Artificial Muscle Constructs

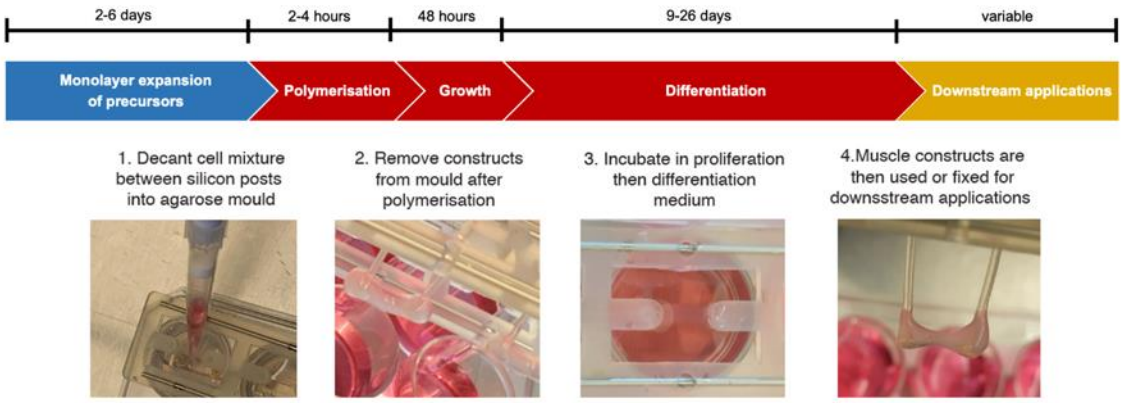

Figure 3-7. *Outline of production of human iPS cell-derived 3D artificial muscle constructs.*

Top panel: Timeline of making 3D muscle constructs in black, with the corresponding key stages of the protocol in the flow chart below. Bottom panel: pictures of the key steps of the protocol showing polymerisation, culturing of constructs and macro morphology at the end of differentiation. Adapted from (Khedr et al., 2022).

3.5.1 Single Lineage Artificial Skeletal Muscle Constructs

The procedures and methods used to produce and culture 3D muscle constructs have been previously reported and detailed by our lab (Maffioletti *et al.*, 2018) and are protocolised in (Khedr *et al.*, 2022)- under review. Both transgene free and transgene-based iPS cell-derived myogenic cells were utilised in the experiments for this thesis.

Myogenic cells (HIDEMs and iPS cell-derived myoblasts), at no more than 90% confluence, were treated with 10 µM ROCK inhibitor (Y-27632) for two hours prior to detachment using TrypLE express. For every 3D muscle construct, 10^6 cells were suspended in a mixture of 10% Matrigel®, human recombinant fibrinogen (3.5 mg/ml), thrombin (3 IU/ml), and inactivated respective media

(HIDEM proliferation media containing heat-inactivated FBS or heat-inactivated Myocea Myoblast media; heat inactivation at 56 °C for 30 mins) in a total volume of 120 µl. The mixture was quickly decanted in 2% autoclaved UltraPure™ Agarose (Thermo Fisher™) moulds fitted with silicone posts (EHT Technologies GmbH, Hamburg) (**Figure 3-7**) and incubated for 2 hours at 37 °C with 5% CO_2. After that, the wells were filled with 1 ml of warmed up media, incubated again for 5 minutes and then the posts with the constructs were replaced in a new plate containing proliferation media (HIDEM proliferation media and Myocea myoblast media, for transgene based and transgene free constructs, respectively) supplemented with 33 µg/µl Aprotinin daily at 37 °C with 5% CO_2 and atmospheric oxygen levels for 48 hours. Muscle morphology is shown in **Figure 3-7**.

To induce terminal differentiation, HIDEM-derived 3D muscle constructs were then cultured in fresh HIDEM proliferation medium supplemented with 1 µM 4-OH Tamoxifen, this is day 0 of differentiation, and 24 hours later, the constructs were then moved to HIDEM differentiation medium supplemented with 1 µM 4-OH Tamoxifen for another 24 hours, after which, media was replaced every other day with fresh HIDEM differentiation media. Culturing media was supplemented with fresh 33 µg/µl Aprotinin every other day from the day of construct polymerisation. For transgene-free derived 3D muscle constructs, medium was changed to skeletal muscle differentiation medium (composed of Dulbecco's Modified Eagle's Medium (Gibco™) with 2% vol./vol. Horse Serum (Thermo Fisher™), 2 mM GlutaMAX™ (Gibco™), 100 IU penicillin and 0.1 mg streptomycin and 2 µM TGFβ- inhibitor (SB 431542)), this is day 0 of

differentiation and media was changed every other day and was supplemented with 33 µg/µl Aprotinin daily.

3.5.2 Bi-Lineage Artificial Skeletal Muscle Construct

For bilineage skeletal muscle constructs, neurospheres derived using the methods detailed in the protocol by (Stacpoole *et al.*, 2011) were used and expanded in neurospheres chemically defined medium (composed of 1:1 Iscove's Modified Dulbecco's Medium and DMEM/F12 (Gibco™) supplemented with 0.5% weight/volume bovine serum albumin (BSA) (Thermo Fisher™), 1% lipids, 100 IU penicillin, 0.1 mg streptomycin and 7 µg/ml human insulin, 15 µg/ml transferrin and 450 µM monothioglycerol) supplemented with 20 ng/ml β-FGF and 5 µg/ml heparin.

To make the constructs the same procedure for single lineage transgene based (HIDEM) artificial skeletal muscle constructs was followed but using inactivated bilineage medium (1:1 HIDEM proliferation medium with heat-inactivated FBS and neuroshperes chemically defined medium) instead. After 48 hours, on day 0, neurospheres suspended in 5 µl fibrinogen were added to the constructs and polymerised by adding 7 µl of medium and left at RT for 3 minutes, following that the constructs were cultured in bilineage media supplemented with 1 µM 4-OH Tamoxifen, 5 µg/ml heparin and 0.1 µM retinoic acid. On day 1 media was switched to bilineage media supplemented with 1 µM 4-OH Tamoxifen and 0.1 µM retinoic acid. From day 2 onwards constructs were cultured in bilineage media supplemented with 0.1 µM retinoic acid. Culturing media was

supplemented with fresh 33 µg/µl Aprotinin every other day from the day of construct polymerisation.

3.6 MOLECULAR BIOLOGY TECHNIQUES

3.6.1 DNA AND RNA EXTRACTION, CDNA SYNTHESIS, AND REAL-TIME PCR

For DNA extraction, the column based DNeasy® Blood & Tissue Kit (Qiagen) was used according to manufacturer's instructions. Cells were washed with DPBS and detached using the appropriate dissociation reagent, pelleted, and stored at -80℃ or used immediately. Following the extraction, DNA was eluted in Nuclease-Free water (Invitrogen) between -20℃ and 4℃ for use in downstream applications.

For RNA extraction, TRIzol™ (Invitrogen™) based method was used. In summary, cells were washed with DPBS and detached by the direct application of TRIzol™ (using approximately 1 ml/10 cm^2), homogenised, and immediately used for RNA extraction or stored at -80℃. First, chloroform was added (1/5 the volume of TRIzol), then the sample was mixed by pipetting up and down, and for 3D muscle constructs the samples were homogenised using the Ultra-Turrax T10 Basic Homogeniser for 30 seconds at max speed, centrifuged at 4 °C at 12,000 RCF for 15 minutes, and finally the upper aqueous phase was carefully isolated using a P200 pipette into a new vial. Afterwards, 500 µl of isopropanol was added and the solution was incubated for 10 minutes at room temperature. Finally, RNA

was pelleted through centrifugation at 12,000 RCF for 15 minutes at 4°C), the pellet was then washed with 1 ml of 75 % ethanol and pelleted again by centrifugation at 7,500 RCF for 5 minutes at 4 °C before resuspension. After being left to dry, RNA pellets were resuspended in 20 µl Nuclease-Free water (Invitrogen). The total RNA yield and purity was assessed with a NanoDrop 1000 (Thermo Fisher™). RNA samples were stored at -80°C.

For RNA expression analyses, variable amounts of RNA (between 200 ng and 1 µg) were retro-transcribed into cDNA using the ImProm-II™ Reverse Transcription Kit (Promega) or High-Capacity RNA-to-cDNA Kit (Applied Biosystems) according to the manufacturer's instructions.

Quantitative real time PCR was used to assess the expression levels of different genes and markers. The reactions were performed using the GoTaq® qPCR Master Mix kit according to the manufacturer's instructions using specific primers listed in **Figure 3-5**, the reactions were run on an Applied Biosystems™ 7500 Fast Real-Time PCR System according to manufacturer's instructions. The reactions were performed in triplicates (three technical repeats) on various numbers of independent experiments. The data points collected were the Ct values, analysed using Microsoft Excel and represented using GraphPad Prism. Data was normalised using an endogenous control using GAPDH or RPLP0 gene expression values and the ΔΔCt method was used (Schmittgen and Livak, 2008), finally presented as a mean value.

Gene	Forward primer Reverse Primer	Comments

MYC	GTAGTGGAAAACCAGCAGCCTC	109 bp
	CGTCGCAGTAGAAATACGGCT	
SOX2	GCGAACCATCTCTGTGGTCTT	109 bp
	CCAACGGTGTCAACCTGCAT	
KLF4	CGGCCAATTTGGGGTTTTGG	129 bp
	GCCAGGTGGCTGCCTCATTA	
OCT3/4	CTGTCTCCGTCACCACTCTG	124 bp
	AAACCCTGGCACAAACTCCA	
GAPDH	GCCCCCATGTTCGTCATGGGTG	199 bp
	CCCGGAGGGGCCATCCACAG	
DMD	TGGAACGCATTTTGGGTTGTT	170 bp, exon 79
	AAAACAATGCGCTGCCTCAAA	
MYH1	GGTCGCATCTCTACGCCAG	140 bp
	TCTGGGCTTCAATTCGCTCC	
MYH2	TCACTTGCTAACAAGGACCTCT	133 bp
	TCAAAGGGCCTATTCTGGGC	
MYH8	AGCAGATAGCAGCGCGAAGA	200 bp
	ACCTCAGCTGGTGCAACACA	
RPLP0	TCTACAACCCTGAAGTG CTTGAT	97 bp
	CAATCTGCAGACAGACACTGG	

Table 3-5. *List of primers used in qRT-PCR reactions.*

MYC: MYC proto-oncogene, bHLH transcription factor; SOX2: SRY (sex determining region Y)-box 2; KLF4: Krüppel-like factor 4; OCT3/4: octamer-binding transcription factor 4, POUF51; GADPH: glyceraldehyde 3-phosphate

dehydrogenase; DMD: dystrophin; MYH: myosin heavy chain; RPLP0: Ribosomal Protein Lateral Stalk Subunit P0.

3.7 Protein Detection and Visualisation

3.7.1 Immunofluorescence Staining

Name	Manufacturer and code	Concentration
mouse anti-dystrophin Dys1	Novocastra; NCL-DYS1	1:25
mouse anti-dystrophin Dys2	Novocastra; NCL-DYS2	1:25
mouse anti-dystrophin Dys3	Novocastra; NCL-DYS3	1:25
Dystrophin	Thermofisher; RB-9024	1:125
Dystrophin	Abcam; Ab15277	1:25
Mandys 106 (dystrophin)	DSHB; MANDYS106(2C6)	1:20
HaloTag	Promega; G9281	1:500
mCherry	Abcam; Ab167453	1:500 (1:100 vale)
Myosin heavy chain, MyHC	DSHB; MF20	1:5
MyoD	Dako; M3512	1:100
Titin	DSHB; TITIN (9D10)	1:100

Table 3-6. *Primary antibodies used in immunofluorescence staining.*

Name	Manufacturer and code	Concentration
Hoechst 33342	Sigma; B2261	1:1000
Donkey anti rabbit 594	Thermofisher; A21207	1:500

Goat anti mouse IgG2b 647	Thermofisher; A21242	1:500
Goat anti mouse IgG2a 546	Thermofisher; A21133	1:500
Goat anti mouse IgG 546	Thermofisher; A11035	1:500
Goat anti mouse IgG1 546	Thermofisher; A21123	1:500
Goat anti mouse IgM 488	Thermofisher; A21042	1:500

Table 3-7. *Secondary antibodies used in immunofluorescence staining.*

3.7.1.1 Monolayer Cell Cultures

At terminal differentiation, monolayer cultures were first washed three times with DPBS for 5 minute each, fixed in 4% PFA at room temperature for 10 minutes and briefly washed twice with DPBS. Fixed cells were either stored in DPBS at 4 °C or used immediately for immunofluorescence staining. To permeabilise, the cells were incubated in Tris-Buffered Saline (TBS) pH 7.4 containing 1% BSA and 0.2% Triton X-100 for 45 minutes at room temperature. For the blocking step, cells were incubated with the permeabilisation buffer containing 10% goat serum. For hybridisation with primary antibodies, cells were incubated with TBS pH 7.4 containing 1% BSA, 0.2% Triton X-100 and the primary antibodies (listed in **Table 3-6**) overnight at 4 °C on a mechanical rocker. The following day, samples were washed thrice with TBS pH 7.4 containing 0.2% Triton X-100 in for 5 minutes. Cells were finally incubated for 1 hour at room temperature with TBS pH 7.4 containing 0.2% Triton X-100 and the appropriate fluorophore conjugated secondary antibodies (listed in **Table 3-7**). Finally, were washed twice with TBS pH 7.4 containing 0.2% Triton X-100, once with TBS and

stored in PBS at 4 °C for imaging and analysis under the microscope. For technique negative control samples, primary antibodies were not added in the primary antibody mix but the rest of the mixture components were used.

3.7.1.2 3D Skeletal Muscle Constructs

At terminal differentiation, 3D muscle constructs were first washed thrice in DPBS for 5 minutes, fixed overnight at 4 °C in 4% PFA, then washed in DPBS thrice again. To permeabilise, the constructs were incubated in Tris-Buffered Saline (TBS) pH 7.4 containing 10% FBS, 1% BSA and 0.5% Triton X-100 for 6 hours at 4°C. Following that, they were incubated with the primary antibodies (listed in **Table 3-6**), suspended in TBS pH7.4 containing, 1% BSA and 0.5% of Triton X-100 overnight at 4°C. After 6 thorough washes of 1 hour each, in TBS, they were incubated with the secondary antibodies (listed in **Table 3-7**), suspended in TBS pH7.4 containing, 1% BSA and 0.5% of Triton X-100 overnight at 4 °C. Note that for 3D muscle constructs' immunofluorescence staining double the concentration listed in the table was used. Finally, the constructs underwent 6 more thorough washes and are stored in DPBS at 4 °C for future use. For technique negative control samples, primary antibodies were not added in the primary antibody mix but the rest of the mixture components were used.

3.7.2 Protein Extraction and Western Blotting

To extract proteins from monolayer cultures and 3D muscle constructs, samples were first washed thrice with ice-cold DPBS, then 200 µl of lysis buffer (10% sodium dodecyl sulphate (SDS), 150 mM NaCl, 50 mM Tris-HCl pH 7.5,

proteases inhibitors (Roche, 04693159001)) was added and mixed well by either pipetting up and down for monolayer cultures or through homogenisation using the Ultra-Turrax T10 Basic Homogeniser for 30 seconds at max speed for 3D muscle constructs, and the mixture was incubated for 5 minutes at room temperature. Afterwards, the solutions were briefly vortexed every 30 seconds for a total of 15 minutes, centrifuged at 13,000 RPM (7,500 RCF) for 15 minutes at 4 °C, supernatant was then isolated, and its concentration determined using the DC protein assay (Bio-Rad laboratories) according to manufacturer's instructions. Prior to loading and running the proteins, the required amounts of protein suspended in loading buffer were denatured by incubation at 70 °C for 10 minutes, running gel (Criterion XT 3-8% Tris-Acetate gradient gel (Bio-Rad, 3450129)) and running buffer (10X TGS (Tris-glysine-SDS): 250 mM Tris (Trizma), 1.92 M glycine, 1% SDS, pH 8,3.) were assembled, and the loading buffer ((Genscript, M00676-10) with 9:1 β-Mercapto-Ethanol (Gibco)) and ladder (HiMark pre-stained protein standard Ladder (Invitrogen, LC5699)) were also set-up. 100 µg of proteins were loaded in each well and the gels were run at 180 V for approximately 1.5 to 2 hours. Proteins were then transferred to a PVDF membrane using the iBlot 2 system (Invitrogen) for 11 minutes at 20 V. Membranes were washed thrice with wash buffer (0.1% Tween in TBS) for 5 minutes on a rocker at room temperature, blocked in blocking buffer (5% non-fat milk in 0.1% Tween in TBS) for two hours on a rocker at room temperature, and hybridised with primary antibody in blocking buffer (antibodies used and their concentrations are listed in **Table 3-8**) overnight at 4 °C. The next day,

membranes were washed thrice in wash buffer for 5 mins on a rocker at room temperature, hybridised with HRP- conjugated secondary antibody by incubating in blocking buffer containing 1:1000 goat anti-mouse IgG antibody (Bio-Rad, 170-6516) for 2 hours at room temperature while rocking. Finally, the membranes were washed thrice using wash buffer, once in distilled water and finally visualised using the ECL immunoblotting detection kit (Amersham) on an Imager 600 (Amersham).

Name	Manufacturer and code	Concentration
mouse anti-dystrophin Dys1	Novocastra; NCL-DYS1	1:250
mouse anti-dystrophin Dys2	Novocastra; NCL-DYS2	1:250
mouse anti-dystrophin Dys3	Novocastra; NCL-DYS3	1:250
Dystrophin	Abcam; Ab15277	1:250
HaloTag	Promega; G9281	1:1000
mouse antiGAPDH	Biogenesis; 4699-9555	1:4000

Table 3-8. *List of antibodies used in western blotting.*

3.8 MICROSCOPY AND IMAGING

3.8.1 *Inverted Fluorescence Microscopy*

For regular monolayer cultures, a Leica DMI6000 inverted fluorescence microscope along with Leica Application Suite (LAS-X) software with various objectives, 4X, 10X and 20X was used. An EVOS M5000 imaging system with the same objectives was also used.

3.8.2 *Confocal Microscopy*

For fixed and stained samples, a Leica confocal SPE2 was used for confocal imaging, along with Leica Application Suite (LAS-X) software with objectives of 4X, 20X and 40X. To image 3D muscle constructs using the confocal microscope, the organoids are mounted on glass slides with a cavity, covered with DAKO Fluorescence Mounting Medium and covered with a coverslip.

For live confocal imaging of cells, a Zeiss LSM 710 run with the Zeiss ZEN 2011 software and using the 10X and 20X objectives was utilised. Cell culture plates with bottoms suitable for imaging, including plate-reader specific 96-well plates with white sides and clear bottoms (Corning®, 3610) and 24 well plates with black sides and no. 1.5 polymer coverslip bottom (Ibidi®, 82426) were used to grow cells. For live imaging, we used media without phenol-red.

3.8.3 *Stereo Microscopy for 3D muscle Constructs*

For live imaging of the 3D muscle constructs, fluorescence stereo microscope; DMI2500 with DFC310 FX digital camera running Leica Application Software (LAS AF) was utilised. Briefly, silicone posts with the 3D constructs

were placed sideways inside a 15 cm dish under the hood while avoiding having the muscle constructs touch the bottom of the dish and placed on the stereo microscope stage. Images were then acquired after adjusting the focus and filter cubes (green fluorescence and brightfield). Keeping the constructs out of the incubator for more than 7-10 minutes was avoided.

3.8.4 Lightsheet Microscopy for 3D Muscle Constructs

Zeiss Lightsheet (ZEISS, Lightsheet Z.1) was used to image and reconstruct the macro structure of the 3D muscle constructs. First, the light sheet fluorescence microscope was set up the objective (5×), the filter cubes and perfusion chamber (filled with PBS till the top of the chamber windows) onto the chamber mount and a syringe was set in the microscope's sample holder disk with syringe adapter ring. Immunofluorescence-stained 3D muscle constructs were then attached to the tip of the needle using a droplet of superglue and attached to the furthest end of the muscle construct. The sample holder disk was then mounted onto the microscope stage and into the upper opening, the construct was located, and imaged using multi-view input field with 100 µm -stack acquisition mode. The acquired image was then exported in CZI file format and imported in Arivis Vision4D software for 3D reconstruction and image and video export.

3.9 AAV Transduction

3D artificial muscle constructs were generated and cultured as detailed in 3.5.1. Initially single lineage transgene based (HIDEM) muscle constructs were used in the initial experiment. AAV9, which has a high skeletal muscle tropism (Zincarelli *et al.*, 2008), containing a plasmid with Green Fluorescence Protein (GFP) under the drive of a CMV promoter was used to transduce the constructs. Seven constructs were used per condition (two different doses/titres). The constructs were transduced with the virus on day 3 of differentiation (measured starting with the day the first tamoxifen pulse is delivered as day 0) with two different doses/titres: high dose titre of 100,000 MOI and low dose titre of 1,000 MOI.

The constructs were then cultured for 12 days and imaged on day 5, 6, 9 and 12 post-transduction. At every time point, the constructs were imaged under a stereo microscope in brightfield and GFP channel, then on an inverted microscope using the GFP channel.

For experiments testing AAV tropism on the 3D muscle constructs, we generated and cultured bilineage constructs as detailed in 0. AAV2, AAV5, AAV8 and AAV9 containing a plasmid with GFP under the drive of a CMV promoter were used. A minimum of four artificial skeletal muscle constructs are infected with each AAV (a total of 23 constructs including uninfected controls) using a titer of 50,000 MOI. Transduction was done on day 4 of differentiation (first tamoxifen pulse delivery day is day 0) and images were acquired on days 3, 6, 8,10, 12, and 14 post-transduction using stereo and inverted microscopes.

3.10 ANTI-SENSE OLIGONUCLEOTIDE (AON) ADMINISTRATION

3.10.1 AONS USED

Antisense oligonucleotides used in this thesis to validate the potential use of the model system to model DMD are two different types but targeting the same exon. The first one is based on the sequence used in the FDA approved drug Golodirsen (Sarepta therapeutics), which is an exon 53 skipping antisense oligonucleotide. The Phosphorodiamidate morpholino oligomer (PMO) was ordered from Genetools, LLC with a 3' end modulus modification (Gene-tools blue fluorescent protein 750 daltons, excitation wavelength: 433, emission wavelength: 465).

The sequence of the molecule is GTTGCCTCCGGTTCTGAAGGTGTTC, the molecular formula is $C_{305}H_{481}N_{138}O_{112}P_{25}$ and the molecular weight is 8647.28 daltons.

The other antisense oligonucleotide is a peptide phosphorodiamidate morpholino oligomer (PPMO) with the same target exon (53), however the sequence of the molecule is confidential. It was acquired from Matthew Wood laboratory in Oxford university. The PMO is conjugated to a cell-penetrated peptide that theoretically increases the uptake of the PMO in muscle tissue. The peptide used in this case is called Pip8b2, consisting of a left arginine-rich motif, hydrophobic core, and right arginine-rich motif.

3.10.2 AON PREPARATION AND ADMINISTRATION

Lypholised PMOs (from Genetools, LLC) were resuspended in non-DEPC-treated water by heating at 65 °C for 30 minutes to a concentration of 0.5 mM. PPMO (peptide-conjugated, from the Wood laboratory, Oxford University) were resuspended in DEPC-treated water at a concentration of 194.5 µM. PMOs were stored at room temperature away from direct light, while PPMOs were stored at 4 °C. The delivery vehicle used for PMO administration was Endo-porter (Genetools, LLC), which is a novel peptide (US patent 7,084,248 B2) that relies on endocytosis-mediated process to channel substances into the cytosol of the cell without damaging the cell membrane. It was delivered as a 1 mM stock solution in 10% polyethylene glycol 1500 MW. Prior to use in PMO delivery, cell toxicity of Endo-porter was tested by incubating differentiating transgene free myogenic cells with it for 72 hours, at which point the cultures were imaged, and then fresh media was added and for a further 5 days to assess effect on cells long term. Skeletal muscle differentiation medium was used for differentiation and Endo-porter concentrations used in testing toxicity are 2 µM, 4 µM and 6 µM.

In all monolayer AON administration experiments, cells were seeded at a very high density of 150,000 cells/cm^2 on Matrigel®or Geltrex™ coated plates, left to recover for 48 hours, after which media was changed to skeletal muscle differentiation media without phenol red and without any antibiotics. For both PMO and PPMO administration, solutions quickly spun down, thoroughly mixed by pipetting up and down and then added immediately to culture media. For PMO administration 6 µM of Endo-porter was added to media prior to the addition of the PMO. PMO concentrations used in various experiments were

500 nM, 1000 nM, 1500 nM and 3000 nM. As for PPMO concentrations, both 1500 nM and 3000 nM (1.5 μM and 3 μM respectively) were used.

3.11 CALCIUM TRANSIENTS: STIMULATION AND MEASUREMENT

Monolayer and 3D muscle constructs (retained on the PDMS post racks) were incubated in Recording buffer (Glucose 10 mM, NaCl 150 mM, KCl 4.25 mM, NaH2PO4 1.25 mM, NaHCO3 4 mM, CaCl2 1.2 mM, MgCl2 1.2 mM, HEPES 10 mM solution with a pH of 7.4) supplemented with Fluo-4AM 5 μM and Pluronic 0.002% (w/v) for 30 minutes at 37°C. Samples were then washed the constructs using recording buffer.

For 3D muscle constructs they were carefully disassembled from the PDMS posts using fine tip forceps, placed in a 10 cm dish and submerged samples in recording buffer making sure most of the construct body is sticking to the bottom of the plate

The Zeiss 880 confocal LSM was set to fluorescence excitation of 488 nm and emission collected at >530 nm and set to acquire as a time series at 33 fps. The CO_2 and temperature settings in the chamber were also set. After selecting a field with aligned myotubes, using a 10× objective, image acquiring was started and caffeine was added to the recoding buffer at a final concentration of 10 mM (making sure to immediately mark the time/captured frame on the microscope software). Caffeine was re-added after fluorescence excitation subsided, for as many times as necessary.

Time series was exported from microscope and open in Fiji software (also known as ImageJ). Areas were selected (each covering a maximum of 10% of a myotube to avoid cross signals from other myotubes) using the “Lasso” tool and data points of “mean gray value” were acquired.

4 RESULTS

4.1 CRISPR/CAS9 GENE EDITING FOR DYSTROPHIN REPORTER HUMAN IPS CELL LINE GENERATION

This section details the CRISPR/Cas9 HDR mediated gene editing of human iPS cell lines to insert the dystrophin reporter cassette and the subsequent screening procedures. The main reasons supporting development of this multifunctional cassette is the difficulty in optimising and complexity in detecting dystrophin protein in human cell cultures, especially iPS cell-derived myogenic cells.

Prior to the development of the dystrophin reporter cell lines, I sought out to characterise dystrophin protein production in monolayer human iPS cell-derived skeletal muscle cells. Dystrophin antibodies have been extensively used and validated in paraffin embedded tissue sections (Kodippili *et al.*, 2014); however, validation and optimisation of protocols in human iPS cell-derived cultures *in vitro* is limited.

UCLi007 HIDEM-derived 8-day differentiated myotubes were used for optimisation of immunofluorescence staining for dystrophin using the most widely used antibodies DYS-1, DYS-2, and DYS-3 in addition to ManDys106 and RB-DYS from Abcam. The results, in **Figure 4-1**, are representative images of different staining conditions. They indicate that the staining of dystrophin using the aforementioned antibodies was not optimal, as the signal intensity matches

that of the secondary antibody only control condition. DYS-1 shows a very faint signal that is visible in all myotubes in the field of view and is thus unreliable.

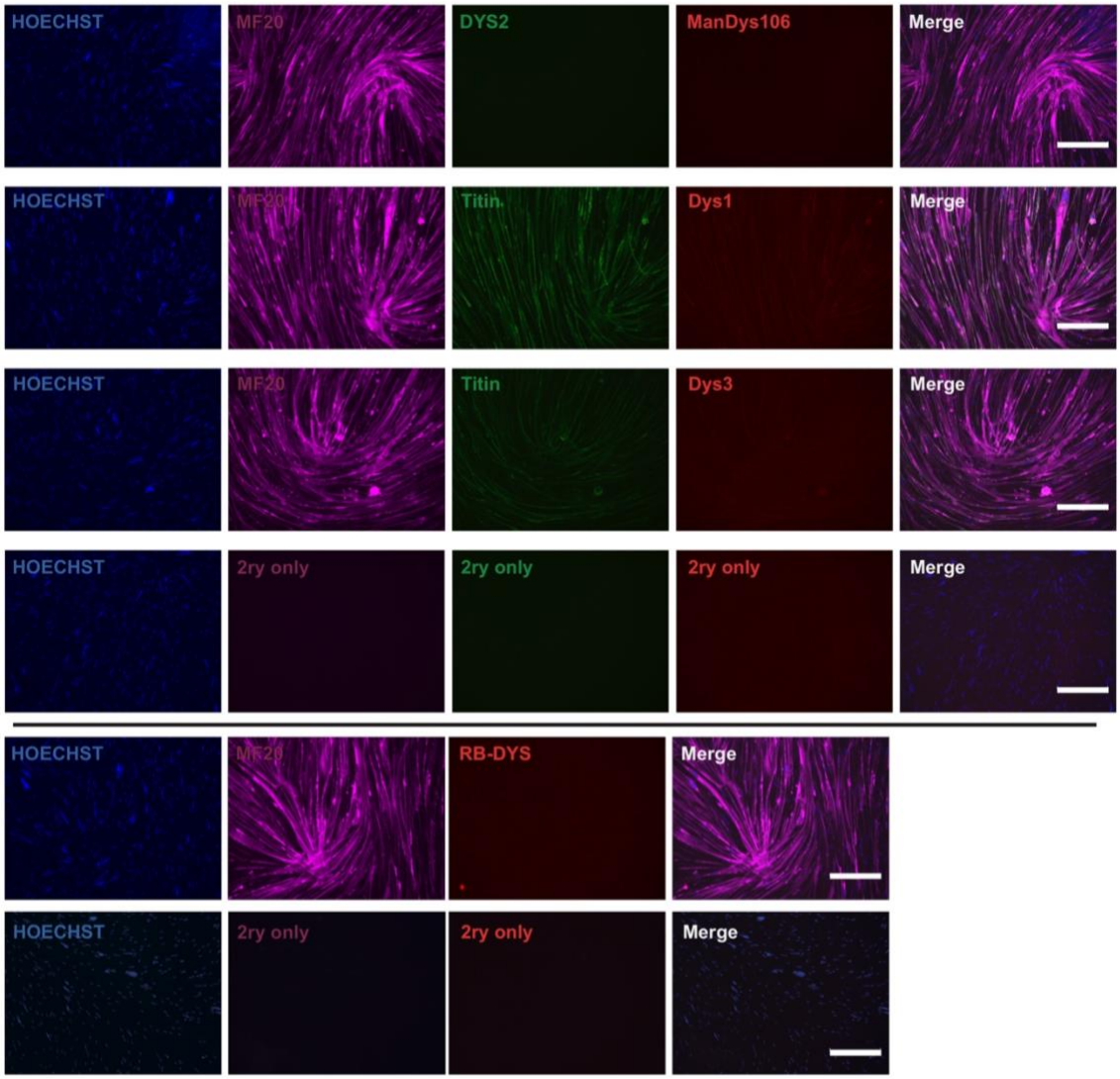

Figure 4-1. *Immunofluorescence staining of iPS cell-derived differentiated skeletal muscles using multiple dystrophin antibodies.*
UCLi007 HIDEMs were differentiated for 8 days. Hoechst in blue, MYH1E (MF20 antibody) in purple, Dystrophin (DYS-1, DYS-2, DYS-3 or RB-DYS) in red. Scale bar = 300 µm.

Overall, the results show very limited dystrophin immunoreactivity, hardly discernible from negative control samples. While this is mainly an issue of optimisation and validation of protocols for different derivation methods of myotubes from iPS cells, as other groups show results, with a range of efficiency, using these antibodies (H. L. Li *et al.*, 2015; Chal *et al.*, 2016; Rao *et al.*, 2018; Al Tanoury *et al.*, 2021), the process is laborious and is a barrier to new researchers and methods. This is further complicated by the fact that the field is moving towards multi-lineage DMD models in which optimisation would also need to account for antibody affinity to the various cell types and their respective *dystrophin* isoforms..

In efforts to confirm detection of dystrophin using western blotting in iPS cell-derived myoblasts, immortalised myoblasts (commonly known to express dystrophin and other skeletal muscle proteins at higher levels than human iPS cell-derived myotubes (Jalal, Dastidar and Tedesco, 2021)) differentiated in a 3D environment, using skeletal muscle constructs, were utilised. The three most commonly used *dystrophin* antibodies were utilised for this experiments, DYS-1, DYS-2, and DYS-3. The results in **Figure 4-2** show that DYS-1 is detecting multiple bands, these include the expected bands corresponding to Dp260 and Dp427. DYS-2 detects a very faint band that corresponds to Dp427m. No bands were detected using DYS-3 antibody. However, the rest are unspecific bands and degradation bands, which has been a long standing issue in the field (Hoffman, Kunkel and Brown, 1988).

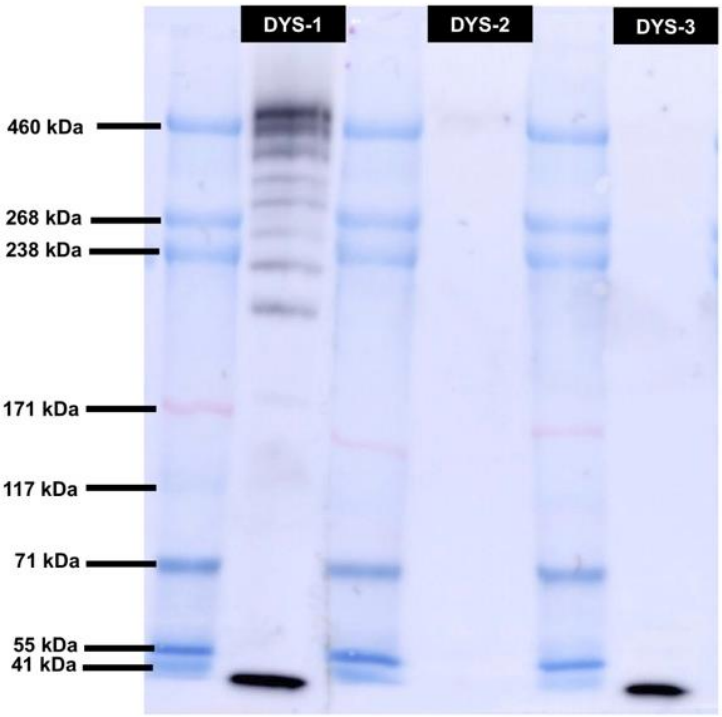

Figure 4-2. *Western blot using differentiated primary myoblast 3D muscle construct (AB1190 cell line).*
3D muscle constructs terminated on day 15 of differentiation. Antibodies used are listed at the top. Bottom band is GAPDH. Western blot performed with Dr. Valentina Lionello.

The results further confirm the difficulty of detecting dystrophin protein *in vitro* and highlight the differential epitope selectivity between different antibodies. Dystrophin antibodies have different epitopes, and unequal affinities to different isoforms, as shown by the original publications of Glen Morris (the developer of the majority of monoclonal dystrophin antibodies) and the common use of an array of dystrophin antibodies in most publications for more reliable detection of the protein (Nguyen *et al.*, 1991, 1992; Man and Morris, 1993; Lam *et al.*, 2010; Kodippili *et al.*, 2014).

This issue is further highlighted in iPS cell-derived cultures in which differentiation is asynchronous, especially in multi-lineage systems, and various isoforms are differentially expressed over time.

Together these results demonstrate the need for a dystrophin reporter capable of capturing most dystrophin isoforms without bias and providing the facility of being able to detect them the same time, to better study DMD disease mechanisms and screen therapies.

In the sections below I report the derivation of cell lines with a versatile multiple reporter for dystrophin, including luminescence, fluorescence and an interchangeable labelling tag. The reporter cassette map is in **Figure 3-4**

4.1.1 Initial Trials Using Triple-plasmid Electroporation (Old Method)

Multiple attempts were conducted to insert the cassette into five iPS cell lines using the triple-plasmid electroporation method, detailed 3.4.2.1.1. After antibiotic selection, colonies were isolated as clones and expanded. The colonies were then screened using PCR for 5' and 3' end insertion of the cassette. NCRM1 electroporation experiment had clones with confirmed 3' and 5' PCRs, 4 in total. Results are summarised in **Table 4-1**.

Cell Line (Type)	Clones Isolated	Clones confirmed
NCRM1 (Wild type)	22	4
NCRM5 (Wild type)	26	0
SBI B3 (Wild type)	41	0
DMD 864 (dystrophic)	50	0
CENSOi007 (dystrophic)	73	0

***Table 4-1**. Summary of gene-editing trials using the original CRISPR/Cas9 based approach on 5 cell lines.*

In addition to that, insertion was assessed using southern blotting on multiple clones, after genomic digestion with NcoI, results shown in **Figure 4-3**. The first panel in the figure is for hybridisation using the 3' probe and shows that multiple clones have acquired the cassette (A7, A11, A15, A22, B3, B11, and B30; in addition to A2, B15, B16 and B30 (heterogenous). The second panel of the figure was using a probe that binds to BSD, in the middle of the reporter cassette insert, and therefore assess off-target inserts of the cassette. Four clones, A7, A11, A15, B22 and B26, were confirmed to have an insertion with no off-target inserts and no heterogeneity.

These four clones, two of which are heterogenous A2, and B30 based on the extra wild type size band in southern blotting, were scanned using PCR to assess the complete insertion of the cassette, using multiple primer pairs that give rise to overlapping amplicons, shown in **Figure 3-6** and listed in **Table 3-2**.

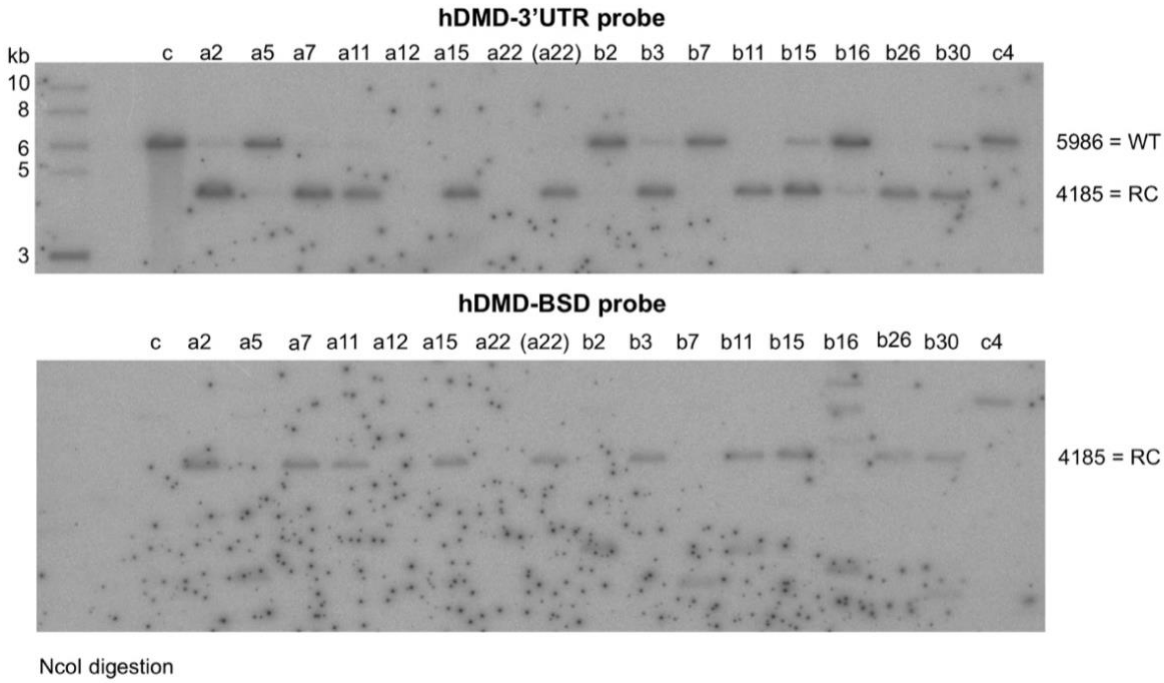

Figure 4-3. *Southern blotting for the control cell line (NCRM1) clones generated using the old method (triple-plasmid electroporation).*
Top panel is the result using a probe that binds to the 3' end of the inserted cassette. Bottom panel is the result using a probe that binds to the blasticidin resistance sequence in the middle of the inserted cassette. WT= wild type, RC= reporter cassette. Souther blot performed in collaboration with Prof Konstantinos Anastassiadis, Dresden University, Germany.

The results in **Figure 4-4** are of PCR screening for NCRM1 human iPS cell clones transfected with the reporter cassette: Clone A (B11), B (B26), C (B30) and D (A2). The results suggested that clones A (B11) and D (A2) possess the entire cassette, however, there are multiple unspecific bands amplified by PCR using L17 and R8 primers; while cells from clones B (B26) and C (B30) likely carry incomplete fragments of the cassettes; which could still be functional.

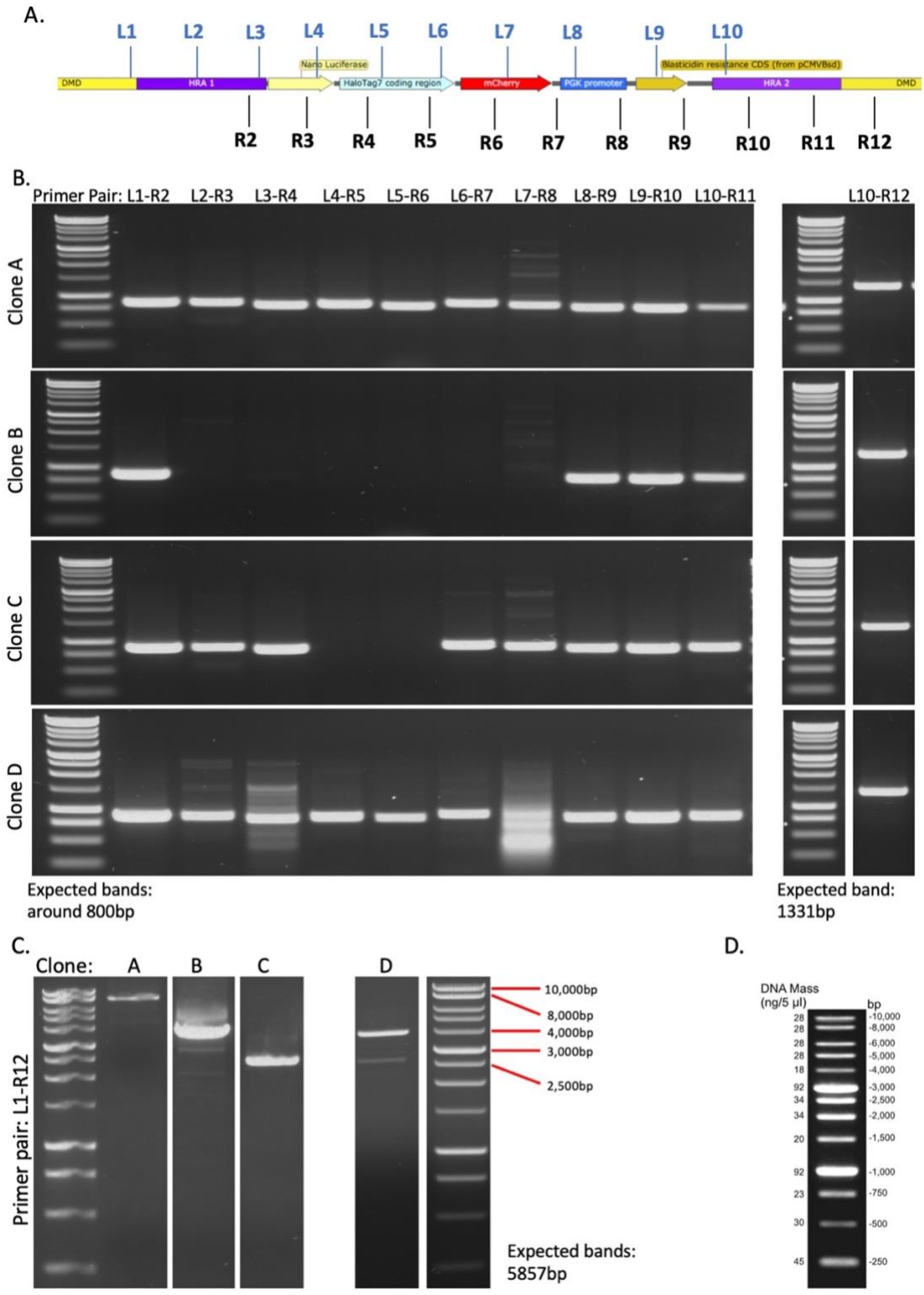

Figure 4-4. *PCR screening of clones obtained from initial CRISPR/cas9 gene editing experiments.*

A. *Map of the primers used to amplify the entire length of the inserted cassette in the obtained clones. B. Gel electrophoresis of clones' PCRs using the primer pairs listed at the top, amplifying overlapping sections of the cassette of around 800 bp. C. Gel electrophoresis for clones' PCR using primers spanning the entire length of the cassette outside the homologous recombination arms. D. 1 kb Ladder used for all gels in this figure, map from http://www.bio-helix.com/products/5. Analyses done together with M. Zhou, MRes student under my and Prof. Tedesco supervision.*

PCRs for clone A, in **Figure 4-5**, using primers encompassing the entire insert, show a predominantly bright band at around 8000 bp, another 2 very faint bands can be seen around 4000 bp, and between 5000 and 6000 bp; top panel of figure. PCR amplification for overlapping primer pairs spanning the length of the cassette have bands of the expected size. These results indicate that the cassette components have been inserted, however, there might be a duplication suggested by the large 8000 bp) band. Another possible cause for the bands detected for clone A is that the population is not clonal and are thus heterogenous.

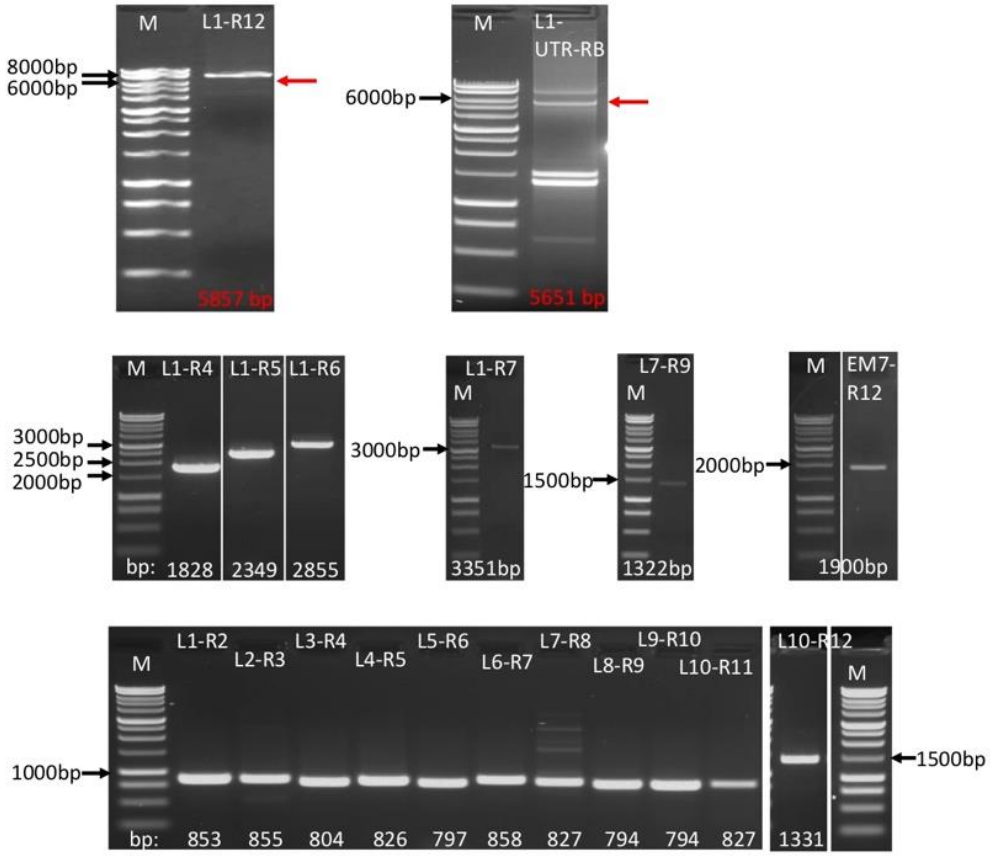

Figure 4-5. *Agarose gel electrophoresis showing PCR amplicons obtained using primer pairs targeting reporter sequence in clone A (aka. B11).*

The expected band size for each reaction was labelled at the bottom of the gel. White colour labelling indicates that the band obtined is of the expected size, while the red colour indicates it is not. M represents 1kb DNA ladder. PCR products amplified by primers listed at the top of the ladder. Two pairs of primers were used to amplify the entire reporter outside of the homology recombination arms: L1-R12 (left panel) and L1-UTR-RB (right panel). Red arrows pointed the faint bands with the correct size. Analyses done together with M. Zhou, MRes student under my and Prof. Tedesco supervision.

PCRs for clone B, in **Figure 4-6**, using primers encompassing the entire insert show a predominantly bright band at 4000 bp with fainter bands of higher and lower weights. The results of PCR amplification for overlapping primer pairs spanning the length of the cassette indicate that there is a huge omission from the cassette, that includes a significant section of the 5’ end of the cassette, including the NanoLuciferase, HaloTag and mCherry sequences, roughly around 2600 bp. The obtained result is not consistent with the result from the first panel as both PCRs show a very clear band at 4000 bp however, there are multiple bands that can be seen at lower weights (particularly, one around 2000 bp, that is 3800 bp smaller than the expected size, corresponding to the size of the missing area from L2-R8), this could be due to duplications or a result of heterogeneity of the population.

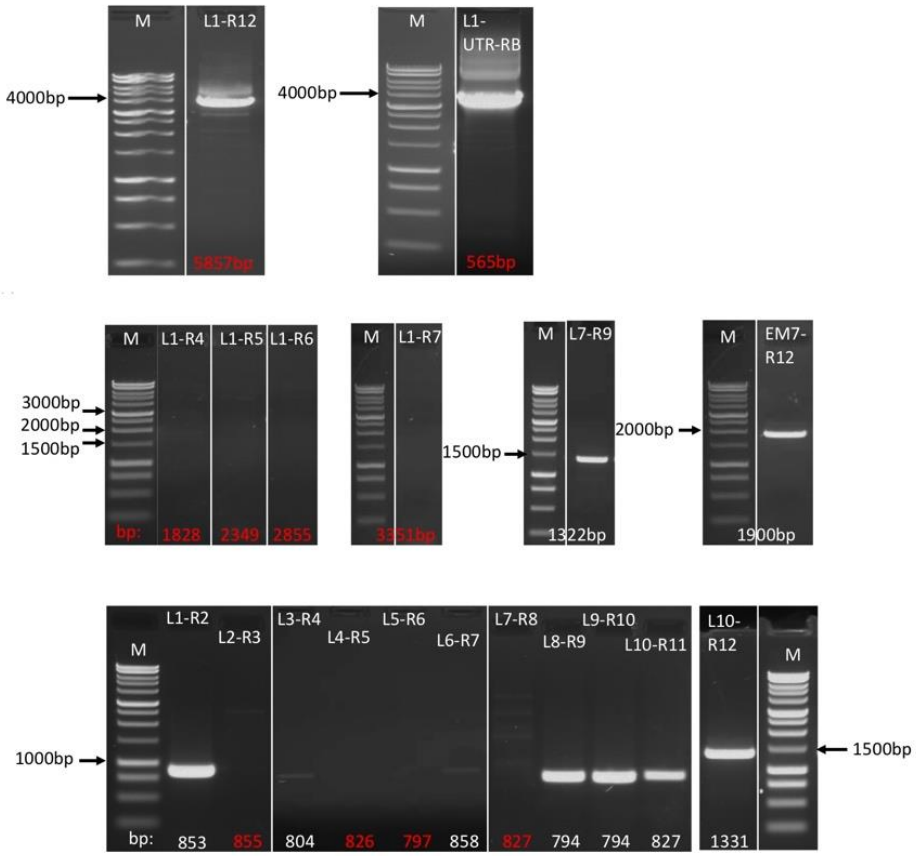

Figure 4-6. *Agarose gel electrophoresis showing PCR amplicons obtained using primer pairs targeting reporter sequence in clone B (aka. B26).*
The expected band size for each reaction is labelled at the bottom of the gel. White colour labelling indicates that the band obtained is of the expected size, while the red colour indicates it is not. M represents 1kb DNA ladder. PCR products amplified by primers listed at the top of the ladder. Two pairs of primers were used to amplify the entire reporter outside of the homology recombination arms: L1-R12 (left panel) and L1-UTR-RB (right panel). Analyses done together with M. Zhou, MRes student under my and Prof. Tedesco supervision.

PCR results for clone C, in **Figure 4-7**, using primers encompassing the entire insert show a predominantly bright band less than 4000 bp and with fainter

bands of higher and lower weights. The results of PCR amplification for overlapping primer pairs spanning the length of the cassette indicate the omission of a central section of the cassette, roughly less than 650 bp, this area falls into HaloTag sequence only. This is not consistent with the result of the PCR shown in the first panel of the figure showing a clear band at 4000 bp, however, there are multiple very faint bands, above and below the 4000 bp, the largest of which is consistent with the result in the bottom PCR panel.

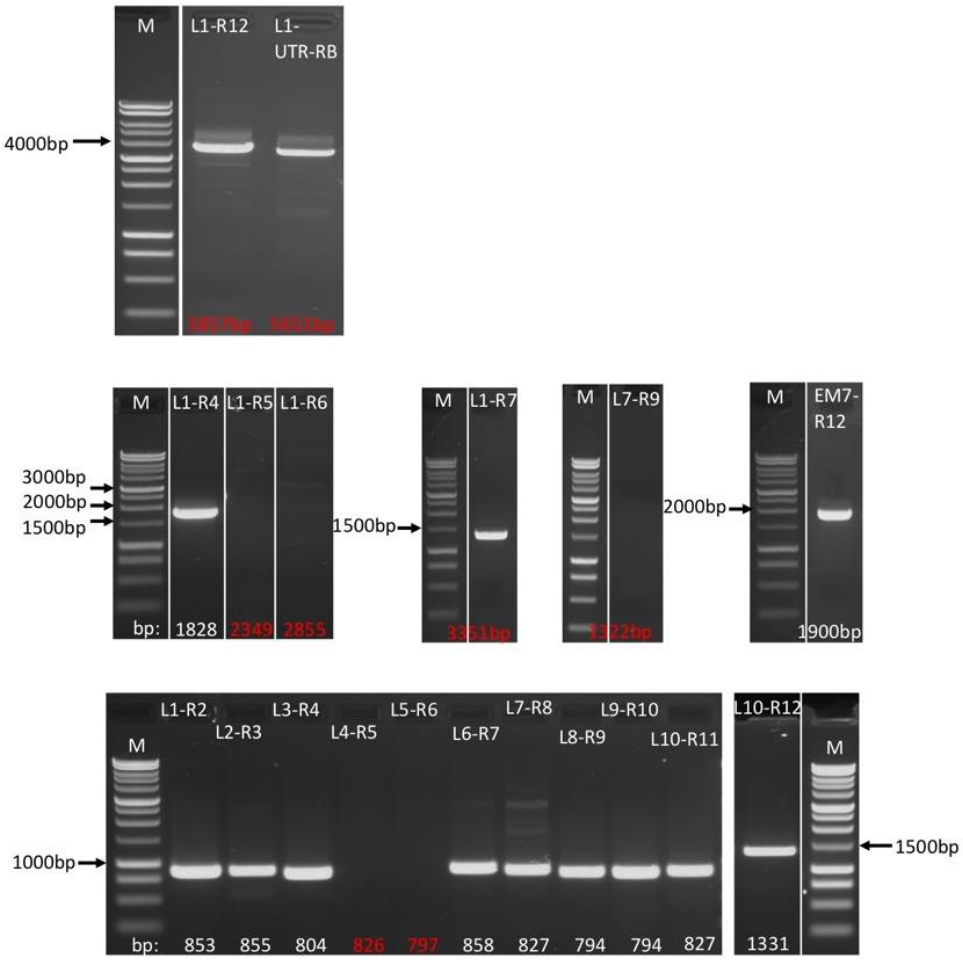

Figure 4-7. *Agarose gel electrophoresis showing PCR amplicons obtained using primer pairs targeting reporter sequence in clone C (aka. B30).*

The expected band size for each reaction was labelled at the bottom of the gel. White colour labelling indicates that the band obtained is of the expected size, while the red colour indicates it is not. M represents 1kb DNA ladder. PCR products amplified by primers listed at the top of the ladder. Two pairs of primers were used to amplify the entire reporter outside of the homology recombination arms: L1-R12 (left panel) and L1-UTR-RB (right panel). Analyses done together with M. Zhou, MRes student under my and Prof. Tedesco supervision.

PCRs for clone D, in **Figure 4-8** using primers encompassing the entire insert, show six distinct bands with sizes between 1500 bp and more than 10,000 bp, one of which is the expected band size, indicated with a red arrow. The results of PCR amplification for overlapping primer pairs spanning the length of the cassette show all the expected screening bands, however, for the amplification using the L7-R8, L2-R3 and L3-R4 primer pairs, multiple bands/nonspecific binding is observed. This corresponds to the results of the PCR sing primers encompassing the entire insert that show multiple clear bands. It is most likely that this is a heterogeneous population with different cells contributing to different bands shown on the PCR, none of which possess the intact full length of the cassette.

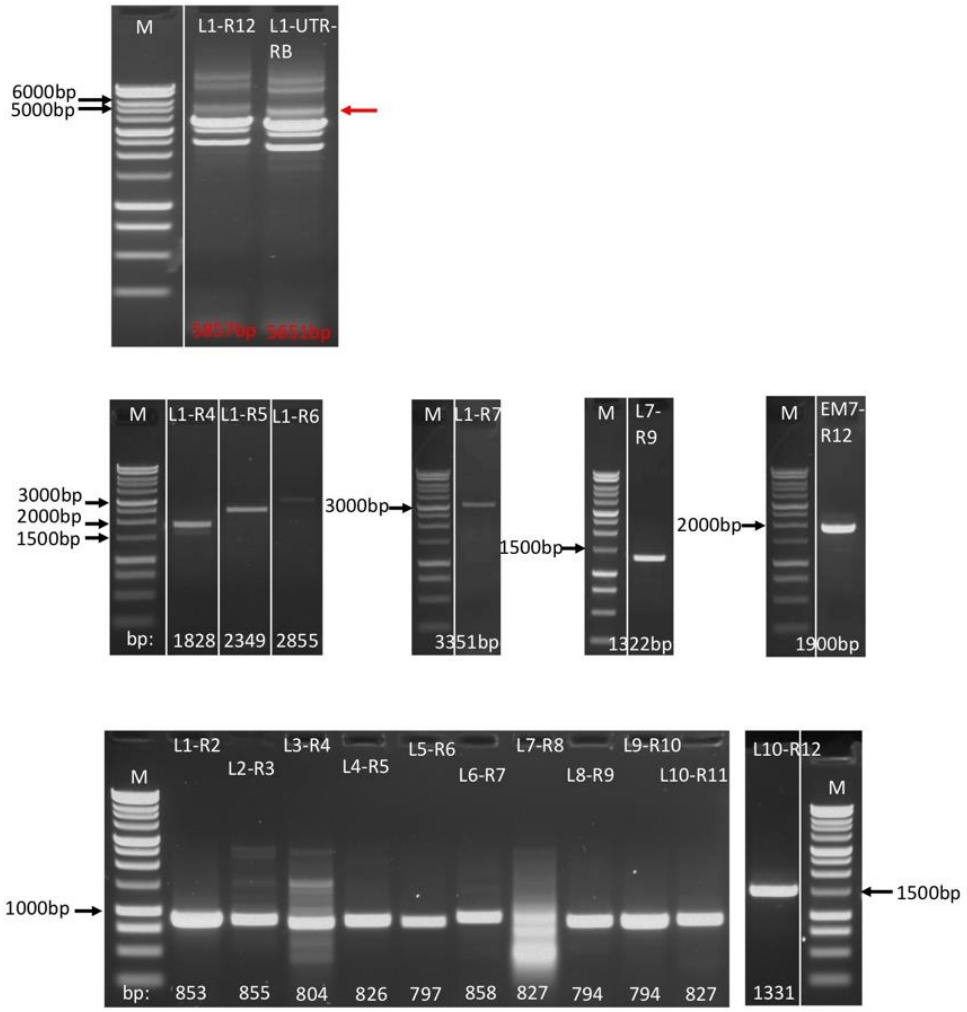

Figure 4-8. *Agarose gel electrophoresis showing PCR amplicons obtained using primer pairs targeting reporter sequence in clone D (aka. A2).*

The expected band size for each reaction was labelled at the bottom of the gel. White colour labelling indicates that the band obtained is of the expected size, while the red colour indicates it is not. M represents 1kb DNA ladder. PCR products amplified by primers listed at the top of the ladder. Two pairs of primers were used to amplify the entire reporter outside of the homology recombination arms: L1-R12 (left panel) and L1-UTR-RB (right panel). Red arrow points the faint band with the correct size. Analyses done together with M. Zhou, MRes student under my and Prof. Tedesco supervision.

The results of the PCRs indicate that the four clones are likely heterogenous, with sub populations that have the cassette or fragments of it inserted in frame. The clones were screened for the functionality of some of the reporter cassette's components, in 4.2.2.1.1 and 4.2.2.2.1.

Overall, the results suggest that disruptions occur during the gene-editing procedures. The results also confirmed that this original triple-plasmid electroporation approach, using an electroporator, was highly inefficient and further CRISPR/ Cas9 gene editing experiments were needed to derive dystrophin reporter clones with the full cassette inserted in frame.

4.1.2 CRISPR/CAS9 OPTIMISATION EXPERIMENTS

The low editing efficiency and genomic changes observed in CRISPR/Cas9 editing using the triple-plasmid electroporation original (old) method prompted the optimisation of the procedures as detailed in 3.4.2.1.2.

4.1.2.1 Guide RNA Optimisation

Using this new approach, the cutting efficiency of the gRNAs was assessed, individually and in a cocktail to determine the most effective guide/combination of guides- without the presence of the reporter cassette template. New single guide RNAs were designed using Benchling.com's CRISPR design tool based on the protospacer adjacent motif (PAM) site for *S. pyogenes* derived Cas9 (we use New England Biosciences' EnGen Cas9 with NLS (nuclear localisation signal)). The three guides within a range of 30 bp upstream and

downstream the desired insertion site, with the highest on-target cutting score (Hsu *et al.*, 2013; Doench *et al.*, 2016) and relatively low off-target effect were chosen. The selected guide binding sequences are shown in **Figure 3-3**. Three sequences were chosen to test, individually and in a cocktail/combination.

The obtained bulk population of edited cells were sequenced, and the sequencing files were uploaded to Synthego's ICE tool, which calculates the cutting efficiency of the guides in relation to thee control sequencing file (a negative control of the nucleofection experiment, same cell line without editing) using the algorithms developed by Doench and colleagues (Doench *et al.*, 2014).

x

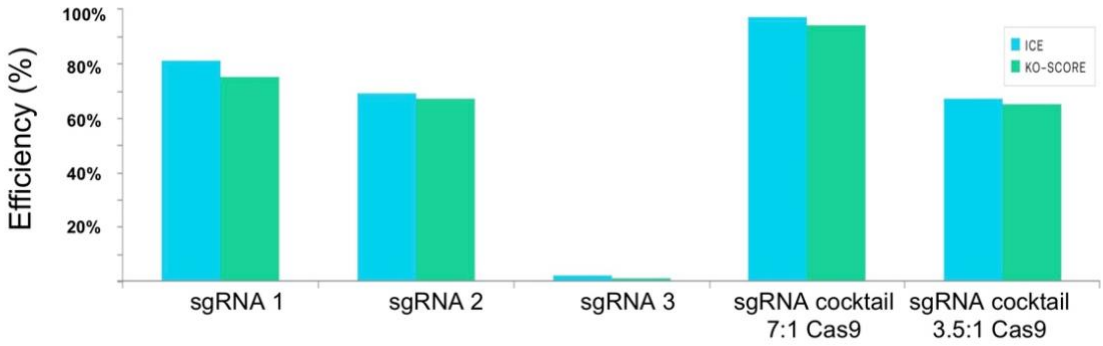

Figure 4-9*. Cutting efficiency of sgRNAs through the Inference of CRISPR Efficiency (ICE) tool by Synthego.*
The ICE score is the score of any edits in the target sequence and the knockout (KO) score represents the proportion of cells that either have a frameshift or 21 bp+ indel. The results are of three technical repeats and thus no statistical analysis has been done.

The results in **Figure 4-9** show that the cutting efficiency in the cocktail of 3 sgRNAs is around 10% higher than the highest performing sgRNA delivered individually. This is in keeping with the literature as it has been demonstrated that a cocktail of guides achieves better cutting rates than individual sgRNAs (Jang *et al.*, 2018).

The results, also show higher cutting efficiency, more than 90% using a high 7:1 ratio of sgRNA particles to the Cas9 protein to the low 3.5:1 ratio, around 65%. sgRNA3 and sgRNA1 are flanking the desired insertion site and overlap as it has been demonstrated that this conformation increases the likelihood of HDR compared to non-homologous end joining (NHEJ) (Jang *et al.*, 2018). This suggests that a cocktail of the 3 at a high concentration offers best cutting efficiency at more than 90%.Interestingly, sgRNA3 showed a very low cutting efficiency, which might be due to an error loading the sgRNA in the sample or during the nucleofection- since showed a good efficiency in the HDR experiment addressed later. Since the two other guides showed a high cutting efficiency, I decided to proceed with editing using a single guide per experiment to avoid the additive nature of potential off target effects in using multiple guides, which are listed in **Table 4-2** and shown in **Figure 3-3**.

I also performed *in silico* analysis to assess the likelihood of the selected sgRNA sequences to target other areas of the genome and lead to off-target cutting events. The results below, in **Table 4-2**, show the top 5 potential off-target sites, the score, their respective gene and genomic locus. The scores, calculated using the methods published by Hsu and colleagues, ranging from 0-100 in which

100 is higher likelihood of cutting at site, indicate that the likelihood of off-target edits is very minor (Hsu *et al.*, 2013). In addition to that, none of the off target sites fall into coding regions of the genome. In addition, any clones that are confirmed to have off-target cassette inserts using southern blotting are not used in further steps. For the purposes of this study, in which the cells were used exclusively *in vitro*, assessing off target potential effects using *in silico* tools prior to administration and using southern blotting for probes within the template cassette in obtained clones, was sufficient to ensure no major disruptions have occurred.

	Target sequence	PAM	Score	Gene	Locus
sgRNA 1 (MK)	TGTAGGAAGTCTTTTCCACA	TGG	100.0	DMD	chrX:-31140024
	GCTAGGCAGTCTTTTCCACA	CAG	1.8	N/A	chr14:-68793612
	TGGATGAAGTCTTTTCCACG	TGG	1.7	N/A	chr2:-20051965
	TGGAGGTTGTCTTTTCCACA	CAG	1.7	N/A	chr9:-27845399
	TTTATGAAGTGTTTTCCACA	GAG	1.7	N/A	chr11:+48008397
	TGGAGGACTTCTTTTCCACA	GGG	1.6	N/A	chr9:+7502366
sgRNA 2	AGATGATTTGGGCAGAGCGA	TGG	100.0	DMD	chrX:-31140000
	AGAAGATCTGGGCAGAGCGA	CAG	6.0	N/A	chr1:+26925659
	TGATCATATGGGCAGAGCGA	AGG	2.8	N/A	chr9:+70717967
	TGATCATATGGGCAGAGCGA	AGG	2.8	N/A	chr9:-44164073
	TGATCATATGGGCAGAGCGA	AGG	2.8	N/A	chr9:-43179073
	TGATCATATGGGCAGAGCGA	AGG	2.8	N/A	chr9:+42322752

sgRNA 3 (LM)	AGACTTCCTACATTGTGTCC	TGG	100.0	DMD	chrX:+31140046
	AGGCTTCACACATTGTGTCC	TGG	1.6	N/A	chr13:+43888285
	AGACATCTTACATTGTGTCA	GGG	1.6	N/A	chr9:+129056439
	ACCCTCCCTACATTGTGTCC	AGG	1.5	N/A	chr17:+48205730
	TTTCTTCCTCCATTGTGTCC	TAG	1.4	N/A	chr9:-86012383
	ATAATTGCTTCATTGTGTCC	AGG	1.0	N/A	chr15:+90582095
sgRNA 4 (KN)	TGTTTTCCAGGACACAATGT	AGG	100.0	DMD	chrX:-31140041
	TTTTTTCCAGGACACAATGA	AAG	8.6	N/A	chr6:-18375620
	TGTTTCCAAGGACACAATGT	AAG	3.3	N/A	chr1:-178031633
	TGTTTTCTTGGACACAATGT	AAG	3.2	N/A	chr1:+92612932
	TTATTTCCAAGACACAATGT	AAG	2.6	N/A	chr4:+20569678
	TGCTGTCCAGGACACAATGA	TGG	1.7	N/A	chr11:+80681390
sgRNA 5	AGACTTCCTACATTGTGTCC	TGG	100.0	DMD	chrX:+31140046
	AGGCTTCACACATTGTGTCC	TGG	1.6	N/A	chr13:+43888285
	AGACATCTTACATTGTGTCA	GGG	1.6	N/A	chr9:+129056439
	ACCCTCCCTACATTGTGTCC	AGG	1.5	N/A	chr17:+48205730
	TTTCTTCCTCCATTGTGTCC	TAG	1.4	N/A	chr9:-86012383
	ATAATTGCTTCATTGTGTCC	AGG	1.0	N/A	chr15:+90582095

Table 4-2. *In-silico analysis of off-target effects of sgRNAs used in reporter insertion CRISPR/Cas9 gene-editing experiment.*

The table lists the main target (dystrophin) and the top 5 off-target loci. For potential off target sequences, bases in black match that of the endogenous target sequence, bases in red are different from target sequence.

4.1.2.2 Comparing Electroporation and Nucleofection

To further optimise the gene editing procedure, I compared the efficiency of using electroporation (using a Bio-Rad Pulser Xcell electroporator) and using a nucleofector (Lonza Amaxa 4D nucleofector), using sgRNA and Cas9 plasmids. To gain some insight on the efficiency of delivering the sgRNA and Cas9 as plasmids compared to a pre-assembled RNP complex, I used a pre-assembled RNP complex composed of the Cas9 protein and sgRNA 2 (LM)).

The results below are the number of clones obtained after antibiotic selection, in **Table 4-3**; they show that clone outgrowth from electroporation is higher and suggest that using an RNP complex leads to higher clone outgrowth compared to using plasmids, however, despite the initial outgrowth rate, both electroporation and nucleofection using RNP complex give rise to just one successfully edited clone.

Method	Electroporation	Nucleofection
RNP complex	64 (1)	14 (1)
Plasmid template and guide	23 (0)	

Table 4-3. *Number of clones obtained from electroporation and nucleofection using sgRNA 2 (LM) after culturing for 7-14 days under antibiotic selection. Between brackets are the number of clones that showed expected bands in PCR and on southern blots, indicating they have been successfully edited.*

4.1.3 Editing Using RNP Complex (New Method)

Based on the previous results, I proceeded with a new CRISPR/Cas9 gene editing approach using an Amaxa 4D nucleofector, sgRNA and Cas9

delivered as a preassembled RNP complex, and, to avoid the additive nature of potential off-target effects, a single guide was used per experiment, as detailed in 3.4.2.1.2.

4.1.3.1 Screening for Successful Edits and Off Target Inserts

NCRM1 human iPS cells at passage 50 were recovered from cryostore and used for CRISPR/Cas9 gene editing. Six reactions were carried out, along with negative controls, using the old cassette and the new cassette (**Figure 3-4**) in combination with one of three guides. The results, summarised in **Table 4-4**, suggest that insertion of a smaller sized cassette using HDR is more efficient based on the number of clone outgrowth.

Name	sgRNA 1 (MK)	sgRNA 2 (LM)	sgRNA 4 (KN)
Original cassette	6 (0)	14 (1)	12 (0)
New cassette	20 (3)	3 (1)	43 (3)

Table 4-4. *Number of clones obtained from nucleofection after culturing for 7-14 days under blasticidin antibiotic selection.*
Between brackets are the number of clones that showed expected bands in PCR and on southern blots, indicating they have been successfully edited, regardless of off-target inserts or heterogeneity.

All clones that survived expansion in culture were used in southern blotting for off target insertions using three different probes (carried out by Professor Konstantinos Anastassiadis in TU Dresden, Germany) in **Figure 4-10**. Multiple clones are confirmed have the correct insertion of the cassette in the target locus; highlighted with a red arrow in the figure. The clones which were selected for

further expansion and the results of their southern blotting analysis are summarised in in **Table 4-5**, and they further confirm PCR results in **Table 4-4**. The clones which have the expected results for probes using southern blotting are NCRM1-L1O, NCRM1-M18N and NCRM1-L14-N, unfortunately, the latter two did not survive further expansion in culture. The NCRM1-L1O was the only clone that had a successful insertion of the long version of the reporter cassette.

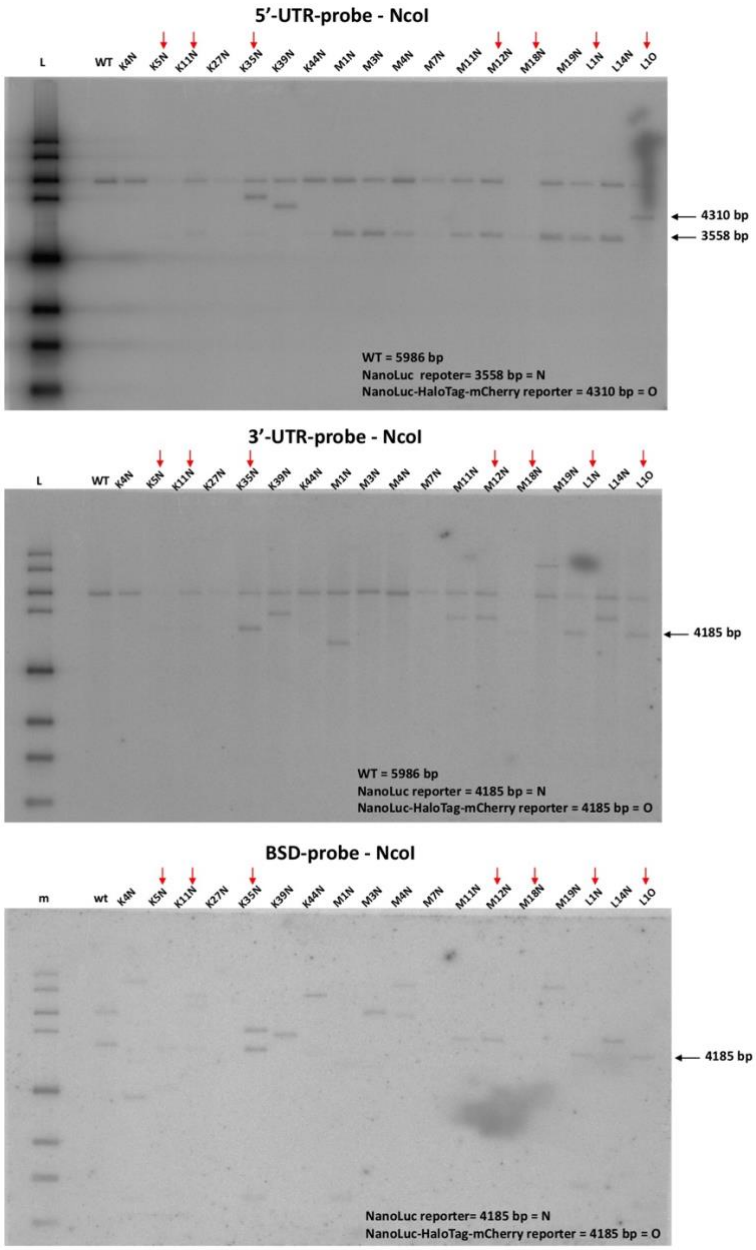

Figure 4-10. *Southern blotting for the control cell line (NCRM1) clones generated using the new method (gRNP complex).*

Top panel: Bands obtained using 5' UTR probe. Middle panel: Bands obtained using 3' UTR probe. Bottom panel: Bands obtained using blasticidin resistance sequence probe. WT refers to the band expected from wild type control cells, N refers to the band expected from the new, shortened, cassette, and O refers the

band expected from the original cassette. This experiment was performed by our collaborator Professor Konstantinos Anastassiadis.

	5' probe	3' probe	BSD probe	Status
K5N	correct	correct	Two bands	Lost
K11N	correct	correct	Two bands	Frozen
K35N	correct + additional band	correct	Two bands	Frozen
M12N	correct	higher than expected	One band	Frozen
M18N	correct	correct	One band	Lost
L14N	correct	correct	One band	Lost
L1O	correct	correct	One band	Frozen

Table 4-5. *Selected clones' summary of southern blotting results and current status.*
Highlighted in blue are clones that have the expected bands for all 3 probes, without duplications of BSD probe. Note: a WT size band is observed in 5' and 3' probes for all clones except M12N.

Going forward, NCRM1-L1O clone was selected for further experiments, and given the WT band obtained in southern blotting using the 5' and 3' probes, it was essential to single cell sort the clone into subclones following the methods in 3.4.2.2. based on the assumption that the cells might be an heterogenous population of edited and non-edited cells. While the use of an antibiotic selection should increase efficiency and eliminate un-edited cells, one explanation for the heterogenous nature of the colonies is that un-edited cells cluster with and are shielded by cells that have acquired resistance. Another possibility is that the

cells might be female cells. The obtained subclones were then assessed using PCRs with primers amplifying 4 different regions of the reporter cassette, shown in **Figure 4-11**. The results are shown in **Figure 4-12** and summarised in **Table 4-6**.

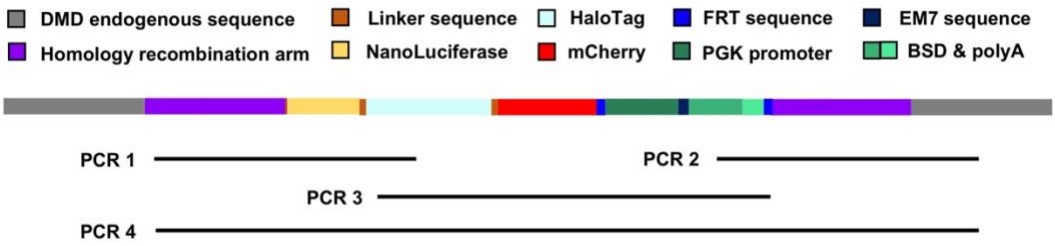

Figure 4-11. *Schematic diagram of construct and primers used for PCR screening.*

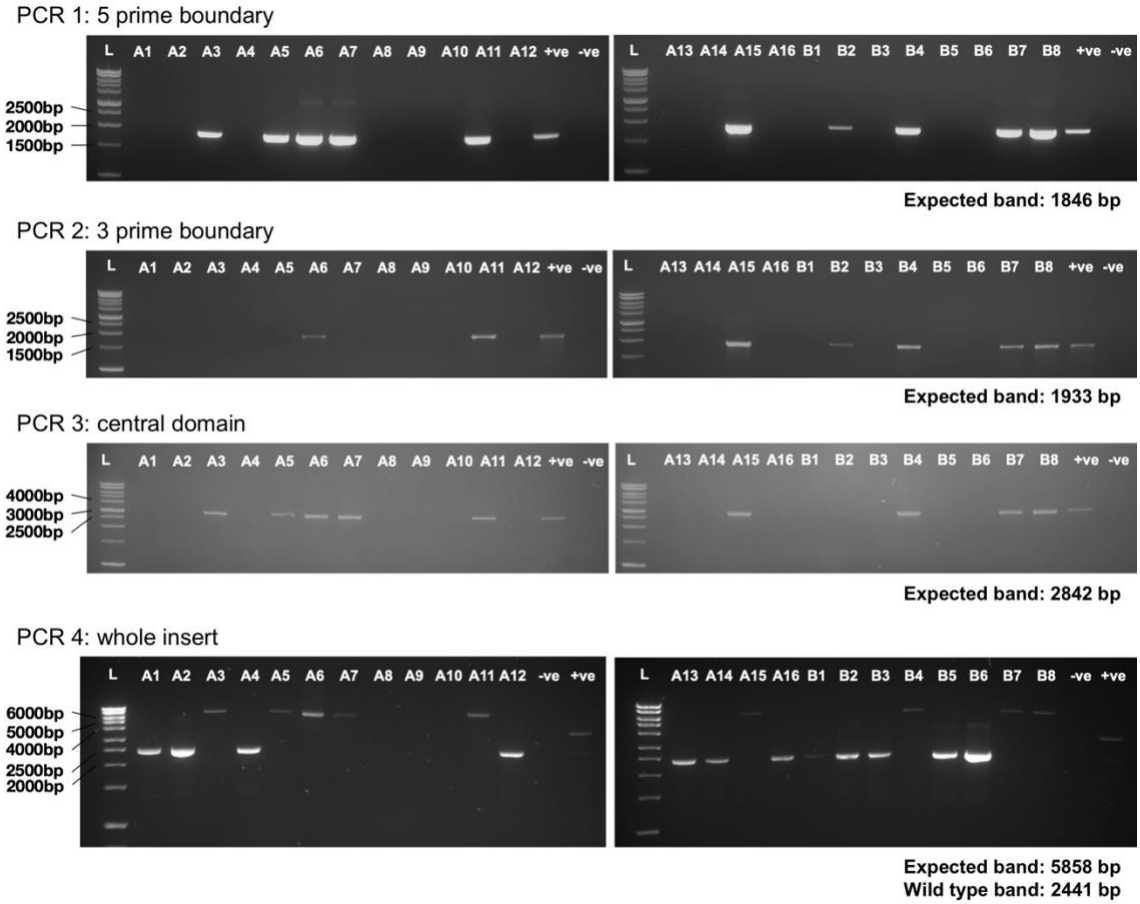

Figure 4-12. *PCR screening for reporter cassette in NCRM1 L1O sub-clones obtained through fluorescence activated single cell sorting (FACS).*
PCR 1: 5 prime boundary screening amplified using primers hDMDSeqL01 and hDMDSeqR04 and the expected band size is 1846 bp. PCR 2: 3 prime boundary screening amplified using primers EM7 and hDMDSeqR12 and the expected band size is 1933 bp. PCR 3: Central domain screening amplified using primers hDMDSeqL04 and hDMDSeqR09 and the expected band size is 2842 bp. PCR 4: Whole insert (including Homology-recombination arms) screening amplified using primers hDMDSeqL01 and hDMDSeqR12and the expected band size is 5878 bp, and 2441 bp for WT. PCR amplicons run on agarose gel (1.5%) electrophoresis and a previously generated clone is used as a positive control.

Subclone	5’ boundary	3’ boundary	Central domain	HRA encompassing

A1	N	N	N	WT
A2	N	N	N	WT
A3	Y	N	Y	Y
A4	N	N	N	WT
A5	Y	N	Y	Y
A6	Y	Y	Y	Y
A7	Y	N	Y	Y
A8	N	N	N	N
A9	N	N	N	N
A10	N	N	N	N
A11	Y	Y	Y	Y
A12	N	N	N	WT
A13	N	N	N	WT
A14	N	N	N	WT
A15	Y	Y	Y	Y
A16	N	N	N	WT
B1	N	N	N	WT
B2	Y	Y	Y	WT
B3	N	N	N	WT
B4	Y	Y	Y	Y
B5	N	N	N	WT
B6	N	N	N	WT
B7	Y	Y	Y	Y
B8	Y	Y	Y	Y

Table 4-6. *Summary of PCR screening for insertion of reporter-cassette.*

The clones that were confirmed to possess the entire sequence inserted in target locus, based on PCR analysis, are highlighted in blue in **Table 4-6**, a

total of 6. Of these, four were selected for analysis to confirm the sex of the lines, using a Taqman multiplex qPCR copy number variation assay for the SRY gene to rule out the possibility that the line is female, **Figure 4-13**. The results further confirm that the lines are male lines and that any WT bands in **Figure 4-10** southern blots were due to heterogeneity of the sample.

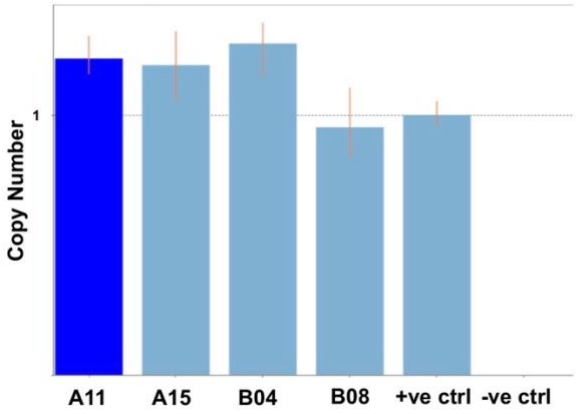

Figure 4-13. *TaqMan Multiplex qPCR copy number variation assay for SRY gene on sub clones A11, A15, B04 and B08.*
Probe used is SRY-408. TERT is used as an endogenous control. Data analysis and representation using CopyColour software. Error bars represent standard deviations for 3 technical repeats each. Analysis done with the help of Ms. Ruta Meleckyte.

Two of the sub-clones: NCRM1-L1O-A15 and NCRM1-L1O-B8 were analysed using sanger sequencing to confirm in-frame insertion of the cassette and detect any mutations that might have occurred during HDR CRISPR/Cas9 gene editing. Primers used for amplicon sequencing are listed in **Table 3-2** and the amplicon sequencing alignment is shown in **Figure 4-14**.

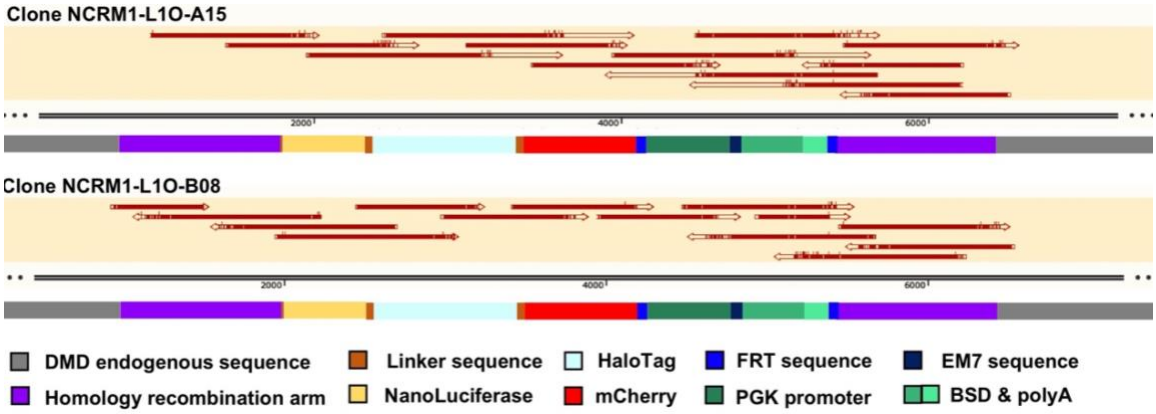

Figure 4-14. Sanger sequencing for subclones L1O-A15 and L1O-B08.
The clone L1O-A15 was sequenced using primers L01, L02, L03, L04, L05, L06, L07, L08, L10, R10, R11, UTR R (B) (12 probes). The clone L1O-B08 was sequenced using primers L01, L03, R03, R04, L05, L06, L07, L08, L09, L10, R10, R11, UTR R (B) (14 probes).

The results indicate that the clones possess some mutations that do not affect the function of the cassette as they are within the blasticidin resistance sequence, used for antibiotic selection of colonies post gene editing. The fact that mutations are identical in both clones suggest that they are likely to be sister clones. The summary of mutations confirmed by sequencing are listed below.

For both clones:

Pos 4598 G—>A (PGK promoter)

Pos 5134 T—>G/C (BSD sequence)

Pos 5173 G—>C (BSD sequence)

Pos 5381/5382. T (Linker post PGK-pA and before FRT SEQ2) insertion.

4.1.3.2 Flip Recombination to Remove Blasticidin Resistance Sequence

The NCRM1-L1O-A15 clone was used in excision of the blasticidin resistance cassette along with its promoter and polyA tail, using flip recombinase, following the steps in 3.4.2.1.3. after initial experiments of myogenic differentiation to assess reporter functions, mainly live mCherry fluorescence as detailed in the next section. However, difficulty in detecting the signal prompted us to proceed with flip recombination to remove the blasticidin resistance sequence as it was a possible cause for the dysfunction of mCherry

An initial confirmation that the human iPS cells have been successfully nucleofected with the flippase plasmid was to observe live expression of GFP, three days post nucleofection. A high percentage of the clones were GFP+ as seen in **Figure 4-15**. The human iPS cells were then single cell sorted to obtain subclones which were subsequently screened for excision of blasticidin resistance sequence using PCRs. Primers used in amplifying regions flanking the excision locus are shown in **Figure 4-16** and summarised in **Table 4-7**.

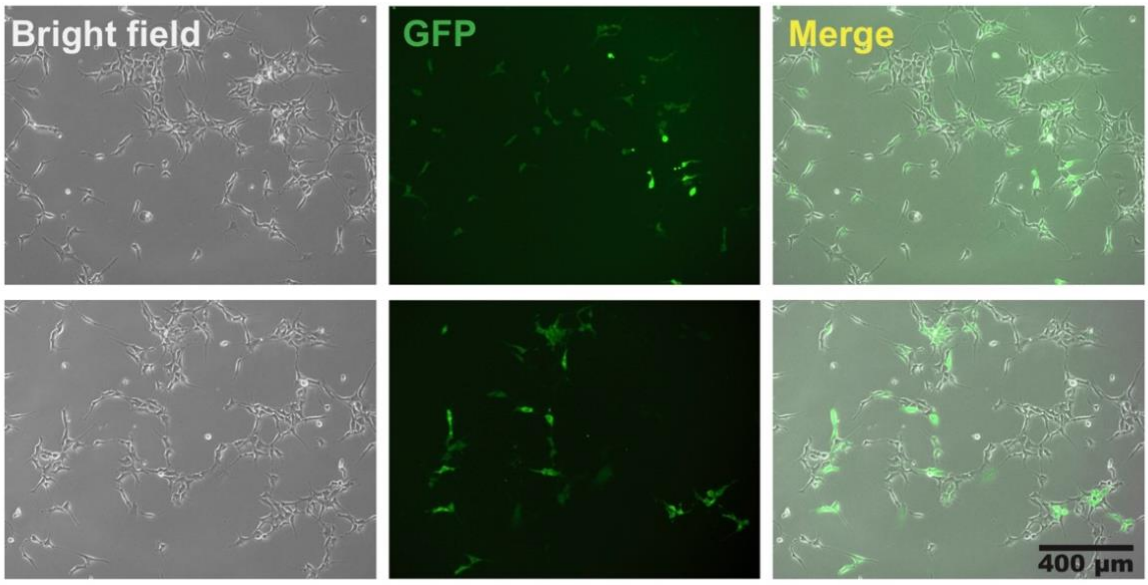

Figure 4-15. *Live imaging of NCRM1-L1O-A15 iPS cells nucleofected with pMax GFP Flpo plasmid.*

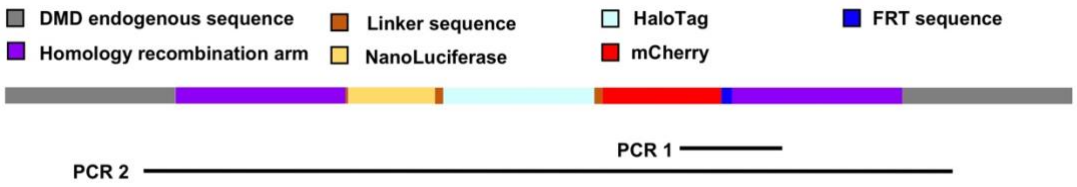

Figure 4-16. *Schematic diagram of construct and primers used for PCR screening for confirmation of removal of blasticidin resistance sequence by FLP recombinase.*

PCR 1 results are in ***Figure 4-17*** *and PCR 2 results are in* ***Figure 4-18****.*

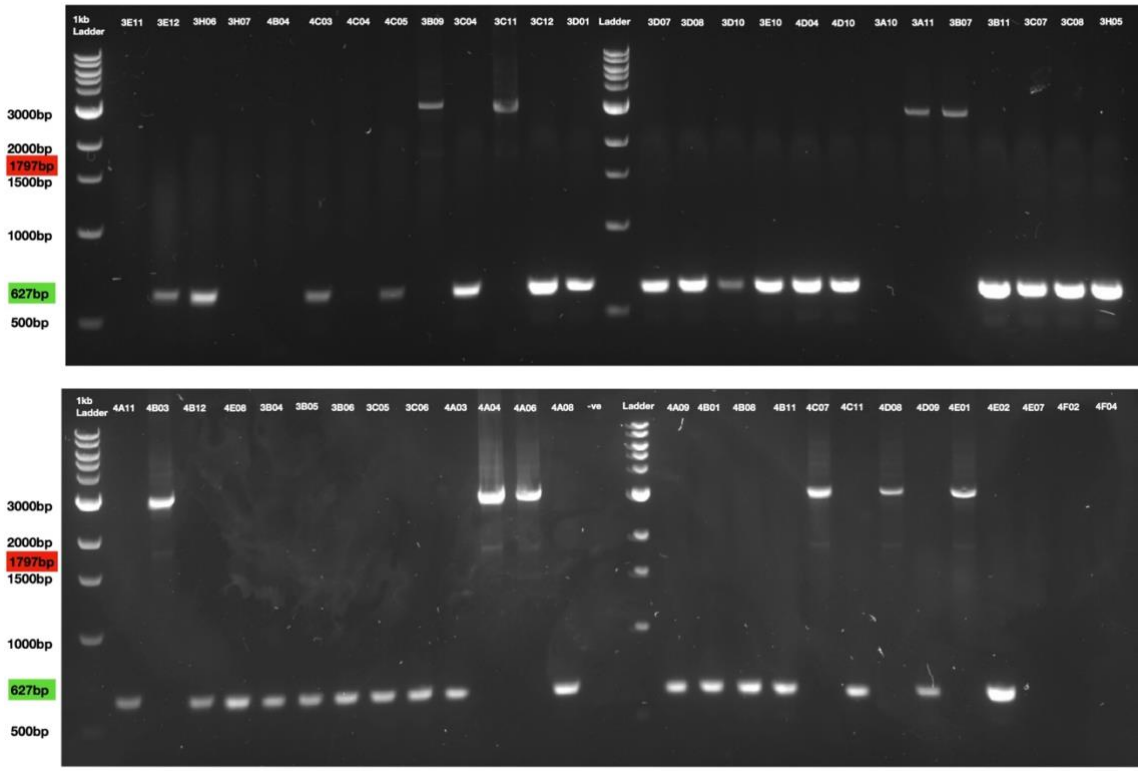

Figure 4-17. *PCR screening for L1O-A15 sub-clones for flip recombination obtained through fluorescence-activated single cell sorting (FACS).*

Using agarose gel (1.5%) electrophoresis of PCR products amplified using primers hDMDSeqL07 and hDMDSeqR10. Expected band size is 627 bp (for correctly flipped and 1797 bp for clones with unsuccessful flipping.

Subclone	Expected band?	Subclone	Expected band?
3E11	N/A- no DNA	4A11	Y
3E12	Y	4B03	N (2.8kb & 1.8kb)
3H06	Y	4B12	Y
3H07	N/A- no DNA	4E08	Y
4B04	N/A- no DNA	3B04	Y
4C03	Y	3B05	Y
4C04	N/A- no DNA	3B06	Y
4C05	Y	3C05	Y

3B09	N (2.8kb & 1.8kb)	3C06	Y
3C04	Y	4A03	Y
3C11	N (2.8kb & 1.8kb)	4A04	N (2.8kb & 1.8kb)
3C12	Y	4A06	N (2.8kb & 1.8kb)
3D01	Y	4A08	Y
3D07	Y	4A09	Y
3D08	Y	4B01	Y
3D10	Y	4B08	Y
3E10	Y	4B11	Y
4D04	Y	4C07	N (2.8kb & 1.8kb)
4D10	Y	4C11	Y
3A10	N/A- repeat	4D08	N (2.8kb & 1.8kb)
3A11	N (2.8kb)	4D09	Y
3B07	N (2.8kb)	4E01	N (2.8kb & 1.8kb)
3B11	Y	4E02	Y
3C07	Y	4E07	N/A- repeat
3C08	Y	4F02	N/A- repeat
3H05	Y	4F04	N/A- repeat

Table 4-7. *Summary of PCR screening of clones after flip recombination. Results from Figure 4-17. Subclones highlighted in blue were selected for further analyses and use. Y= 34, N=10, N/A=8. Y/(Y+N) = 77.3%.*

Multiple clones underwent successful flip recombination, highlighted in green in **Table 4-7** as screened. Clones selected for sequencing analysis are highlighted in blue. Prior to proceeding, another PCR that encompasses the entire sequence from outside of the homology recombination arm was conducted, map in **Figure 4-16**. Results for the four clones are shown in **Figure 4-18**, along

with the bulk heterogenous population subjected to flip recombination obtained from sorting, which suggests that a very high percentage of them were successfully flipped given that the WT size band is very faint and barely detectable.

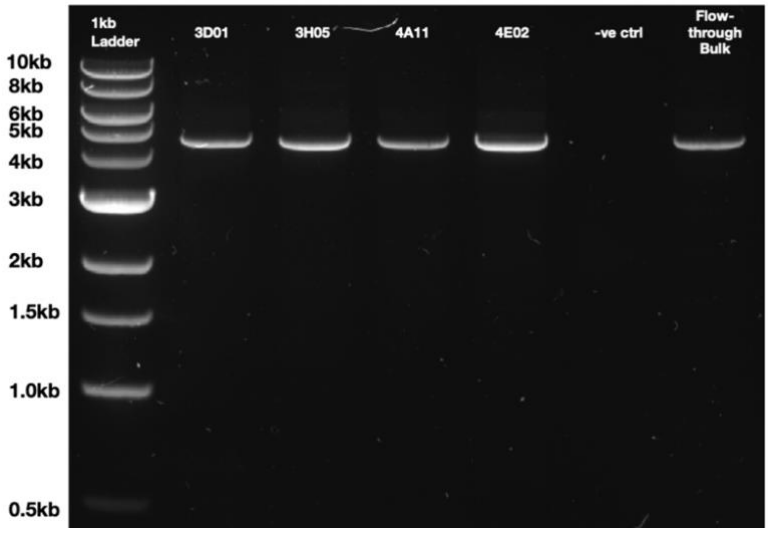

Figure 4-18. *PCR screening for four L1O-A15 sub-clones obtained through fluorescence-activated single cell sorting (FACS) to confirm full length of cassette including HRAs without FLP'd region, using agarose gel (1.0%) electrophoresis of PCR products amplified using primers hDMDSeqL01 and hDMDSeqR12. Expected band size is 4708bp.*

The NCRM1 L1O-A15-4E02 clone was lost during expansion. The PCR fragments amplified in **Figure 4-18** were then used in sanger sequencing using primers shown in **Figure 4-19**. Multiple amplicon alignment for the three clones are shown in **Figure 4-20**, confirming successful flipping and no other edits.

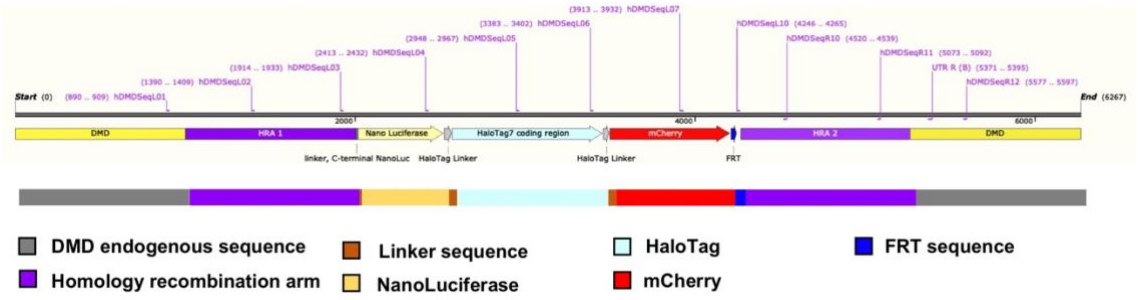

Figure 4-19. *Schematic diagram of construct after FLP recombination and primers designed for sequencing target loci.*

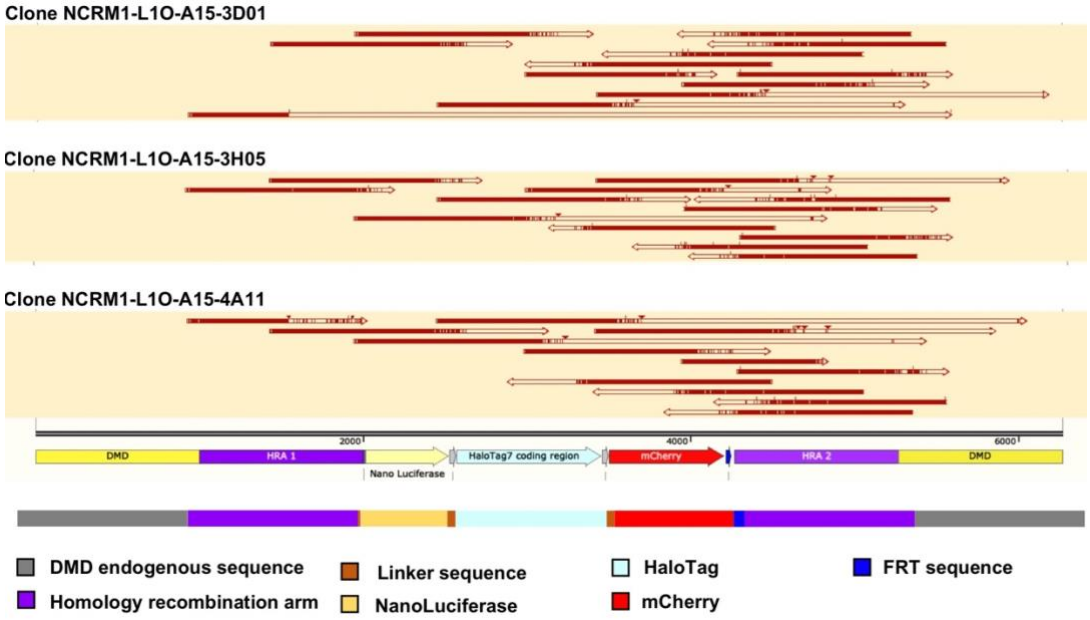

Figure 4-20. *Multiple amplicon alignment of L1O-A15 subclones 3D01, 3H05 and 4A11 after FLP recombination to expected sequence.*
Sanger sequencing using primers L01, L02, L03, L04, L05, L06, L07, L10, R10, R11, UTR R (B), R12 (12 probes) for sub-clone A. L1O-A15-3D01, B. L1O-A15-3H05 and L1O-A15-4A11.

4.1.3.3 Generation of Exon 52 Deleted Isogenic Reporter Clones

The three clones obtained from the previous step (NCRM1-L1O-A15-3D01, NCRM1-L1O-A15-3H05, NCRM1-L1O-A15-4A11) were then used in generating isogenic reporter clones with dystrophin exon 52 deletion Following the procedure in 3.4.2.1.4, using gRNAs in **Table 3-1**, and visualised in **Figure 3-5** showing intronic target sites upstream and downstream of exon 52. A cocktail of three guides, two targeting downstream of the exon and one targeting upstream of the exon were used, based on validated sequences published in the literature (Al Tanoury *et al.*, 2021; Soblechero-Martín *et al.*, 2021) were used.

In silico analysis to assess the likelihood of the sgRNA sequences to target other areas of the genome and lead to off-target cutting events was carried out. The results in **Table 4-8** show the top 5 potential off-target sites, the score, their respective gene and genomic locus. The scores, calculated using the methods published by Hsu and colleagues, indicate that the likelihood of off-target edits is very minor, (Hsu *et al.*, 2013). In addition to that, none of off target sites fall into coding regions of the genome.

	Target sequence	PAM	Score	Gene	Locus
sgRNAEx52Upstream1	GCTGAAGAACCCTGATACTA	AGG	100.0	DMD	chrX:-31747889
	GCTGGAGAACCCTGATACTG	TGG	5.7	N/A	chr1:+25457023
	CCTGAAGAAACCTGATGCTA	GGG	1.4	N/A	chr12:-6701429
	GGTCAACAAACCTGATACTA	AAG	1.0	N/A	chr2:-19956013
	GCAGCAGCTCCCTGATACTA	TAG	0.9	N/A	chr19:-56367209

	GGTGCAGCACCCTGATACTG	AAG	0.9	N/A	chr11:+113448770
sgRNAEx52Downstream1	GACCAACAGCCAAGGATATG	AGG	100.0	DMD	chrX:+31747401
	GACTATCAGACAAGGATATG	AAG	1.5	N/A	chr7:+110137577
	GACCGACTGCCAAGGATAGG	GGG	1.2	N/A	chr1:-3589169
	GTAAAAGAGCCAAGGATATG	AGG	0.9	N/A	chr10:+54507795
	TACCATCAGCCAAGGAGATG	AAG	0.9	N/A	chr9:+1184601
	GATCGACATACAAGGATATG	TAG	0.8	N/A	chr8:+8765163
sgRNAEx52Downstream2	GTTGCAAAGCATGCATTGAT	GGG	100.0	DMD	chrX:-31747505
	GTTCCAATGCATGCATTGAT	CAG	6.0	N/A	chr6:-166648998
	TTTACAAAGAATGCATTGAT	CAG	2.7	N/A	chr8:-50976184
	GATACAAAGCATGCATTGAA	AAG	1.9	N/A	chr6:-79893093
	TTTGGAATGTATGCATTGAT	CAG	1.5	N/A	chr10:+108581822
	GTTTAAAGGAATGCATTGAT	AAG	1.4	N/A	chr5:-164176223

Table 4-8. *In-silico analysis of off-target effects of sgRNAs used in dystrophin exon 52 deletion CRISPR/Cas9 gene-editing experiment.*
The table lists the main target (dystrophin) and the top 5 off-target loci. For potential off target sequences, bases in black match that of the endogenous target sequence, bases in red are different from target sequence.

After editing, single cell sorting for the cells obtained was carried out according to the methods in 3.4.2.2. PCRs using primers upstream and downstream of the cutting sites of the gRNAs, as shown in **Figure 4-21** were used to screen for successful editing. The PCR results in **Figure 4-22** show a

high percentage of edited clones with various band sizes. A band at 436 bp indicates a successful excision of the region from sgRNAEx52Upstream1 to sgRNAEx52Downstream2, a band at 532 bp indicates a successful excision of the region between sgRNAEx52Upstream1and sgRNAEx52Downstream1.

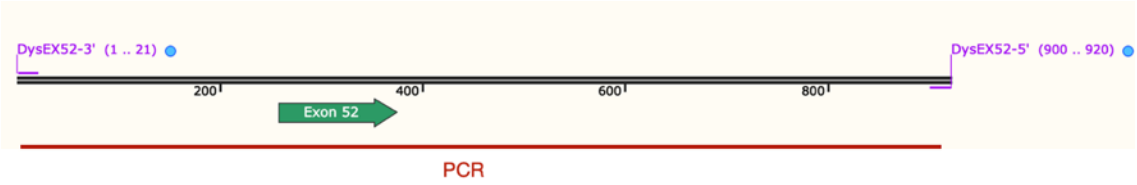

Figure 4-21. *Schematic diagram of endogenous sequence of Exon 52 with flanking regions and primers used for PCR screening for confirmation of exon deletion.*

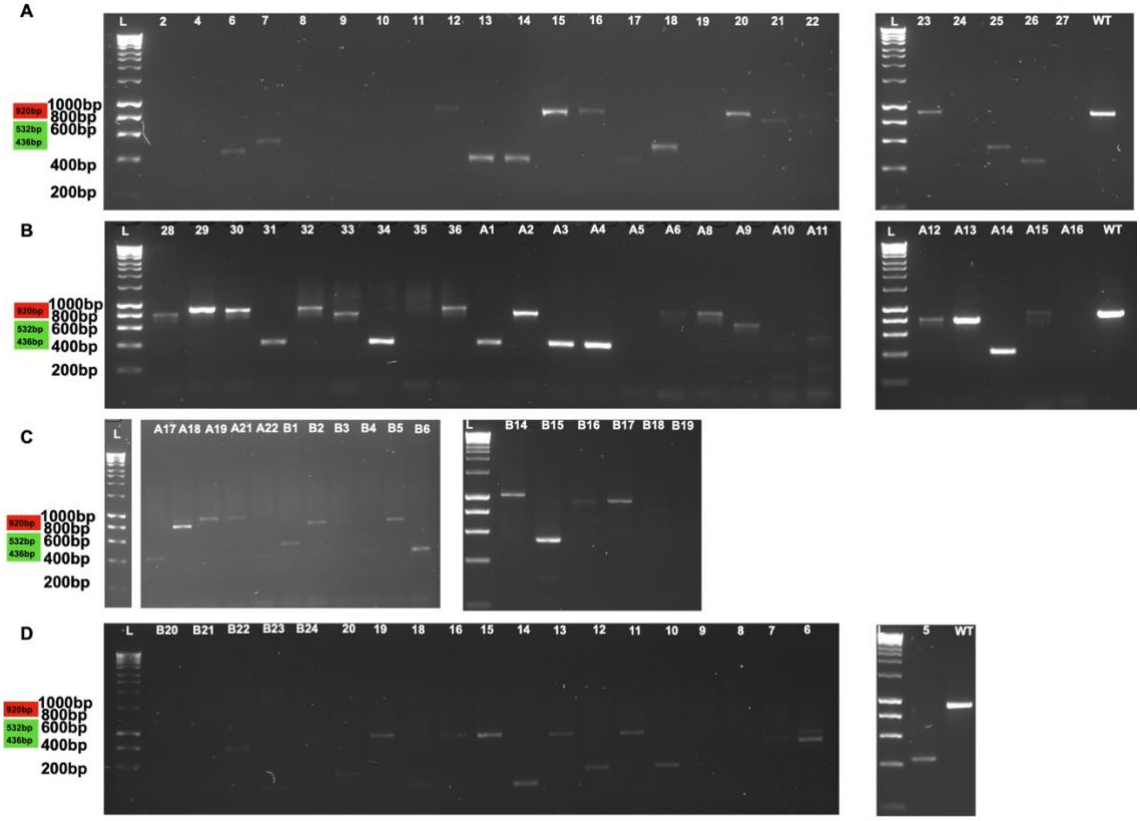

Figure 4-22. *PCR screening for CRISPR-Cas9 mediated Exon 52 Deletion for L1O-A15-(3D01, 3H05 and 4A11) sub-clones obtained through fluorescence-activated single cell sorting (FACS).*
PCR products were amplified using primers in ***Figure 4-21.*** *And resolved using agarose gel (2%) electrophoresis* ***A.*** *L1O-A15-3D01 subclones.* ***B.*** *L1O-A15-3D01 subclones (28-36) and L1O-A15-3H05 subclones (A1-A16).* ***C.*** *L1O-A15-3H05 subclones.* ***D.*** *L1O-A15-3H05 subclones (B20-B24) and L1O-A15-4A11 subclones (20-5). Expected band size for successfully excised exon 52 is 436 bp or 532 bp, WT=920 bp.*

All clones with the expected band sizes and some of the clones with unexpected bands were used in sanger sequencing to better understand the nature of the cuts. **Table 4-9** shows a summary of the PCRs and sequencing results. PCR results match that of the sequencing, however, many of the edits have extra indels around the cutting sites, which does not affect the overall outcomes of the editing as the exon has been excised. Highlighted in green are satisfactory results, clones highlighted in yellow have a successful excision of exon 52 and the ones highlighted in blue are the ones expanded and used in subsequent experiments; namely:

NCRM1-L1O-A15-3D01-17, NCRM1-L1O-A15-3D01-31, NCRM1-L1O-A15-3H05-A4, NCRM1-L1O-A15-3H05-B1, NCRM1-L1O-A15-4A11-05.

Subclone	Expected band?	Sequencing
3D01-2	No Band	
3D01-4	Very faint, between 600-	
3D01-6	Yes, closer to 400	Del. US1-DS1 (Ins.: +18 bp US1 & +1 bp DS1- 7 bp from cut site. Del.: -1 bp

3D01-7	Yes, closer to 500	Del. US1-DS2 384 bp (del.: -10 bp
3D01-8	NO, WT size	
3D01-9	NO. Slightly less than WT	
3D01-10	No Band	
3D01-11	No Band	
3D01-12	NO. WT size	
3D01-13	Yes, closer to 400	Del. US1-DS1 505 bp (del: -1 bp US1 & -14 bp DS1)
3D01-14	Yes, closer to 400. Extra band.	Del. US1-DS1 497 bp (del: -8 bp US1 & -1 bp DS1)
3D01-15	NO. WT size	
3D01-16	NO. WT size	
3D01-17	Yes, closer to 400	Del. US1-DS1 488 bp (5'seq bad)
3D01-18	Yes, closer to 500	Del. US1-DS2 376 bp. Flip DS2-DS1.
3D01-19	No Band	
3D01-20	NO. WT size	
3D01-21	NO. Slightly less than WT	
3D01-22	NO. Slightly less than WT	
3D01-23	NO. WT size	
3D01-24	No Band	
3D01-25	Yes, closer to 500	Del. US1-DS2 385 bp (Ins.: +10 US1)
3D01-26	Yes, closer to 400	Del. US1-DS1 487 bp (Ins.: +13 US1)
3D01-27	No Band	
3D01-28	NO. Less than WT	NO
3D01-29	NO. WT size	NO
3D01-30	NO. Slightly less than WT	NO
3D01-31	YES	Del. US1-DS1 488 bp.

3D01-32	NO. Slightly less than WT	Ins.: +1 bp US1 (A) and del. –9 bp at DS1
3D01-33	NO. Slightly less than WT	Del. DS1-DS2. Del.: -10 bp at US1, -7 DS1 and ins.: 1 bp DS2 (G).
3D01-34	YES	Del.: US1-DS1 492bp (del.: -1 bp DS1, -3 bp US1 and ins: +1bp US1)
3D01-35	No Band	NO
3D01-36	NO. Slightly less than WT	NO
3H05-A1	YES	Del. US1-DS1 488 bp (5'seq bad)
3H05-A2	NO. Slightly less than WT	Del. DS1-DS2 (Ins.: +33 bp and del.: -
3H05-A3	YES	Del. US1-DS1 490BP (Ins.: +3 bp and del.: -2 bp US1)
3H05-A4	YES	Del. US1-DS1 488 bp
3H05-A5	No band	NO
3H05-A6	Fragmented band	NO
3H05-A8	Fragmented band	NO
3H05-A9	Significantly less than WT	Del. US1-DS1. Ins.: + 286 bp US1, +1 bp DS1 (T)
3H05-A10	Very faint bands.	NO
3H05-A11	Many bands around	NO
3H05-A12	NO. Less than WT.	Del. DS2-DS1. Del.: - 6 bp US1
3H05-A13	NO. Less than WT.	Many indels, no KO.
3H05-A14	YES	Del. US1-DS1 488 bp.
3H05-A15	NO. WT, with shadow	NO
3H05-A16	No band	NO
3H05-A17	Yes, closer to 400	Del US1-DS1 494BP (del+4 US1 &del+2 DS1) 5'seq bad

3H05-A18	No, double band, slightly less and bit less than WT	
3H05-A19	NO WT size	
3H05-A21	NO WT size	
3H05-A22	NO, 3 bands (400-600 bp)	
3H05-B1	Yes, closer to 500	Del. US1-DS2 384 bp.
3H05-B2	NO. Slightly less than WT	
3H05-B3	NO. Slightly less than WT	
3H05-B4	Yes, closer to 400	Del. US1-DS1 493 bp (Ins.: +12 US1. Del.: -3 bp US1, -1 bp DS1)
3H05-B5	NO Slightly less than WT	
3H05-B6	Yes, closer to 400	Del. US1-DS1 488 bp (Del.: -1 DS1)
3H05-B7	Inconclusive PCR result	
3H05-B8		
3H05-B9		
3H05-B10		
3H05-B11		
3H05-B12		
3H05-B13		
3H05-B14	NO, slightly larger than WT	
3H05-B15	Yes, 500 bp and 300 bp	
3H05-B16	NO. WT size	
3H05-B17	NO. WT size	
3H05-B18	NO, slightly less than WT	
3H05-B19	No Band	
3H05-B20	No Band	

3H05-B21	No Band	
3H05-B22	NO. bit less than WT	
3H05-B23	Yes, closer to 400	Del. US1-DS1 488 bp.
3H05-B24	No Band	
4A11-20	Yes, closer to 500	Del. US1-DS2 384 bp (Del.: -5 bp
4A11-19	NO. WT size	
4A11-18	Yes, closer to 400	Del. US1-DS1 492 bp (Ins.: +4 bp
4A11-16	NO. WT size	
4A11-15	NO. WT size	
4A11-14	Yes, closer to 400	Del. US1-DS1 501 bp (Ins.: +12 bp US1)
4A11-13	NO. WT with extra band	
4A11-12	Yes, closer to 500	Del. US1-DS2 384 bp (Del.: +5 bp DS1)
4A11-11	NO. WT size	
4A11-10	Yes, closer to 500	Del. US1-DS2 383 bp (Ins.: +2 bp
4A11-9	No Band	
4A11-8	NO. WT size	
4A11-7	NO. Slightly less than WT	
4A11-6	NO. WT & less than WT	
4A11-5	Yes, closer to 400	Del. US1-DS1 488 bp.

Table 4-9. *Summary of PCR and sequencing results for L1O-A15-(3D01, 3H05 and 4A11) subclones edited using CRISPR/Cas9 to excise Exon 52.*
In blue are clones selected for further experiments (expansion and functional analyses), green are clones with precise edits excising exon 52 but not selected for further analyses, yellow are clones that have exon 52 excised but have other modifications as well and not selected for further analyses.

The steps and procedures taken and followed to edit NCRM1 human iPS cells using the new method (sgRNA and Cas9 RNP complex) to obtain dystrophin reporting lines and their isogenic exon 52 deleted dystrophic lines are outlined in **Figure 4-23**. The figure also clarifies the nomenclature used for all the clones and subclones and are consistently used in the results and discussion sections of this thesis.

.

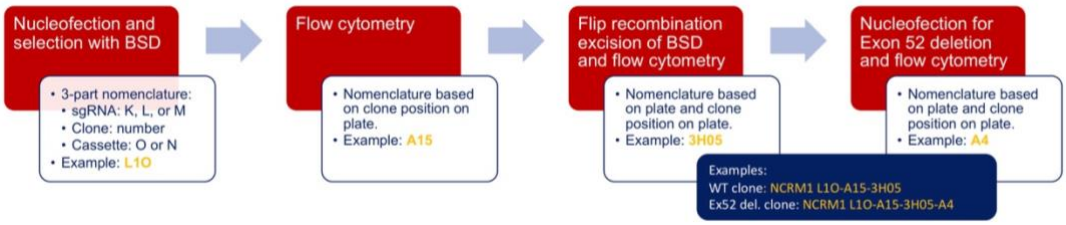

Figure 4-23. *Outline of key procedures and steps in generating NCRM1 wild type and their isogenic dystrophin exon 52 deleted human iPS cells, clarifying the nomenclature.*
sgRNAs M, L, K are sgRNA 1 (MK), sgRNA 2 (LM) and sgRNA 4 (KN), respectively. The reporter cassette O and N refer to old and new cassette, respectively. Old cassette is the one including Nanoluciferase, mCherry and HaloTag sequences while the new cassette only includes Nanoluciferase; refer to ***Figure 3-4****.*

4.2 Assessment of Dystrophin Reporter Function in Human iPS Cell-Derived Skeletal Muscle Cells

Following the results in the previous chapter, I sought out to confirm the function of the reporter sequences in the obtained human iPS cell lines. The

clones were subject to multiple quality control steps prior to use in differentiation and assessment of the functionality of the cassette. The quality control procedures' results discussed below are for the clones obtained from the second gene editing method.

4.2.1 Quality Control Steps and Assessment of Myogenic Capacity of Human iPS Cells

Prior to the use of the edited cells in further experiments, it was essential to confirm their identity and assess their pluripotency. The NCRM1-L1O clone was used for low-pass whole genome sequencing to assess genomic stability, according to the steps in 3.4.2.3.1.

The results in **Figure 4-24** show variance of reads per chromosome (orange line) from the average for all reads and confirms that chromosomes 1-22 are within expected range, confirming the lack of chromosomal aberrations. The reads for the X and Y chromosomes are much lower, half that of other chromosomes, in keeping with the signal expected from a male cell line. These results confirm overall genomic stability of the edited reporter line matching the result obtained from the low-pass whole genome sequencing experiment, figure 4.13.. Low-pass whole genome sequencing has been extensively validated in assessing copy number variations (Chau *et al.*, 2021; Mareschal *et al.*, 2021).

However, in contrast to traditional karyotyping, it is not capable of detecting large structural variants.

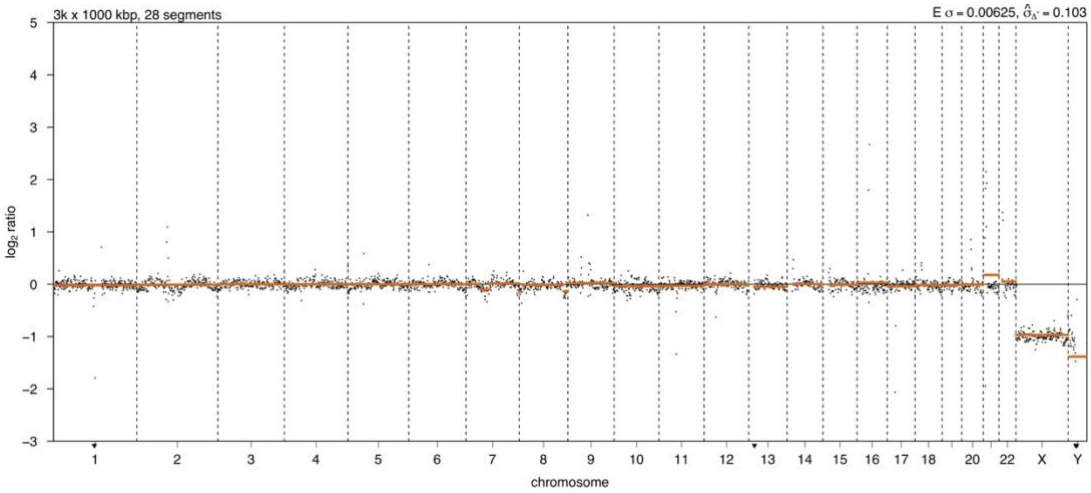

Figure 4-24. *Genomic stability analysis of NCRM1-L1O clone using low-pass whole genome sequencing.*
Sequencing amplicons (0.1x coverage) reads' log2 ratio plotted and mapped to each genomic locus, chromosomes listed on the x-axis, in comparison to the average number of reads mapping to all other comparable loci. Black dots represent individual reads and orange bars represent the consensus.

The sub-clones of NCRM1-L1O, A15 and B08, obtained through single cell sorting were assessed for their pluripotency capacity using qRT-PCR for pluripotent associated markers *OCT3/4, SOX2, KLF4* and *MYC*. The results indicate that A15 clone had expression levels similar to or higher than the L1O mother clone, **Figure 4-25**, and therefore, this clone was selected for further experiments.

Further functional tests for pluripotency include embryoid body formation and assessment of the potential of iPS cells to differentiate into the three germ lineages using qRT-PCR for these markers confirming their pluripotency (Sheridan, Surampudi and Rao, 2012). Another approach to confirm the lines' pluripotency is a bioinformatics assay in which the transcriptomic pluripotent gene expression profile of iPS cells is compared to established iPS cells, known as Pluritest (Müller *et al.*, 2011). However, this method does not provide direct information on differentiation biases or malignancy propensity ("The Stem Cell Initiative", 2018). For the purposes of the proposed applications in this thesis, qPCR analysis of pluripotency markers is sufficient.

It is good practice to periodically carry out one or more pluripotency tests on iPS cells prior to use in downstream experiments and applications especially if the aim is to obtain multiple cell lineages.

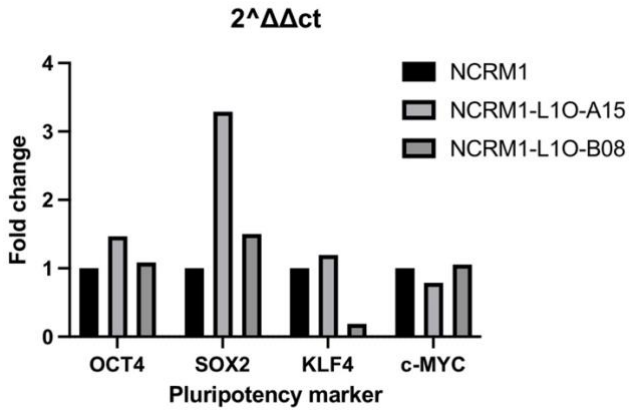

Figure 4-25. *Expression levels of pluripotency markers for NCRM1-L1O subclones using qRT-PCR.*

To confirm the myogenic capacity of the cell lines before conducting the functional analysis experiments, the cells were differentiated used the transgene free method, and the cultures were stained for key myogenic differentiation markers at two different stages. The myoblast (precursor) stage was confimed using MyoD and the terminal skeletal muscle differentiation marker Myosin Heavy Chain (*MYH*) was used at the end of the differentiation protocol. Results, in **Figure 4-26** indicate that the obtained populations at the end of the myoblast stage are indeed myogenic, expressing MyoD and that terminally differentiated myotubes at day 7 are expressing the differentiation marker myosin heavy chain, MyHC.

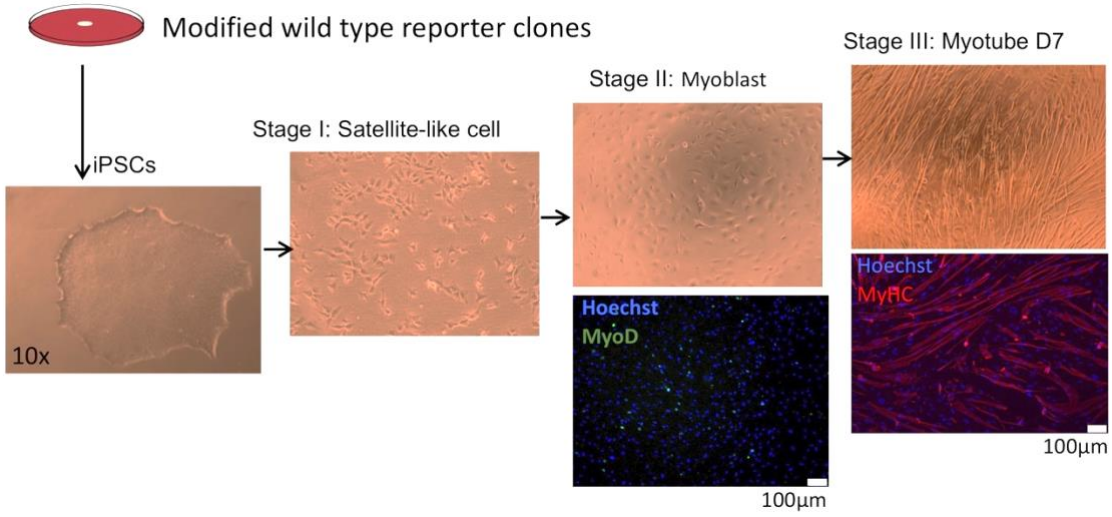

***Figure 4-26**. Derivation of skeletal muscle cells from human iPS cells using the transgene free protocol.*

Representative brightfield images of stages of differentiation based on the transgene-free method, bottom panel is immunofluorescence staining for MyoD in stage 2 cells to confirm myogenicity of the population and myosin heavy chain

MyHC for terminally differentiated myotubes (pictures are from differentiation procedure of one of the initial NCRM1 reporter clones).

4.2.2 Functional Analysis of The Different Reporter Sequences

Functional analysis of the reporter cassette sequences was assessed in two different stages: 1) the first was using the original clones obtained from triple electroporation reported in section 4.1.1 and referred to as clone A, B, C and D; 2) The second stage was using the clones obtained from the second method reported in section 4.1.3 referred to using the nomenclature in **Figure 4-23** (e.g. NCRM1 L1O-A15-4A11-05).

4.2.2.1 Nanoluciferase

To start, the expression of the NanoLuciferase luminescence was assessed as the first functional assay.

4.2.2.1.1 Triple-plasmid Electroporation CRISPR/Cas9 Clones

The clones obtained from the initial CRSPR/Cas9 experiments in 4.1.1 were then differentiated to assess the function of the cassettes' first component. NanoLuciferase expression analysis was conducted as detailed in 3.4.3.1. **Figure 4-27** shows that clones A and C to have much higher expression of levels of luminescence compared to the negative control. Clones B and D have slightly higher expression levels compared to the negative control.

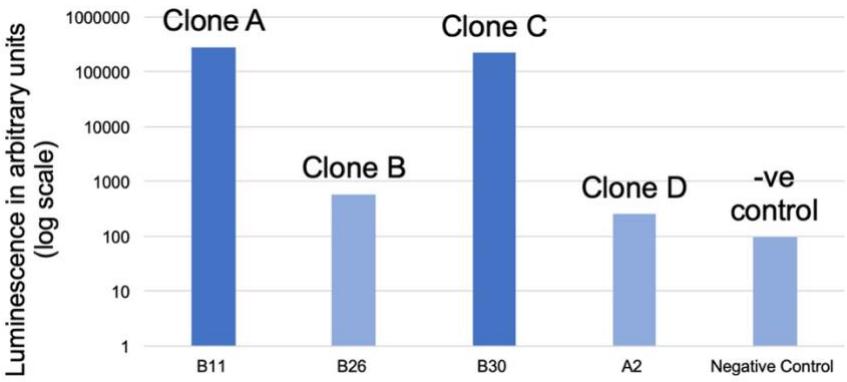

Figure 4-27. *Luminescence levels (arbitrary units) of NCRM1 clones from initial experiment using original CRISPR/Cas9 approach (triple-plasmid electroporation).*
The x-axis shows the clones tested, letters A to D nomenclature corresponds to clone nomenclature in section 4.1.1. The results for each clone are the average of 4 technical repeats. Luminescence levels acquired using a Tecan Spark plate reader.

PCR analysis in **Figure 4-4** shows that the NanoLuciferase sequence is mostly preserved. Clone B has a large omission as seen in **Figure 4-6** and hence no expression of NanLuciferase was expected. Clone D's PCR result also confirms that the sequence is present; however, no expression is detected. This might be due to two possible reasons: 1) The populations are not clonal (heterogeneous) and thus different subpopulations contribute fragments of DNA that would amplify with different primer pairs, even though there is not a sub-population that has a full intact insert; or 2) Upon insertion, a random indel occurred that caused a frameshift or an early termination of expression. Indeed, multiple bands of various sizes can be seen using primers L3-L4 that encompasses the NanoLuciferase sequence in **Figure 4-8**. The same pattern is

observed in the amplification that encompasses the entire HDR cassette as well. Another possible cause for these findings is the fact that identical linker sequences upstream and downstream of the HaloTag sequences of the reporter cassette were used, which can lead to spontaneous rearrangements, including deletions, inversions and excisions.

Afterwards, I decided to further investigate the NanoLuciferase reporter functionality in the clones that showed high expression levels, to investigate temporal expression patterns. The two clones, A and C, were used for this experiment; however, clone A cells spontaneously detached from the bottom of the culture dishes, hence no results are included for this clone. It was not possible to repeat this experiment at that stage, as this was limited by the limited backup of stage 2 myogenic progenitor cells and the commercial media used to expand these cells was discontinued. Alternative media were then tested and this is detailed in 4.5.1.

Clone C myoblasts were differentiated, and the luciferase levels were measured in comparison to iPS cells and myoblast of the same cell line as a control to further confirm that the expression is linked to dystrophin production. In **Figure 4-28** a general trend of consistent increase was observed between day 4 and day 8 in stage 3 of differentiation.

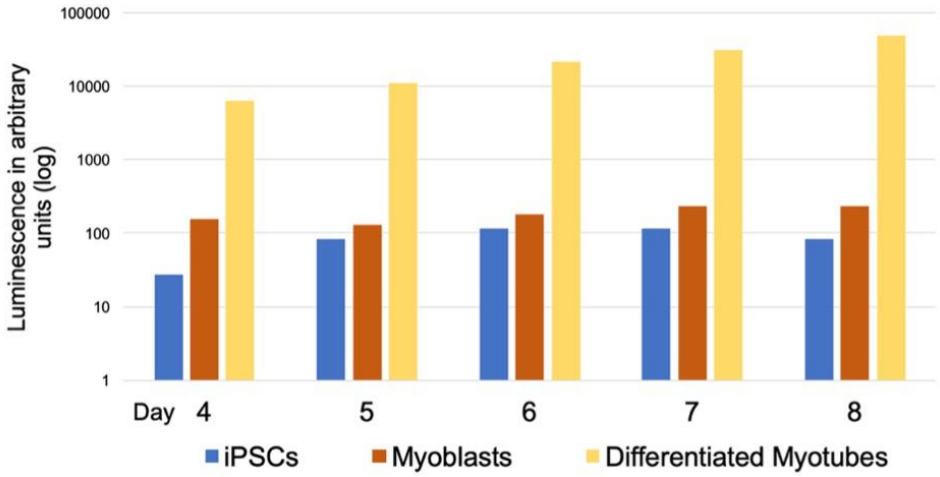

Figure 4-28. *Luminescence levels of NCRM1 clone C obtained using original CRISPR/Cas9 approach (triple-plasmid electroporation).*
The x-axis shows the days in terminal (stage 3) differentiation culture media. Human iPS cells and Myoblasts of the same cell line/clone were used as a negative control for this experiment. The results for each day are the average of 4 technical repeats.

These results were then compared to the differentiation levels of the myotubes assessed using immunofluorescence staining with the Titin antibody and quantified in the form of % of nuclei in Titin+ myotubes in **Figure 4-29** and a correlation was observed. Titin is an essential skeletal muscle elastic protein, and hence is used as a skeletal muscle differentiation and maturation marker (Monroy *et al.*, 2012).

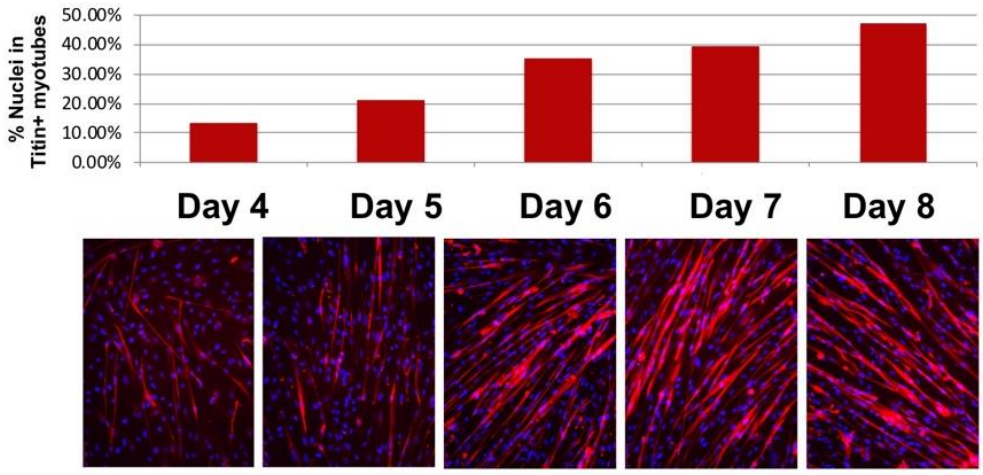

Figure 4-29. *Progressive increase of levels of titin detected using immunofluorescence staining of differentiating myotubes from day 4 to 8 of differentiation.*

Top graph shows quantification of the number of nuclei in dystrophin positive cells. No statistical analysis was done as all repeats were technical from the same experiment.

4.2.2.1.2 gRNP CRISPR/Cas9 Clones

The clones obtained from the RNP based CRISPR/Cas9 editing procedures and their subclones, in 3.4.2.1.2, 3.4.2.1.33.4.3.1, 3.4.2.1.4, were used to assess the luminescence functionality of NanoLuciferase in the cassette as detailed in 3.4.3.1. The first clonal reporter cell line population, NCRM1-L1O-A15, was used initially. The results show that myotubes differentiated for 9 days have a significantly higher luminescence level than their myoblast progenitors, in **Figure 4-30** confirming the functionality of the cassette.

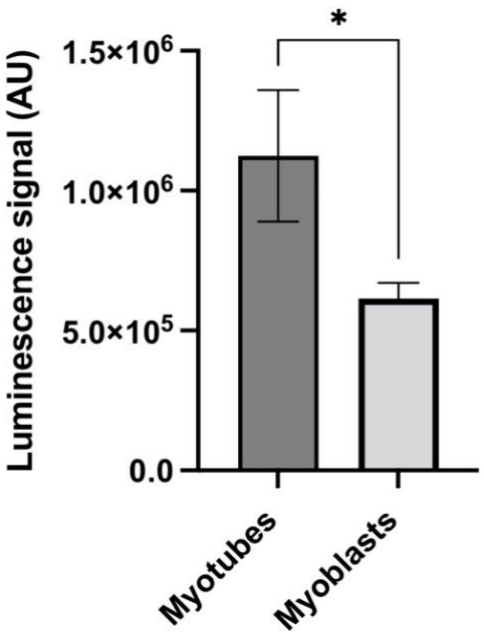

Figure 4-30. *Luminescence levels of NCRM1 L1O-A15 clone myoblasts and myotubes.*
*Myocea myoblast cells were differentiated for 9 days into myotubes, myoblast cells are plated shortly in advance to ensure same confluence. The results for both conditions are of 3 technical repeats. * = P <0.05.*

To assess the suitability of the NanoLuciferase in reporting dystrophin protein production levels at different stages of differentiation, iPS cells, myoblasts and myotubes at different days of differentiation were grown and levels were detected on a plate reader. Results in **Figure 4-31** show increased signal of luminescence in reporter iPS cells compared to isogenic wild type NCRM1 cells without a reporter, suggesting iPS cells express basal levels of a dystrophin isoform. Reporter myoblasts and differentiated myotubes show increasing signal over time, significantly higher than that of reporter iPS cells.

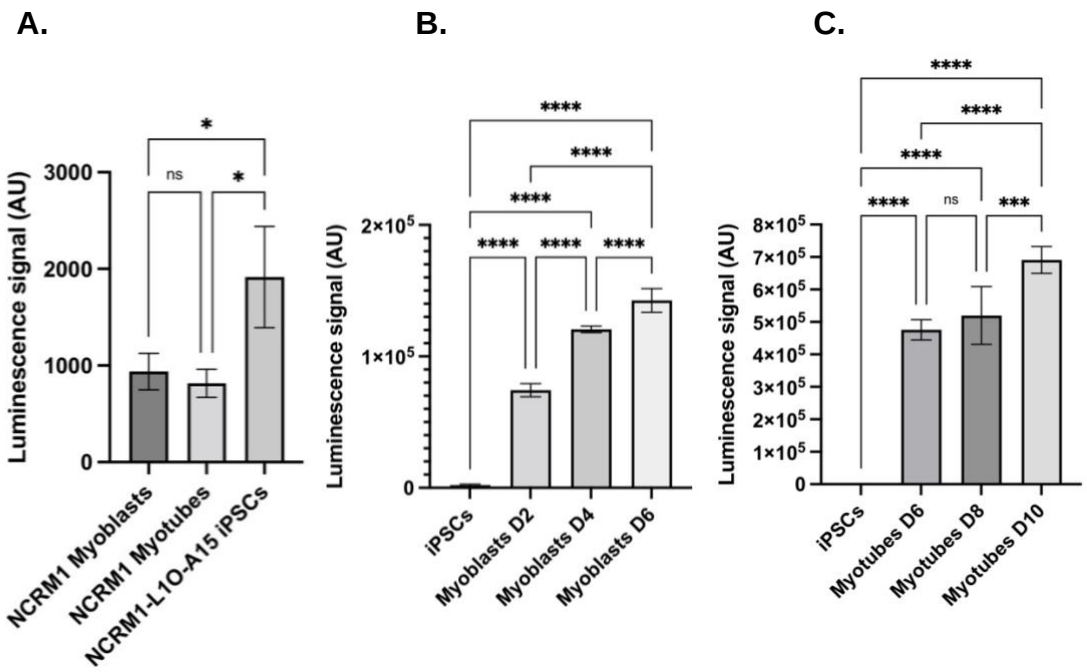

Figure 4-31. *Luminescence levels of NCRM1 L1O-A15 reporter clone at different timepoints and stages of differentiation.*

A. *Reporter iPS cells in comparison to wild type no reporter NCRM1 cells at myotube and myoblast differentiation stages. All cells were seeded at the same confluence.* ***B.*** *Reporter myoblasts seeded at the same confluence as iPS cells and readings taken at days 2, 4 and 6 in expansion media.* ***C.*** *Reporter clone differentiated myotubes at 100% confluence for 6, 8 and 10 days in differentiation media. Results are of at least 3 technical repeats for each condition. ANOVA results confirm that there is a significant difference of P= 0.0126 for the first graph and P <0.0001 for the two other panels. Tukey's multiple comparison test was used to determine pairwise differences and *= P <0.05, ***= P <0.001, **** = P <0.0001.*

Subclones obtained after flip recombination, to excise the blasticidin resistance cassette, were differentiated for 13 days. Results for luminescence, in **Figure 4-32**, confirm significantly higher luminescence levels than the isogenic

wild type control, which in this experiment is DMD L1 myotubes that do not possess the reporter and were differentiated for the same duration.

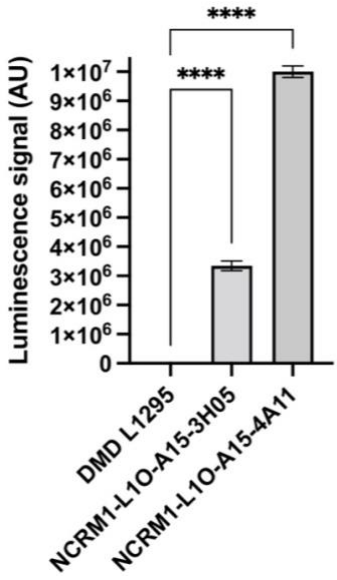

Figure 4-32. *Luminesce levels of NCRM1 L1O-A15-3H05 and NCRM1 L1O-A15-4A11 myotubes.*

*Myocea myotubes cells were differentiated for 13 days. DMD L1 myotubes were used as a negative control. The results for all conditions are of 3 technical repeats. ANOVA results confirm that there is a significant difference of P <0.0001 between clones and control; they also confirm that there is no significant difference between repeats of every cell line. Dunnett's multiple comparison test was used to determine pairwise differences and **** = P <0.0001.*

Reporter lines were also used to assess dystrophin protein production at different time points in monolayer differentiation, beyond day 8. Results in **Figure 4-33** show absence of significant differences between luminescence levels obtained on days 8, 10 and 12 of differentiation.

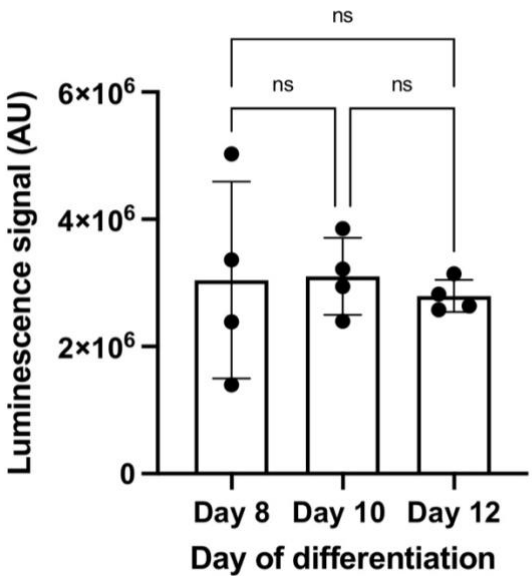

Figure 4-33. *Luminesce levels of NCRM1 L1O-A15-3H05 and NCRM1 L1O-A15-4A11 myotubes over time.*
Myocea myoblast cells were differentiated for up to 12 days into myotubes. The results shown are that of biological repeats of each clone (4 in total). ANOVA results confirm that there is no significant difference between timepoints due to technical repeats.

Finally, the isogenic exon 52 deleted clones were differentiated for 4 days, and their luminescence levels assessed. **Figure 4-34** shows that all the exon 52 deleted clones have significantly lower levels of luminescence compared to the wild-type reporter line NCRM1-L1O-A15-3H05. They also have a much higher luminescence than NCRM1 wild-type line without the reporter, which is likely due to Dp71 isoform expression.

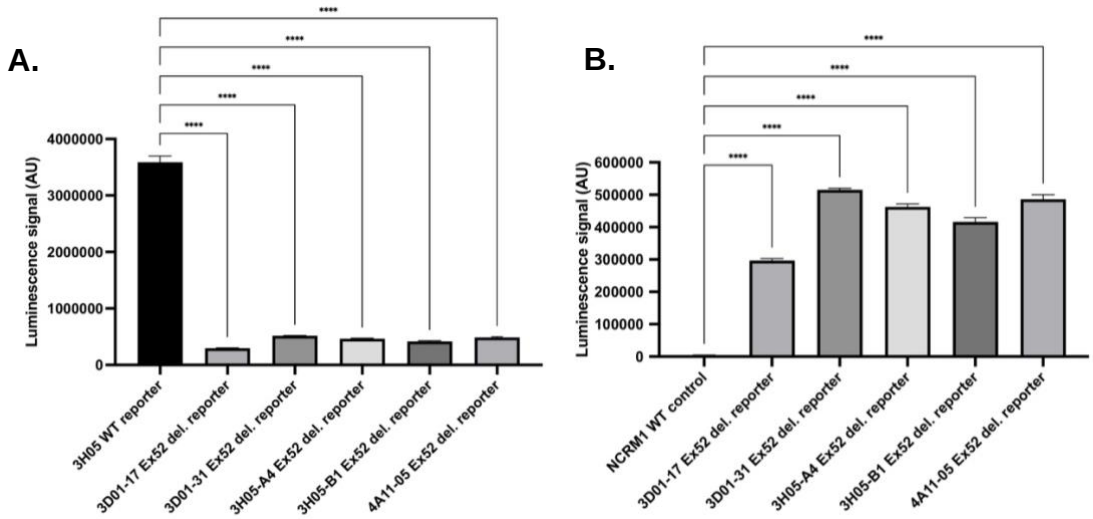

Figure 4-34. *Basal luminescence levels in exon 52 deleted reporter clones in comparison to controls.*

*Luminescence levels acquired from 4 day differentiated Myocea myoblast cells according to standard methods. Results are for 3 repeats per condition. ANOVA results confirm that there is a significant difference of P <0.0001 between Exon 52 deleted clones and controls; they also confirm that there is no significant difference between repeats of every cell line. Tukey's multiple comparison test was used to determine pairwise differences and **** = P <0.0001. Error bars represent standard deviation.*

4.2.2.2 mCherry Fluorescence

The second component of the reporter cassette, the mCherry sequence, was checked throughout differentiation of various clones and their derivatives.

4.2.2.2.1 Triple-plasmid Electroporation CRISPR/Cas9 Clones

During differentiation of the four clones in monolayer cultures, no expression signal of live mCherry fluorescence was detected. Clones were then differentiated on 96-well plates for 8 days and run on the microplate reader in efforts to detect faint signals that are invisible to the naked eye, however mCherry

fluorescence was not detected as seen in **Figure 4-35.** Cultures were then fixed and immuno-stained using an mCherry antibody, to assess the protein expression regardless of the signal intensity, which might be weak due to masking or misfolding and therefore difficult to detect under the microscope and using a plate reader.

The PCR results in **Figure 4-4** show the absence of the bands corresponding to the mCherry sequence in the cassette, which explains the absence of mCherry expression in clones B and C. PCRs for clones A and D showed the expected band sizes but the mCherry live fluorescence signal was not detected. This can be due to two reasons, either the populations are not clonal, or a random rearrangement occurred in the cassette. Another possibility that might explain the lack of mCherry fluorescence is that the signal might be weak and, therefore, an immunofluorescence staining using an mCherry antibody was carried out. However, no signal was detected, **Figure 4-35**. Overall, the results are consistent with the genomic aberrations found through PCR, and the fact that populations could be heterogenous. This led me to repeating the CRISPR/Cas9 gene editing experiments and generating new clones.

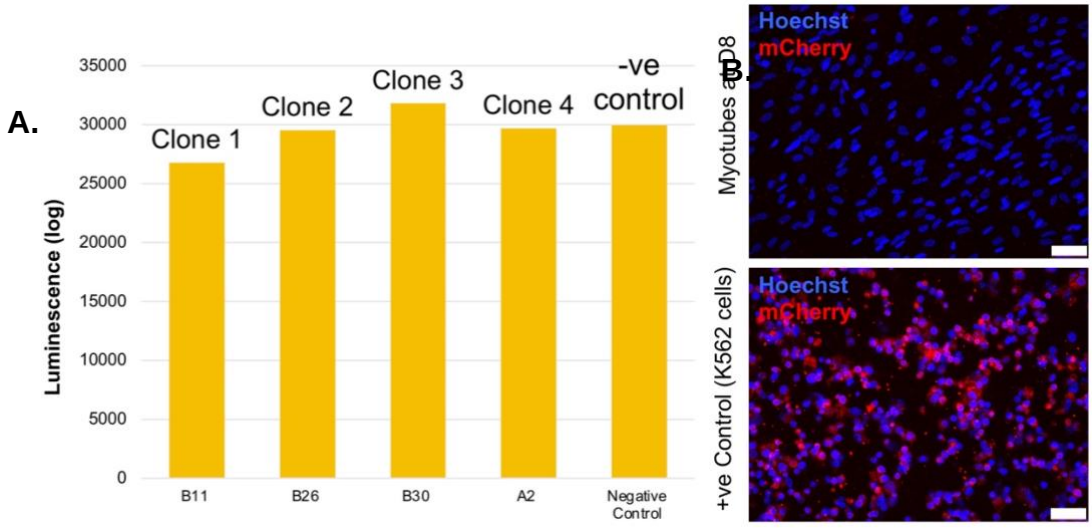

Figure 4-35. *Immunofluorescence staining and live fluorescence intensity graph for mCherry in reporter clones obtained from original triple-plasmid electroporation CRISPR/Cas9 experiments.*

A. *Immunofluorescence staining for mCherry of a representative well from one of the differentiated clones at day 8 of differentiation media step, compared to a positive control cell line from our lab.* ***B.*** *Live fluorescence levels obtained using plate reader for the four clones compared to iPS cells of NCRM1 D clone (used as a negative control). Results are of 3 technical repeats. Scale bar = 75 µm.*

4.2.2.2.2 gRNP CRISPR/Cas9 Clones

Upon single cell sorting of the NCRM1 L1O into clonal populations and confirmation of reporter insertion in-frame in the target locus of sub-clone NCRM1 L1O-A15, it was differentiated, along with the isogenic wild type NCRM1 cells without the reporter as a negative control, using the transgene free method (Myocea kit). After 12 days of differentiation from myoblasts to myotubes, cells were incubated with the fluorescent HaloTag ligand according to the manufacturer's instructions and acquired images using the Evos microscope system. No significant discernible mCherry fluorescence was observed in

NCRM1 L1O-A15 reporter cells, compared to a faint but discernible signal in the positive control HeLa cells, in **Figure 4-36**. Very high background levels of HaloTag ligand fluorescence were observed in both NCRM1 L1O-A15 reporter and isogenic wild type no reporter control cells, compared to a specific signal in MDCK positive control cells.

I hypothesised that this can be due to protein misfolding or steric hinderance as a result of the interaction of the blasticidin resistance sequence that is just after the HaloTag sequence (Chen, Zaro and Shen, 2013). While our reporter design includes linker sequences between the dystrophin exon 79, NanoLuciferase, HaloTag and mCherry sequences, there is no linker between the mCherry sequence and the blasticidin resistance cassette. It is well documented that the lack of linkers between components of a fusion protein leads to misfolding and impaired bioactivity (Chen, Zaro and Shen, 2013).

In efforts to restore the functionality of the reporter sequences, I used flip recombination to remove the blasticidin resistance sequence and the obtained NCRM1 L1O-A15 clone based on the hypothesis that the presence of the blasticidin resistance sequence disrupts mCherry expression, discussed in detail in 5.15.2. The clone NCRM1 L1O-A15-3H05 was differentiated for 9 days and used for the detection of mCherry signal. Under some of the differentiation conditions, a weak mCherry signal was observed and nuclear exclusion, highlighted with white arrows in **Figure 4-37** suggest that the signal is true and detectable.

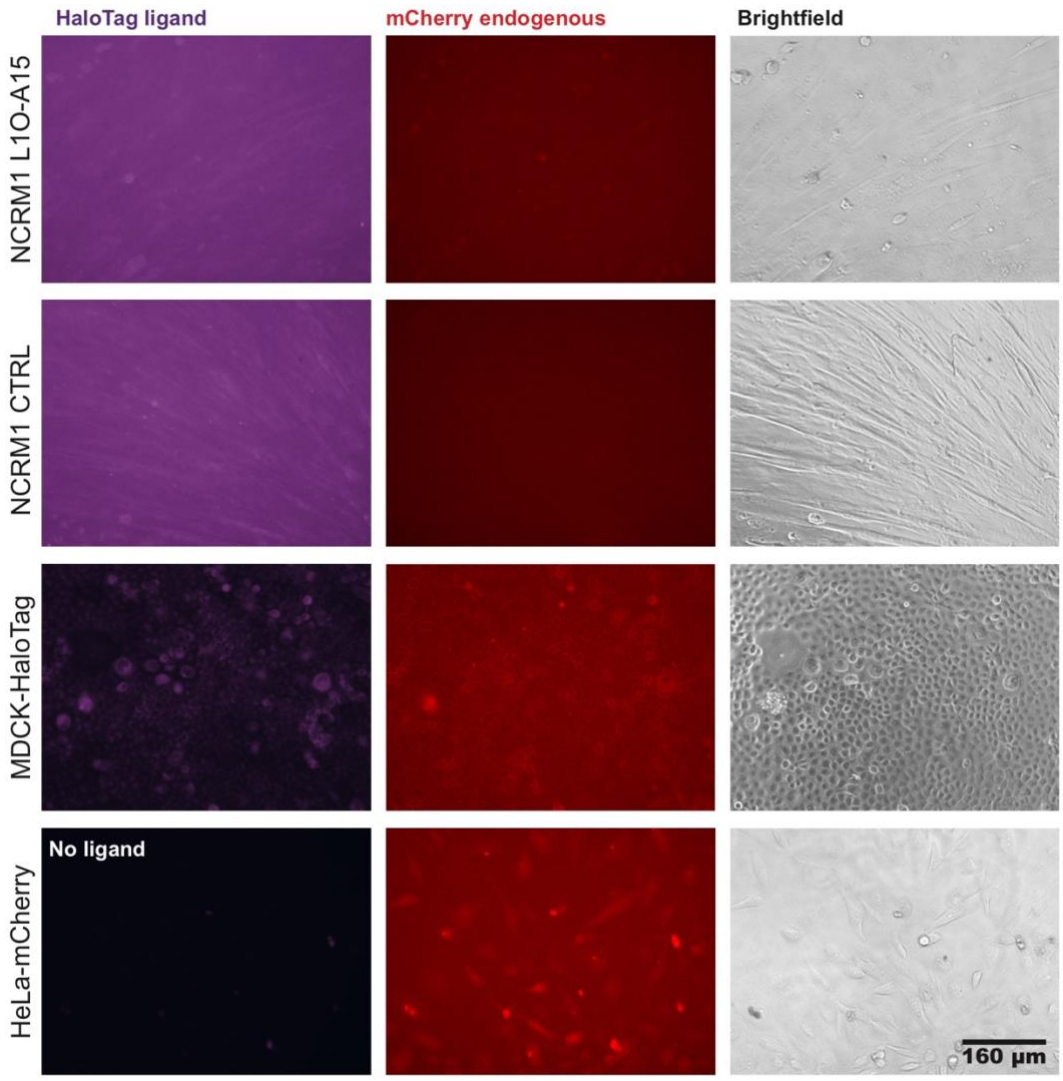

Figure 4-36. *Live imaging of NCRM1 L1O-A15 wild-type reporter clone for mCherry and HaloTag live ligand.*

Representative images of myocea myoblast derived 12 day differentiated myotubes imaged using Evos inverted microscope. For HaloTag ligand, Janelia Flour 646, cells were incubated overnight according to manufacturer instructions. MDCK cells expressing HaloTag and HeLa cells ubiquitously expressing mCherry were used as controls.

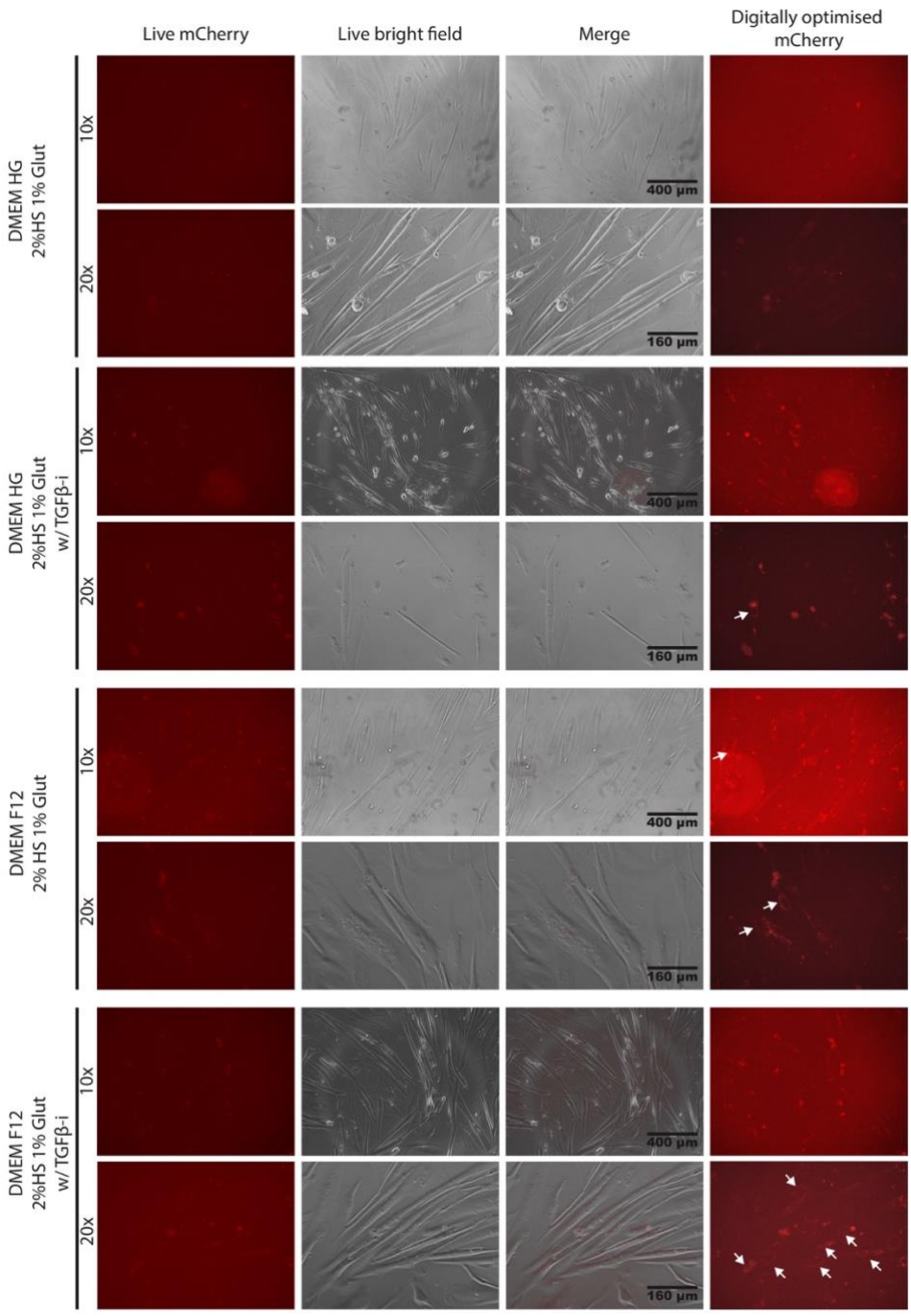

Figure 4-37. *Live fluorescence images of WT dystrophin reporter iPS cell-derived skeletal myotubes at 9 day of muscle differentiation (clone NCRM1-L1O-A15-3H05) using inverted Evos microscope.*

Differentiation was started with different media for 24 hours and (listed on Y axis of figure) then all conditions were switched to DMEM High Glucose with 2% Horse

serum and 1% Glutamax supplemented with TGFB-inhibitor for 72 hours, media was subsequently changed to DMEM without phenol and incubated until imaging. Images on the right are enhanced using FIJI software. Arrows indicate areas of nuclear exclusion.

In order to better assess the mCherry signal, another differentiation of the same clone was conducted on optical bottom dishes to better visualise the signal under a confocal microscope imaging to capture the localisation of the signal across a slice. A clearer mCherry signal under the confocal microscope with very distinct membrane localisation in most of the myotubes was observed using this alternative setup (**Figure 4-38**). The figure also shows the signal intensity along a transverse line plotted on one myotube from every field.

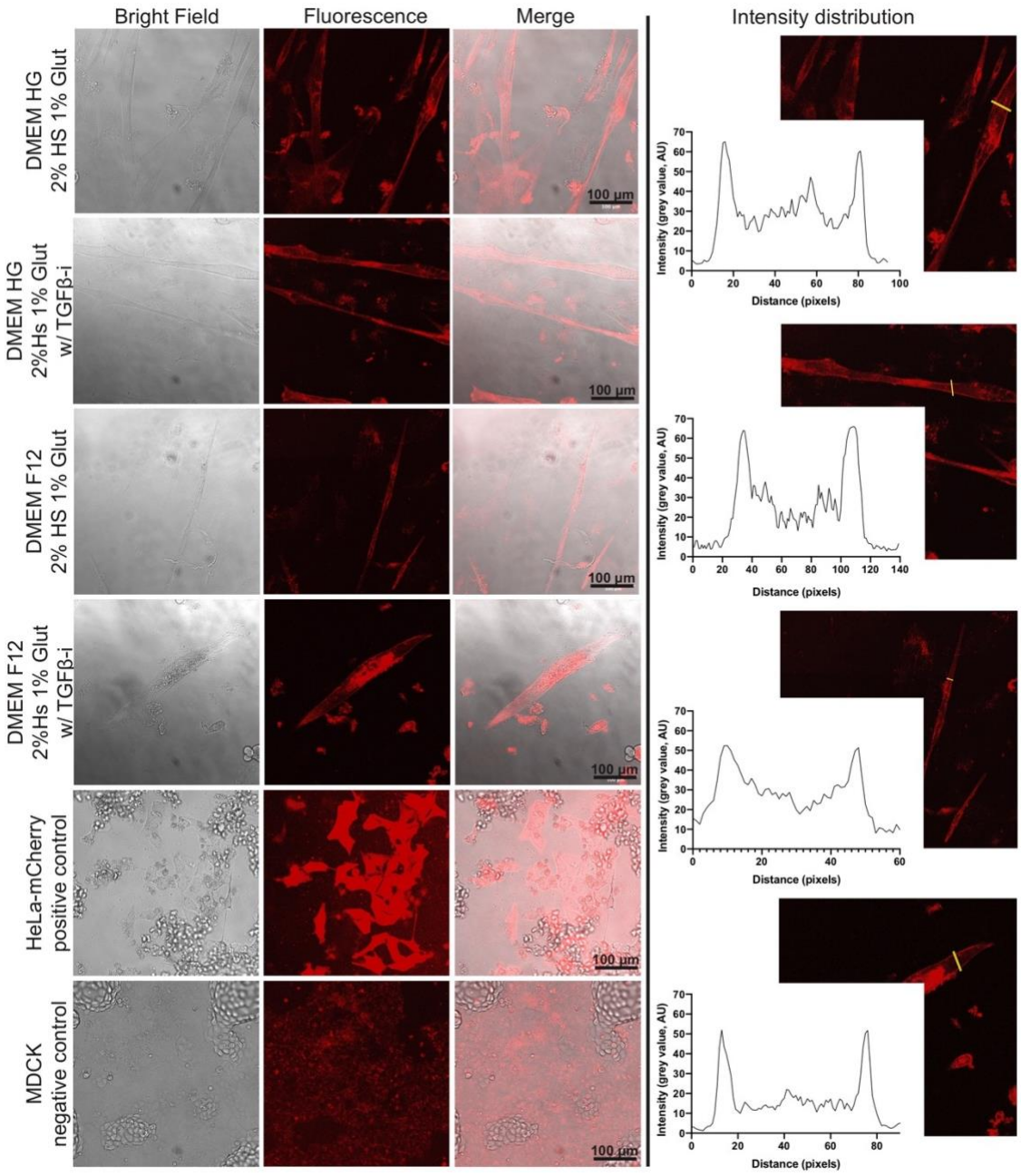

Figure 4-38. *Live fluorescence images of WT dystrophin reporter iPS cell-derived skeletal myotubes at 9 day of muscle differentiation (NCRM1-L1O-A15-3H05; days 9-12 of differentiation).*

Imaging was carried out using a Zeiss LSM 710 confocal microscope, with a 20X apochromatic objective using a 2 µm slice. **A.** *Differentiation media listed on y-axis.* ***B.*** *Intensity values along 10 pt line plotted using FIJI.*

Further differentiation of the WT reporter NCRM1 L1O-A15-3H05 for 13 days shows better localisation and distribution signal of live mCherry reporter, **Figure 4-39**. The acquired images show a very clear membrane localisation of the signal in myotubes “sliced” centrally (**Figure 4-39**). In myotubes where the slice encompasses the membrane, distribution of the signal can be seen clearly and white arrows in **Figure 4-39** point towards areas where the signal arranges longitudinally parallel to each other along the axis of elongation of the myotube. This distribution is characteristic of differentiating myotubes and markers such as myosin and titin (Chal *et al.*, 2016).

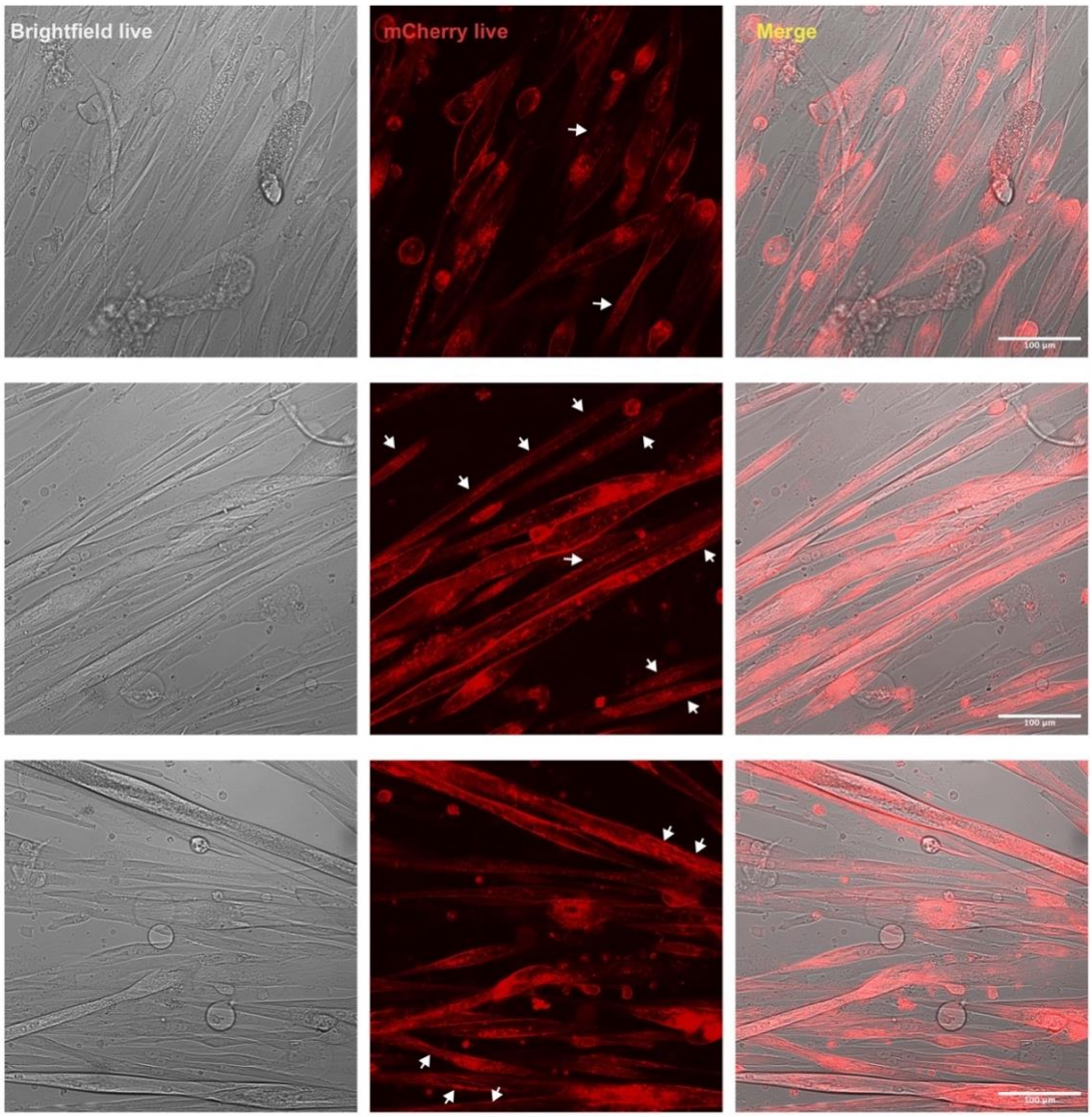

Figure 4-39. *Live confocal imaging of day 13 differentiated wild-type dystrophin reporter 12-day differentiated myotubes showing mCherry localisation.*
Three representative images of day 13 differentiated NCRM1-L1O-A15-3H05 cells. Images are of one slice of 1 µm. Membrane localisation can be clearly seen, and arrows highlight myotubes in which mCherry signal arranges longitudinally. Scale bar = 100 µm.

To further validate the mCherry signal, an immunofluorescence staining of wild-type NCRM1 cells with no reporter, NCRM1-L1O-A15-3H05 cells with

reporter, and NCRM1 L1O-A15-3H005-A4 exon 52 deleted isogenic cells with the reporter differentiated for 12 days was carried out. An mCherry antibody, dystrophin antibody from abcam (which has been validated in western blotting by in iPS cell-derived cultures, (Rani *et al.*, 2020)) and MF20 antibody for myosin heavy chain as a differentiation marker were used. As the dystrophin antibody and the mCherry antibody were both raised in rabbit, sister differentiated cell culture wells were stained to check whether the staining patterns match before finding an alternative mCherry antibody. Results in **Figure 4-40 A)** show that the dystrophin antibody gives a positive staining in the three conditions, including the exon 52 deleted clone, which is thus unreliable under the conditions tested here suggesting presence of background or non-specific signal. This is supported by the fact that signal seems to be non-specific, staining all cells, and with no apparent localisation. However, the mCherry staining, in **Figure 4-40 B)** show no fluorescence in the wild type NCRM1 cells without the reporter, as they do not possess an mCherry fusion protein, a strong signal in the wild type reporter 3H05 cells and a very faint signal in exon 52 deleted 3H05-A4 clones which is consistent with the NanoLuciferase results obtained earlier.

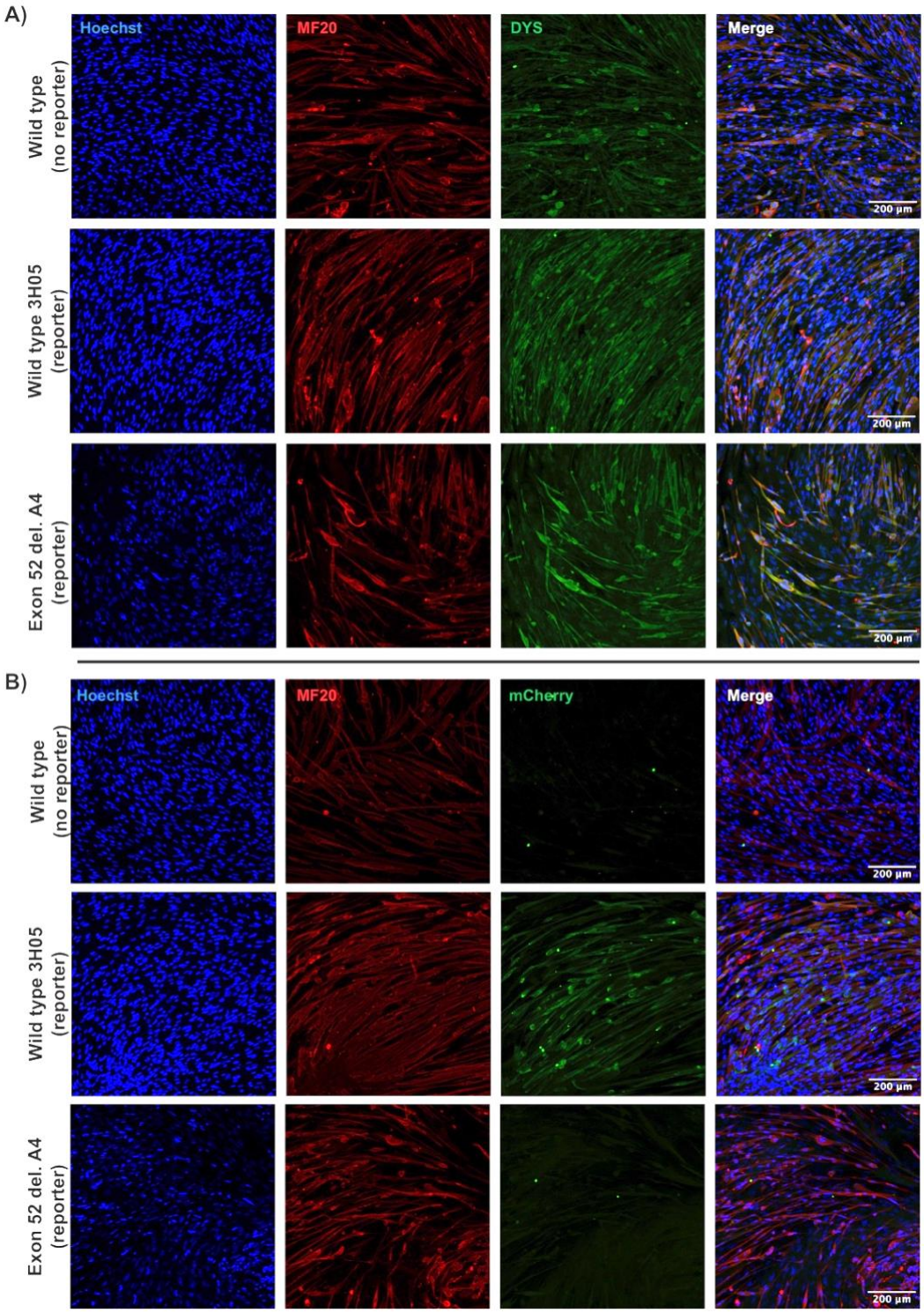

Figure 4-40. *Confocal imaging of immunofluorescence stained fixed differentiated myotubes from clones for mCherry and dystrophin.*

Clones were differentiated for 12 days. A) Immunofluorescence staining using Hoechst, MF20, and mCherry. B) Immunofluorescence staining using Hoechst, MF20 and dystrophin (Abcam). Imaging using LSM710 confocal microscope using a 20X objective. Images are one slice of 1µm. Immunofluorescence staining and imaging performed by Dr. Valentina Lionello.

To further confirm the signal, high resolution imaging for mCherry was performed and dystrophin stained 4A11 cells and exon 52 deleted 3H05-A4 cells for dystrophin, the results, in **Figure 4-41**, a clear membrane localisation of the mCherry antibody signal in differentiated wild type reporter 4A11 clone is observed. However, staining of dystrophin on the same cells shows a generalised signal with no localisation. The generalised signal is also very clear in exon 52 deleted 3H05-A4 differentiated cells, further highlighting the challenges of using the dystrophin antibody, and the effort required to optimise and validate protocols for its detection for staining, and confirming the accuracy and ease of use of the reporter in detection of dystrophin through mCherry staining.

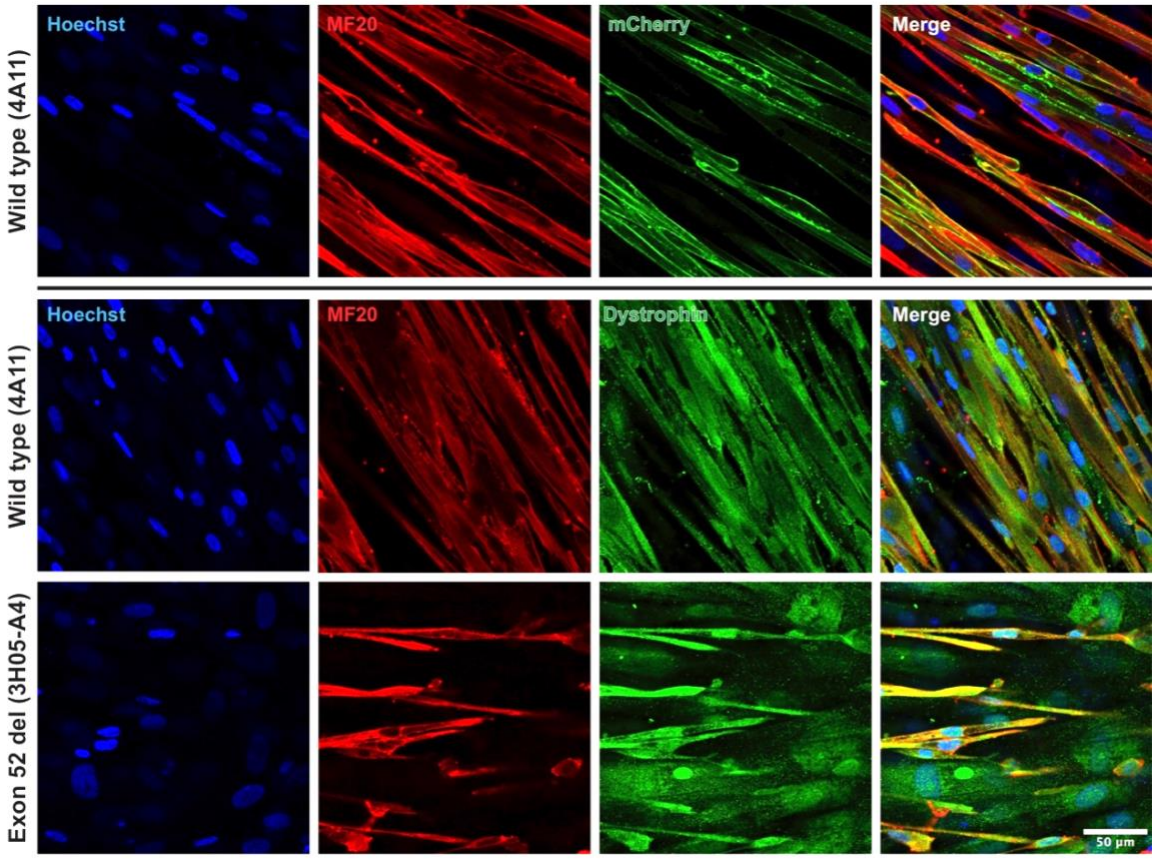

Figure 4-41. *High magnification confocal imaging of fixed differentiated myotubes from 4A11 wild type reporter and 3H05-A4 exon 52 deleted clones for mCherry and dystrophin.*

Clones were differentiated for 12 days. Immunofluorescence staining using Hoechst, MF20, and mCherry or dystrophin (Abcam). Imaging using LSM710 confocal microscope using a 40X objective. Images are one slice of 1 µm. Scale bar = 50 µm. Immunofluorescence staining and imaging performed with Dr. Valentina Lionello.

4.2.2.3 HaloTag for Western Blotting

To confirm the functionality of the HaloTag protein, a western blot for wild-type reporter cells (NCRM1 L1O-A15-3H05) differentiated for 12 days in a monolayer was utilised for this experiment. For dystrophin, the RB-DYS antibody

from Abcam, that has been validated for use in iPS cell-derived cultures' western blots by another lab (Rani *et al.*, 2020), the HaloTag antibody and an mCherry antibody were used. The bands detected by the HaloTag and mCherry antibodies match those of the RB-DYS antibody and match the expected bands for Dp427 and Dp260, detect dystrophin muscle isoform Dp427 and retinal isoform Dp260, which has been shown to be expressed in skeletal muscle cells (Warner *et al.*, 2002),in **Figure 4-42**, taking into account the size of the reporter insert, which is 82.8 kDa. Extra bands between 171 and 238 kDa can be seen with the dystrophin antibody but they do not correspond to any known isoform sizes. Finally, the HaloTag antibody and, at lower efficiency, the mCherry antibody, were able to detect the ubiquitously expressed Dp71.

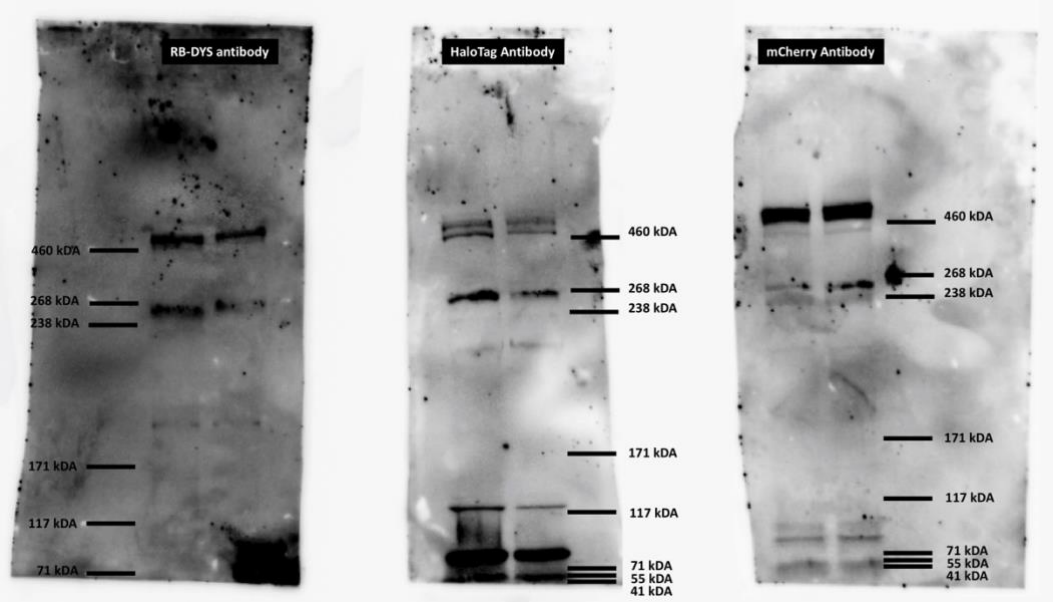

Figure 4-42. *Western blotting analysis for wild-type reporter clone 3H05 using dystrophin, HaloTag and mCherry antibodies.*

Image is an overlay of chemiluminescence images and colorimetric image, both acquired on same imaging device, but automatic overlay/combination gives very poor image. For every antibody, two lanes of proteins are loaded, for the left lane the protein extract was boiled at 70 °C and the right lane proteins were boiled at 90 °C. Western blot performed with Dr. Valentina Lionello.

The results reported here in section 4.2 confirm the functionality of the dystrophin reporter components, luminescence, fluorescence and interchangeable labelling in the clones obtained from section 4.1. The section also details and demonstrates uses and techniques for the various components.

4.3 Towards a Dystrophin Reporter Bioengineered Human Muscle: Assessment of Human iPS Cell-Derived Skeletal Myotubes in 2D and 3D cultures

Parallel to the work reported in the previous two sections, the potential of 3D skeletal muscle constructs as a modelling platform for DMD and its advantages over monolayer cultures was investigated; with the aim of combining the reporter lines platform with the 3D cultures to develop a more advanced DMD modelling platform.

4.3.1 Assessment of Myotube Maturity in 2D vs 3D Cultures

In efforts to better characterise the 3D skeletal muscle construct model and assess its potential as a more faithful and biologically relevant model, qPCR analyses for maturation markers and dystrophin was performed. The initial results from UCLi007 HIDEMs show that at the same time points/days of differentiation,

3D cultures express higher levels (at least 8-fold increase) of dystrophin and while the levels in monolayer cultures seem to have plateaued between days 8 and 12 of differentiation, 3D muscle constructs show an increase from 8 to around 13-fold by day 14 in comparison to monolayer cultures **Figure 4-43**. The following graphs in **Figure 4-43** show the three myosin heavy chains maturation markers, *MYH1* and *MYH2* (adult markers), and *MYH8* (neonatal marker), all of which 3D muscle constructs show a higher expression, but particularly in *MYH1* and more importantly *MYH8*.

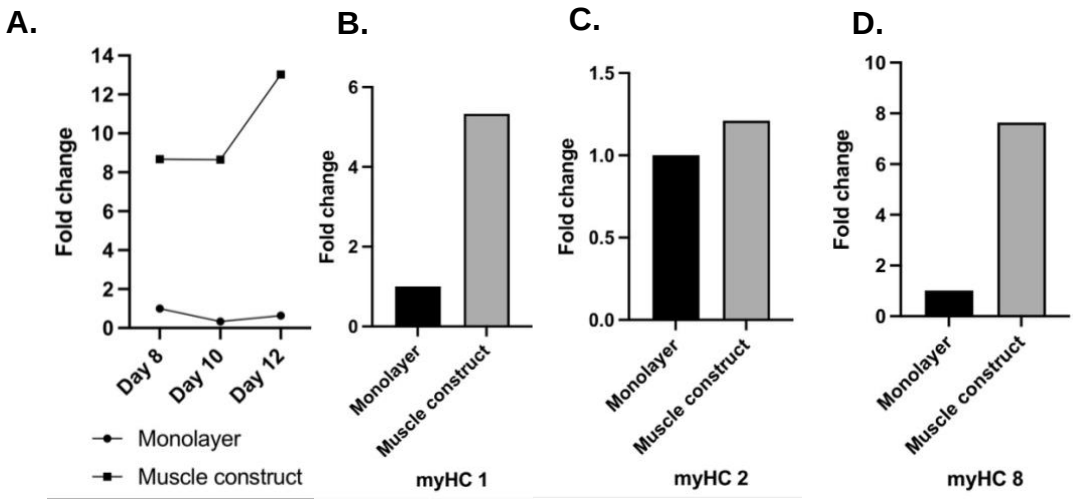

Figure 4-43. *qRT-PCR analysis for myosin heavy chains muscle maturation markers and dystrophin in monolayer and 3D muscle constructs.*
qRT-PCR analysis for the muscle maturation markers myosin heavy chain 1, 2 and 8 in addition to dystrophin showing their level of expression in UCLi007 HIDEMs 3D muscle constructs with compared to monolayer cultures of the same line (an established line in the lab). ***A.*** *Dystrophin expression over time.* ***B. C.*** *and* ***D.*** *Expression at day 8 for the myosin heavy chain markers. Expression (ΔCt) is normalised against RPLP0, and relative expression (ΔΔCt) is normalised against levels of expression of day 8 of monolayer cultures. Results presented as fold*

change. Results are of a minimum of 3 technical repeats and thus no statistical analyses have been done.

The results confirm that differentiation of iPS cell-derived myogenic cells in a 3D environment leads to higher expression of maturation markers than monolayer cultures.

4.3.2 Dystrophin Protein Production in 3D muscle Constructs Through Macro-imaging

One major advantage of 3D cultures over monolayer cultures is that they provide a more physiologically relevant morphological and spatial environment. For skeletal muscle, this also entails exerting tension along one axis which drives the alignment and organisation of differentiating myotubes. To visualise the macro-architecture of the muscle constructs, light sheet microscopy was used for whole 3D muscle construct imaging. With the extra aim of imaging of the entire depth of the construct, generating a complete view of fluorescence signal and reconstructing it in 3D.

The resultant scan obtained from light sheet microscopy does not go beyond a few myotubes in depth, around 100 µm, **Figure 4-44**. However, the external structure, organisation, and alignment of the skeletal muscle cells of the construct can be clearly visualised in the 3D reconstruction. The left image in the figure is a snapshot of digital-3D reconstruction of the artificial skeletal muscle, showing fair alignment, albeit twisted which might be a post-culture artefact, during fixing for example. The image on the right is a digital 3D reconstruction and rendered

video showing more clearly how superficial the image acquisition/staining detection is- video attached.

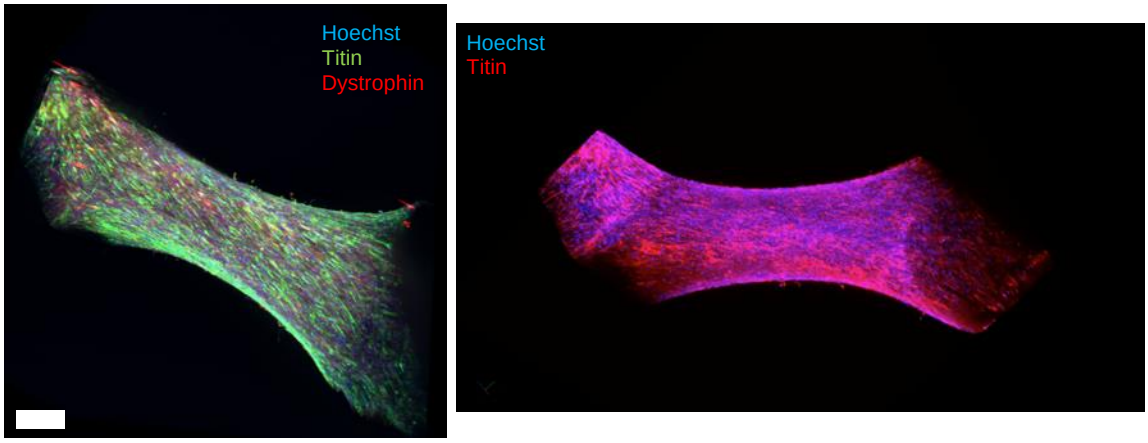

Figure 4-44. *3D reconstructions of a light-sheet image acquisition of an Immunofluorescent stained single lineage 3D muscle organoid derived from NCRM1 HIDEMs.*
On the left is a rendered image with Hoechst in Blue, Titin in Green and Dystrophin in Red. The right pane is a video (right click to play) of the same 3D reconstructed organoid but Titin is rendered in Red, and Hoechst is in blue. Scale bar = 1 mm.

Once the reporter lines were established and differentiated to the progenitor stage, 3D muscle constructs were made using the wild-type reporter NCRM1 L1O-A15-3H05 and its isogenic exon 52 deleted lines, A4 and B1 to assess their differentiation and architecture. For this, maximum intensity projection images from images obtained via spinning disk confocal microscopy were used to better visualise signal localisation and myotube architecture at a higher resolution than light sheet microscopy. The top panel of **Figure 4-45** shows maximum projection

images of CMFDA stained 3D muscle constructs from exon 52 deleted lines, shows presence of aligned multinucleated myotubes in exon 52 deleted clones along the axis of tension. The bottom panel is an image of one slice of the WT reporter line, in which live mCherry can be clearly seen at day 4 of differentiation, showing proper alignment, a higher magnification is required to better visualise sub-cellular localisation of the mCherry, however, this is limited by the working distance of higher magnification objectives available, as the constructs are suspended away from the bottom of the plate. For context, in monolayer cultures, mCherry live signal is only visible starting day 6 of differentiation and reaches similar fluorescence intensity levels on day 10. This is possibly due to the increased maturation and accelerated differentiation provided by the 3D muscle construct.

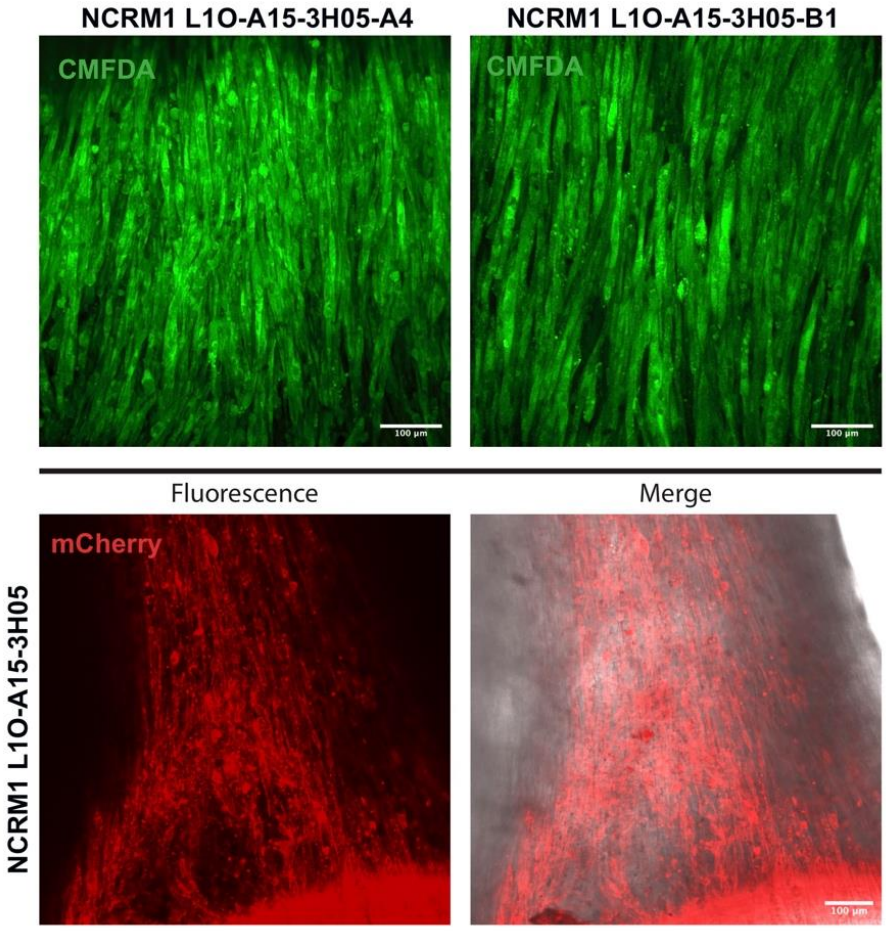

Figure 4-45. *Spinning disk and confocal live imaging of exon 52 deleted and wild type DMD reporter 3D muscle constructs.*

Top panel are representative images of day 10 differentiated 3D muscle constructs made with exon 52 deleted lines, to visualise myotubes, constructs were stained with CellTracker TM green CMFDA dye according to manufacturer's instructions and imaged using spinning disk microscope. Images are maximum intensity projections of a z-stack of 113 µm. Bottom panel is a representative image of live mCherry in day 4 differentiated 3D muscle construct made with wild type reporter line imaged with confocal microscope. Slice size is 1 µm and a line averaging of 4 was used.

Collectively, the results of section 4.3 confirm the potential and advantages of 3D muscle constructs over monolayer cultures. These results also lay the foundations for the use of DMD reporter lines in the 3D muscle constructs using live imaging strategies. Potential outcomes for this combined system would be useful in studies to assess therapies (such as AONs) penetration within the muscle construct- if conjugated with a fluorescence protein- and its correlation to restoration of dystrophin and its localisation using mCherry fluorescence.

4.4 USING DYSTROPHIN REPORTER IPS CELLS FOR DMD THERAPY TESTING USING AONS

4.4.1 SCREENING ANTISENSE OLIGONUCLEOTIDES (AONS)

Having confirmed the reporter cassette's components' functionality, as a proof of concept for the potential use as a therapy screening platform, reporter cell lines generated in section 4.1.3 were used to test AONs mediating exon skipping of dystrophin exon 53.

First, Commonly used methods to deliver AONs to myoblasts and iPS cell-derived myotubes are either chemicals such as Lipofectamine, Lipofectamine 3000 and Endo-porter, or transfection-based such as nucleofection. Since differentiating cells are impossible to detach without significant damage, options were limited to chemical reagents. Lipofectamine is a reagent that contains lipid subunits that form liposomes in aqueous environments, they also contain complexes that bind to negatively charged nucleic acids, this allows for the formation of a liposome containing the payload with a positively charged surface which fuses with the negatively charged plasma membrane allowing the payload to cross into the cytoplasm (Dalby *et al.*, 2004). It is considered the industry "gold standard" for the delivery of exogenous nucleotide sequences into the cells and has been the most cited transfection reagent family with tens of thousands of citations since its launch in 1993 (Cardarelli *et al.*, 2016). In contrast, Endo-porter is a novel peptide (Patented: US 7,084,248 B2) that binds to the plasma membrane of the cell, it acts through endocytosis (Summerton, 2005). Aung-Htut

and colleagues evaluated the transfection efficiency of PMOs in myoblasts using different reagents, their findings indicate that Endo-porter transfection efficiency is similar to that of Lipofectamine 3000 and nucleofection (Aung-Htut *et al.*, 2019). Henceforth, all experiments were conducted using Endo-porter given that the PMO manufacturer cites it as the transfection reagent of choice.

4.4.1.1 Evaluation of Transfection Reagent Toxicity

AONs that are not conjugated to a cell penetrating peptide need a delivery vehicle to channel them into the cytosol. Prior to use of the AONs' manufacturer's recommended channelling agent, Endo-Porter, a preliminary experiment to assess the toxicity of it on differentiating myotubes was run. Cells were seeded at a density that is half of what is normally used for differentiation experiments in this thesis, to allow for more surface area of the cell membrane to be exposed to the media and the agent, which was supplied at 3 different concentrations. Results show that there is no significant difference in cell death and/or detachment of the myotubes from the bottom of the plate between 2 µM, 4µM and 6 µM Endo-porter after 72 hours in culturing media as seen in **Figure 4-46**. The figure also shows that after keeping the cells in culture for a further 5 days, no significant detachment or cell death can be observed. Based on these results, I proceeded with using 6 µM for all experiments in which PMO AON is assessed.

It is worth noting that in all conditions where the Endo-porter was applied, several cell shape anomalies were observed (**Figure 4-46**), such as blobs. This may be due to the promotion of extra endo-cytosis by the reagent, stressing the

cell membrane. Despite being a natural biological process in most cell types, I hypothesise that the degree of endocytosis promoted by Endo-porter exceeds pathway capacity in normal endogenous functions. This observation was also reported by Aung-Htut and colleagues and upon investigation, they found reduced expression of Cited2, a protein downregulated under ER stress (Aung-Htut *et al.*, 2019). For the aforementioned reasons, Endo-porter is not used *in vivo*.

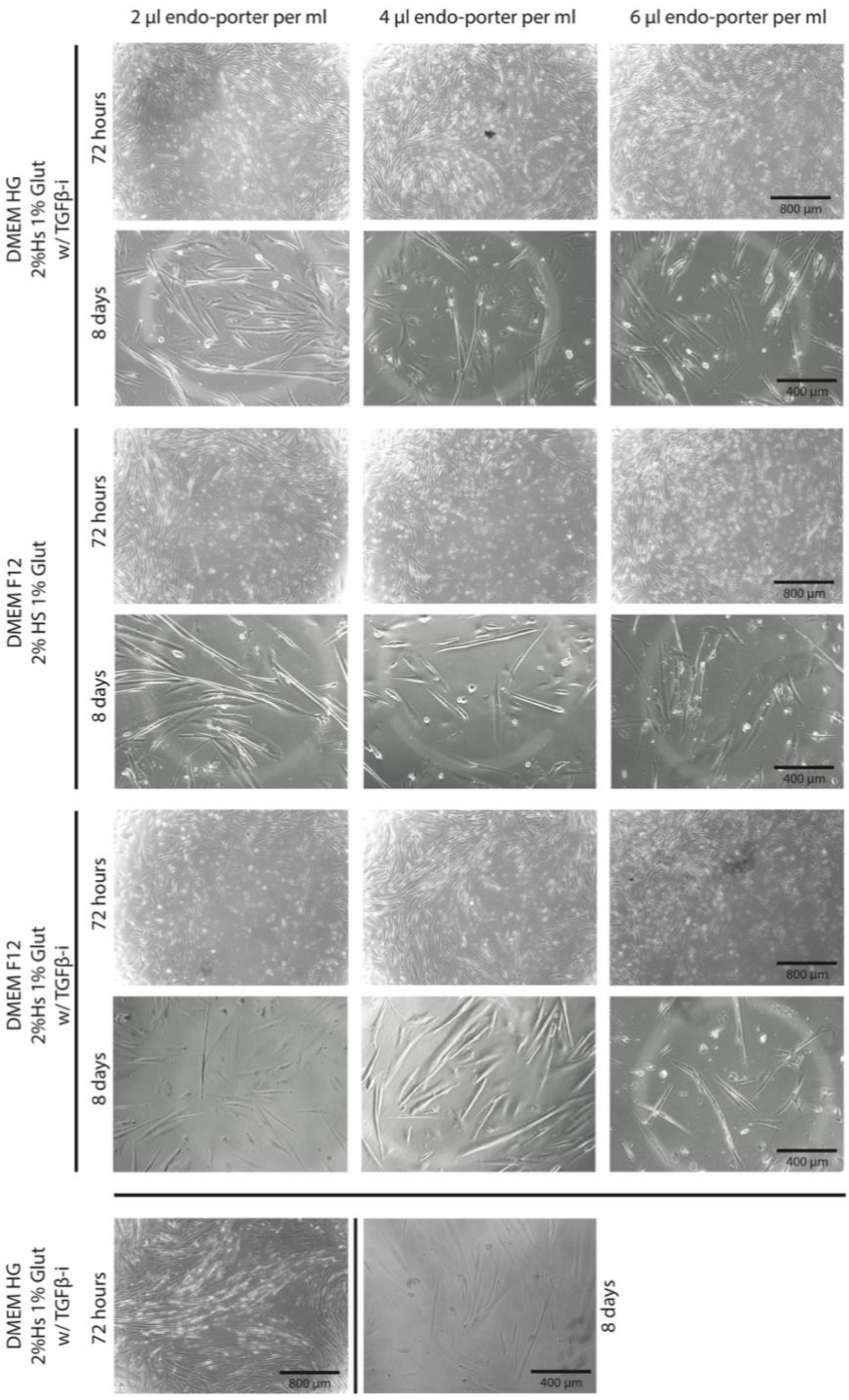

Figure 4-46. *Live imaging of iPS cell-derived transgene-free 9 day differentiated WT dystrophin reporter clone (NCRM1-L1O-A15-3H) incubated with different concentrations of Endo-porter.*

Imaging using Evos microscope at 4X and 10X magnification. Differentiation was started with different media for 24 hours and then all conditions were switched to DMEM High Glucose with 2% Horse serum and 1% Glutamax supplemented with TGFB-inhibitor for delivery of Endo-porter. After 72 hours, media was changed to DMEM without phenol red for imaging and incubated for a further 5 days for day 8 imaging.

4.4.1.2 Trials Using PMO AONs

To examine the potential of using the generated reporter lines in screening AON therapies, two experiments to test the delivery of a PMO using Endo-Porter were conducted. In the first experiment 1 day differentiated myoblasts were incubated for 72 hours with the PMO at 3 different concentrations (500 nM, 1000 nM and 1500 nM), along with 6 μM Endo-porter, then luminescence level was measured using the plate reader. No significant difference or increase in the luminescence levels between treated and untreated conditions was observed, as seen in **Figure 4-47**. This is possibly due to the fact that the AONs were added to cells on day 1 of differentiation, a timepoint at which Dp427 dystrophin protein production is not high, especially when factoring in the time needed for fully transcribing dystrophin (Tennyson, Klamut and Worton, 1995). In fact, results in **Figure 4-28** show that protein production levels in undifferentiated iPS cell-derived myotubes are around 100 times less than in day 4 differentiated myotubes.

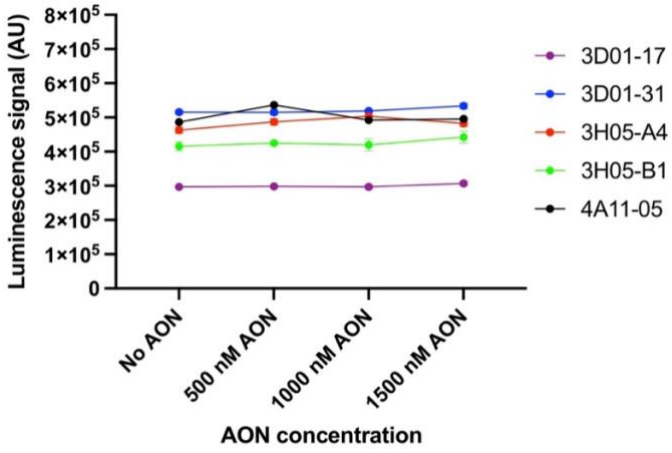

Figure 4-47. *Luminesce levels of exon 52 deleted clones treated with low doses of PMO AON.*

Myocea myoblast cells were switched to differentiation media and after 24 hours the AON was added, along with 6 μM Endo-porter. Cells were incubated with the AON for a further 72 hours and then luminescence levels were measured. The results for all conditions are of 3 technical repeats. ANOVA results show that there is no significant difference due to treatment conditions in comparison to no-treatment control; they also confirm that there is no significant difference between technical repeats of every cell line.

This prompted me to conduct another trial using myotubes differentiated for a longer time in culture, in combination with administering multiple doses of the PMO AON, at a higher concentration than what was used in the previous trial, specifically, 1.5 μM and 3 μM. This experiment was conducted using 3H05-A4 and 4A11-05 exon 52 deleted reporter clones only due to constraints in expansion of other clones, starting at day 6 of differentiation, was conducted. The AON was added at multiple timepoints either on days 6, 8 and 10, or days 8 and 10, or day 10 only. The results in **Figure 4-48** show that there is a significant increase in NanoLuciferase signal due to the number of doses administered, for

3H04-A4 this was only true at the 3 μM concentration, while for 4A11-05 this was true for both 1.5 μM and 3 μM. The results also show a significant increase in NanoLuciferase signal over control untreated cells at both day 10 and 12 (after 2 and 3 doses of the treatment, respectively) but not at day 8 (1 day of treatment) due to treatment, and the effect of number of doses of treatment is significant in increasing luminescence. Once again, for 3H04-A4 this was only true at the 3 μM concentration, while for 4A11-05 this was true for both 1.5 μM and 3 μM.

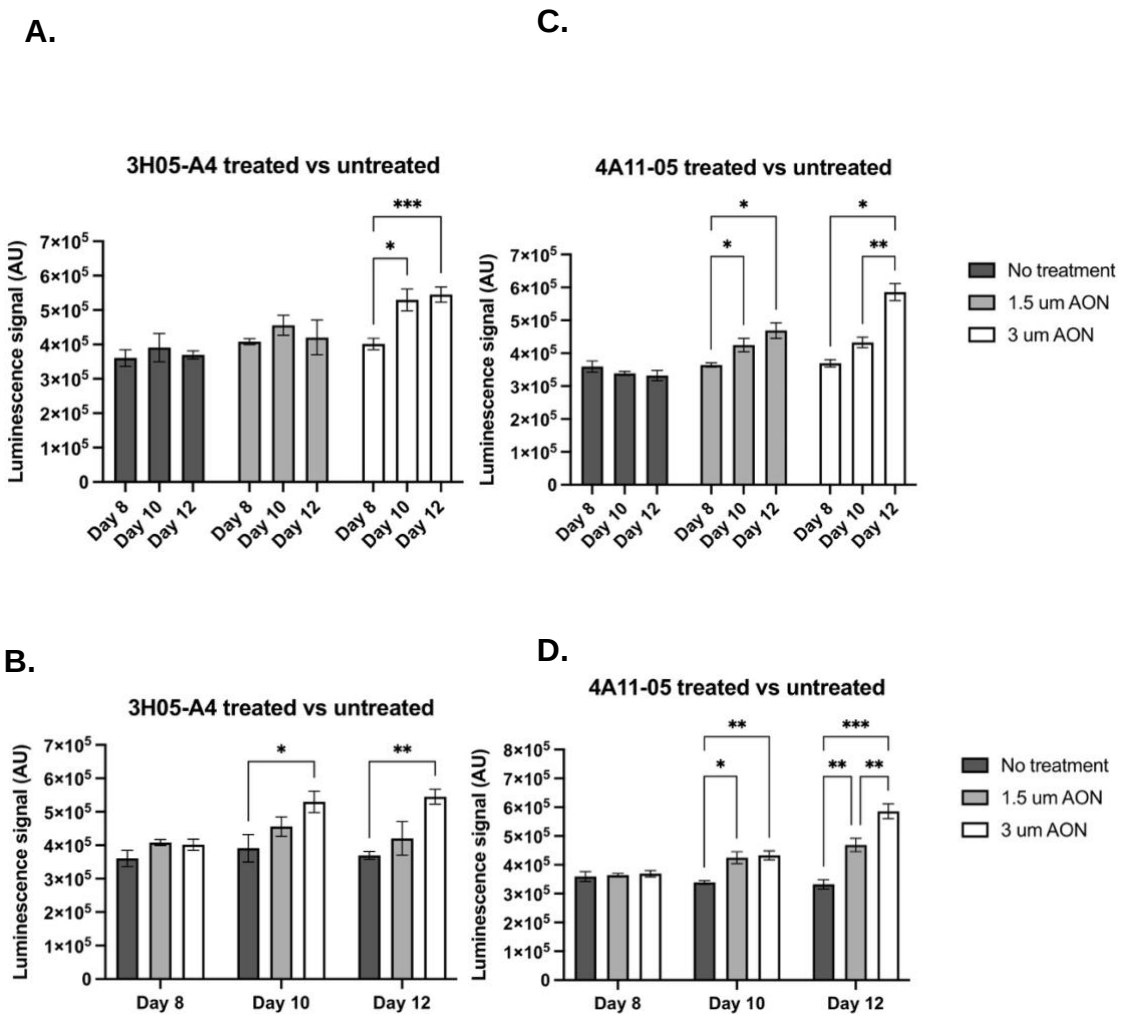

Figure 4-48. *Luminesce levels of exon 52 deleted clones treated with high doses of PMO AON.*

Differentiating myocea myoblast cells are incubated with AON and 6μM Endo-porter on days 6, 8 and 10. Luminescence levels measured on day 8 have been treated on day 6 only, wells measured on day 10 have been treated on day 6 and 8, and wells measured on day 12 have been treated on days 6, 8 and 10. ***A.*** *and* ***B.****, and* ***C.*** *and* ***D.,*** *are of the same data set, respectively, split for better visualisation of pairwise comparisons. The results for all conditions are of 3 technical repeats. ANOVA results show that there is a significant difference of P <0.0001 due to treatment duration and also the treatment concentration; they also confirm that there is no significant difference between technical repeats of every condition. *= P <0.1, **= P <0.01, ***= P <0.001. Error bars indicate standard deviation.*

4.4.1.3 Trials Using Peptide-conjugated PMO AONs

Another approach to the delivery of AONs to cells, is using arginine-rich cell-penetrating peptide conjugates that increase the efficiency of penetration while reducing cellular stress.

Prior to using the peptide-conjugated AON on exon 52 deleted clones, wild-type reporter clones were used to assess its efficacy, given the difficulties faced using the unconjugated PMO AON. The results show a significant decrease in fluorescence levels in peptide-AON treated differentiating cells compared to baseline (baseline being the average expression of day 8, 10 and 12 differentiated myotubes of the same line), in **Figure 4-49** indicating that peptide-conjugated PMO has a very high efficiency in cell penetration and skipping exon 53; given that skipping exon 53 in wild type reporter cells leads to knocking down dystrophin protein production. Levels of luminescence dropped to around 6×10^5 to 1×10^6 compared to more than 3×10^6 in the untreated group. These levels are

within the same range of exon 52 deleted reporter clones (in **Figure 4-48**) and can be used as an indirect indicator of the very high efficacy of the PPMO in skipping exon 53, leading to an out of frame mutation, knocking down the production of dystrophin isoforms larger than Dp116.

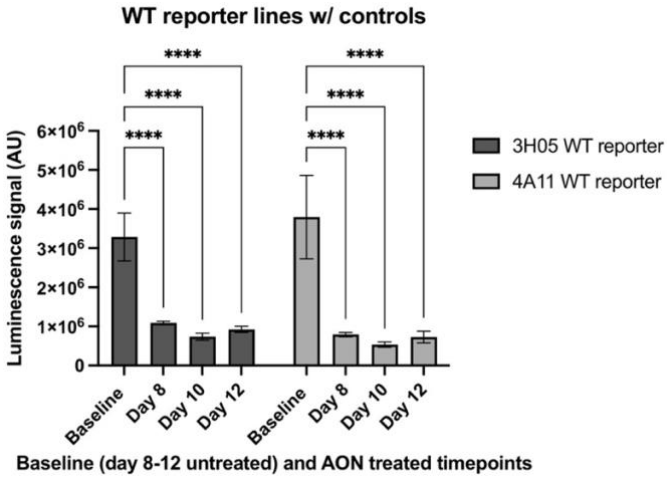

Figure 4-49. *Luminesce levels of wild-type reporter clones treated with 3 doses of 1500 nM peptide-conjugated PMO AON.*

*Differentiating Myocea myoblast cells are incubated with AON on days 2, 4 and 6 of differentiation for 48 hours each. Luminescence levels measured on days 8, 10 and 12 of differentiation using plate reader. The results are for 3 technical repeats. ANOVA results show that there is a significant difference of P <0.0001 due to treatment; they also confirm that there is no significant difference between technical repeats of every condition. Tukey's multiple comparison test was used to determine pairwise significance and **** = P <0.0001. Error bars indicate standard deviation.*

Subsequently, I trialled PPMO in restoring the reading frame in dystrophic reporter lines. Five of the exon 52 deleted reporter lines were used and the same concentration of AON, 1500 nM, was applied to day 11 differentiated exon 52 deleted reporter myotubes for 48 hours. All five exon 52 deleted lines results show that there is a statistically significant increase in luminescence levels in treated vs untreated cultures of approximately 60%, in luminescence levels in treated compared to untreated conditions, in **Figure 4-50**, indicating the efficacy the PPMO AON and the robustness of the reporter system.

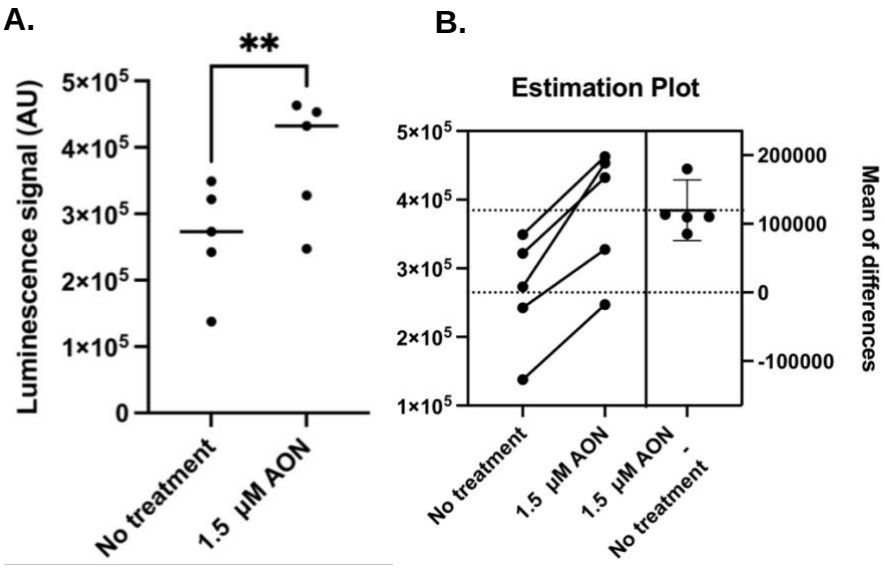

Figure 4-50. *Luminesce levels of exon 52 deleted clones treated with 1500 nM peptide-conjugated PMO AON.*

Differentiating myocea myoblast cells are incubated with AON on day 11 of differentiation for 48 hours. Luminescence levels measured on day 13 of differentiation using plate reader. The results are for 5 biological repeats (the 5 exon 52 deleted clones) and each of the measures plotted are of 3 technical repeats. ANOVA results show that there is a significant difference of P 0.0017

due to treatment; they also confirm that there is no significant difference between technical repeats of every condition.

Finally, to assess whether multiple doses of the treatment would lead to a higher protein production of dystrophin, two of the dystrophin exon 52 deleted reporter lines were used. Cultures were treated with one dose or two doses of 1500 nM of peptide-AON (each treatment is 48 hours). The results, in **Figure 4-51**, show that there is no significant difference between luminescence expression levels (and hence dystrophin protein production) in one dose compared to two dose treated cultures, however, the dystrophin levels in NCRM1-L1O-A15-3H05-A4 are around half that of its isogenic control NCRM1-L1O-A15-3H05. They also show that treatment leads to a highly significant increase in luminescence levels compared to untreated cultures. The luminescence levels obtained after treatment are more than 5-fold the level of untreated cultures, approaching 30%-50% of the levels measured in the isogenic wild type reporter lines.

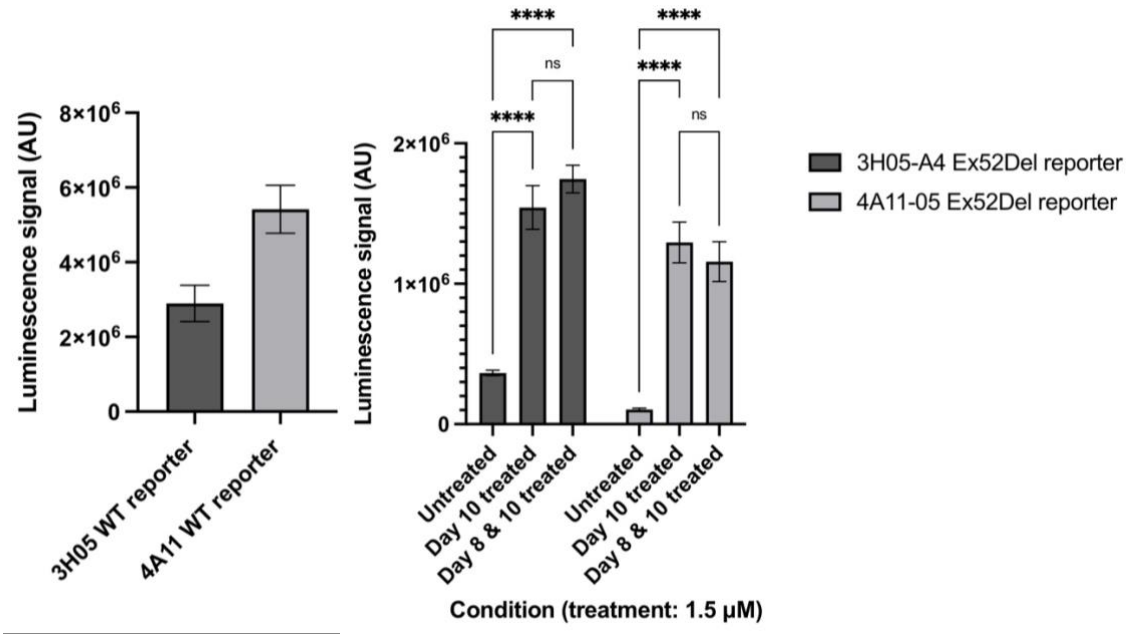

Figure 4-51. *Luminesce levels of exon 52 deleted reporter clones treated with up to 2 doses of 1500 nM peptide-conjugated PMO AONs.*
A. *Luminescence levels of wild-type reporter isogenic controls without treatment.*
B. *Differentiating myocea myoblast cells are incubated with AON on day 8 or day 8 and 10 of differentiation for 48 hours each. Luminescence levels measured on days 12 of differentiation using plate reader. The results are for 3 technical repeats. ANOVA results show that there is a significant difference of P <0.0001 due to treatment; they also confirm that there is no significant difference between technical repeats of every condition. Tukey's multiple comparison test was used to determine pairwise significance and **** = P <0.0001. Error bars indicate standard deviation.*

Taken together, the results presented in section 4.4 demonstrate the potential and the highly sensitive quantitative luminescence function of the reporter in detecting dystrophin protein production in cultures and upon administration of AONs inducing exon skipping.

4.5 ADDITIONAL LINES OF INVESTIGATION:

4.5.1 ALTERNATIVE EXPANSION MEDIA FOR IPS CELL-DERIVED MYOGENIC CELLS

In the middle of this project, the Genea Skeletal Muscle Differentiation kit and its media were discontinued by the producing company. In efforts to replace the expansion media, saving time by allowing us to use already available iPS cell-derived myoblasts in our lab; rather than having to use and optimise another published transgene-free protocol from the literature. We trialled growing the Genea stage 2 myoblasts in Megacell Culture media. Megacell is the media used in our standard transgene-based expansion of human iPS cell-derived inducible myogenic cells (Maffioletti *et al.*, 2015). Other myogenic expansion media exist and, in a lab-wide effort have been tried out by other lab members; however, in this section I am only reporting results I have produced. One cell line was used for this experiment, NCRM1 Genea myoblasts, grown for 10 days in Megacell compared to NCRM1 Genea myoblasts grown in Genea Myoblast media. The results in **Figure 4-52** show a higher differentiation index for Megacell-grown myogenic progenitors compared to Myoblast media-grown myogenic progenitors. Interestingly they show that Myoblast grown myogenic progenitors have a slightly higher fusion index.

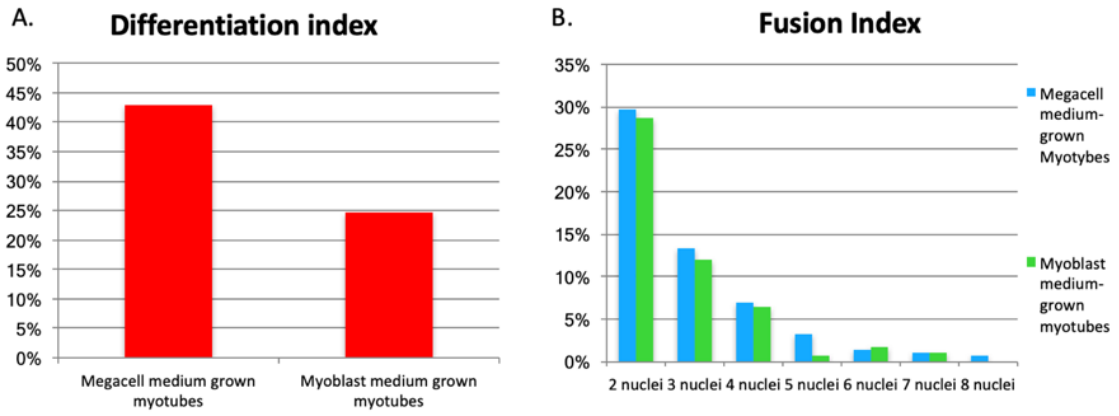

Figure 4-52. *Differentiation and fusion Indices for differentiated Genea myoblasts expanded in different media prior to differentiation.*

Indices calculated for Immunofluorescent stained terminally differentiated myotubes at day 8 of differentiation ***A.*** *Differentiation index (percentage of nuclei in MyHC+ cells) for Megacell media-grown Myoblasts and Genea Myoblast media-grown Myoblasts.* ***B.*** *Fusion index (percentage of nuclei in multinucleated myotubes) divided into 7 groups) for Megacell media-grown Myoblasts and Genea Myoblast media-grown Myoblasts. There are three technical repeats, each with over 1000 nuclei counted, for every value represented on this graph. These experiments were done in collaboration with, Mengmeng Zhu, MRes student under my supervision. Results are of 2 biological repeats, each of 3 technical repeats.*

4.5.2 Setting Up Functional Readouts for 3D Muscle Constructs

4.5.2.1 Calcium Transients in 3D Muscle Constructs

After confirming the advantages and potential of 3D skeletal muscle constructs over monolayer cultures, in section 4.3 and in efforts to set up a system in which calcium transients are used as functional readouts along with live imaging of dystrophin in 3D muscle constructs, a Zeiss confocal microscope

was used to image calcium transients of HIDEM 3D muscle constructs, removed from the posts and rested on the bottom of a tissue culture dish. The results plotted show that there are responding cells and regions in the 3D muscle constructs in **Figure 4-53**. However, the traces also show that the response is not as clear and instantaneous as that of monolayer cultures.

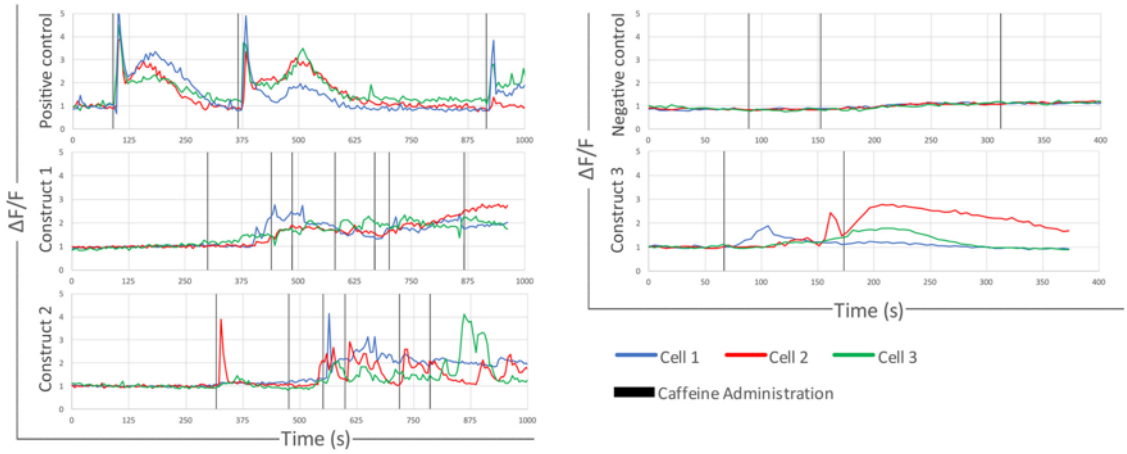

Figure 4-53. *Calcium transients traces in HIDEM 3D skeletal muscle constructs upon caffeine administration.*

3D muscle constructs were loaded with Fluo-4AM (fluorescent Ca2+ indicator) upon caffeine administration. Three representative responding cells per sample are plotted. Caffeine administration (10 mM) is indicated with vertical back lines. Positive control: monolayer of differentiated HIDEM myotubes; negative control: undifferentiated cells from the same line used in the constructs). ΔF/F: measured fluorescence relative to the average resting fluorescence of each point before the administration of caffeine. Image analysis was conducted with ImageJ and Microsoft Excel was used to generate the graphs. Figure adapted from (Maffioletti et al., 2018). Results are of 3 technical repeats.

This experiment is a collaborative effort in which Dr. Shilpita Sarcar made the 3D muscle constructs and I conducted the analysis and made the figure with the help of the Duchen lab.

4.5.3 Transduction of 3D Muscle Constructs with Adeno Associated Viral vectors

To assess the AAV vectors' transduction efficiency of the skeletal muscle constructs, and thus confirming the possibility of using them for screening AAV-delivered therapies a preliminary screening using 2 doses (HD= 100,000 MOI, LD=1,000 MOI as reported in 3.9 was carried out. The results show that up to day 5 post-infection, the GFP expression levels cannot be detected by the naked eye under the stereo microscope as seen in **Figure 4-54 A**., however, in the HD replicates, some GFP signal was picked starting day 6, its clearer in the LD samples that the detection of expression begins later, however, it is very difficult to detect under the stereo microscope, it is better visualised when the muscle constructs are rested on the bottom of a plate and observed under an immunofluorescence microscope as seen in **Figure 4-54 A**. The results show progressive increase in GFP expression, till saturation in both the high dose, and a slow increase low dose, suggesting that the platform can be used for dose determining experiments. **Figure 4-54 B** shows the co-localisation of GFP signal along with MF20. **Figure 4-54 C** shows that the fold increase in GFP mRNA quantified using qRT-PCR in high dose over low dose is more than 100X and this

is also reflected in the AAV copy number per unit of DNA extracted in **Figure 4-54 D**.

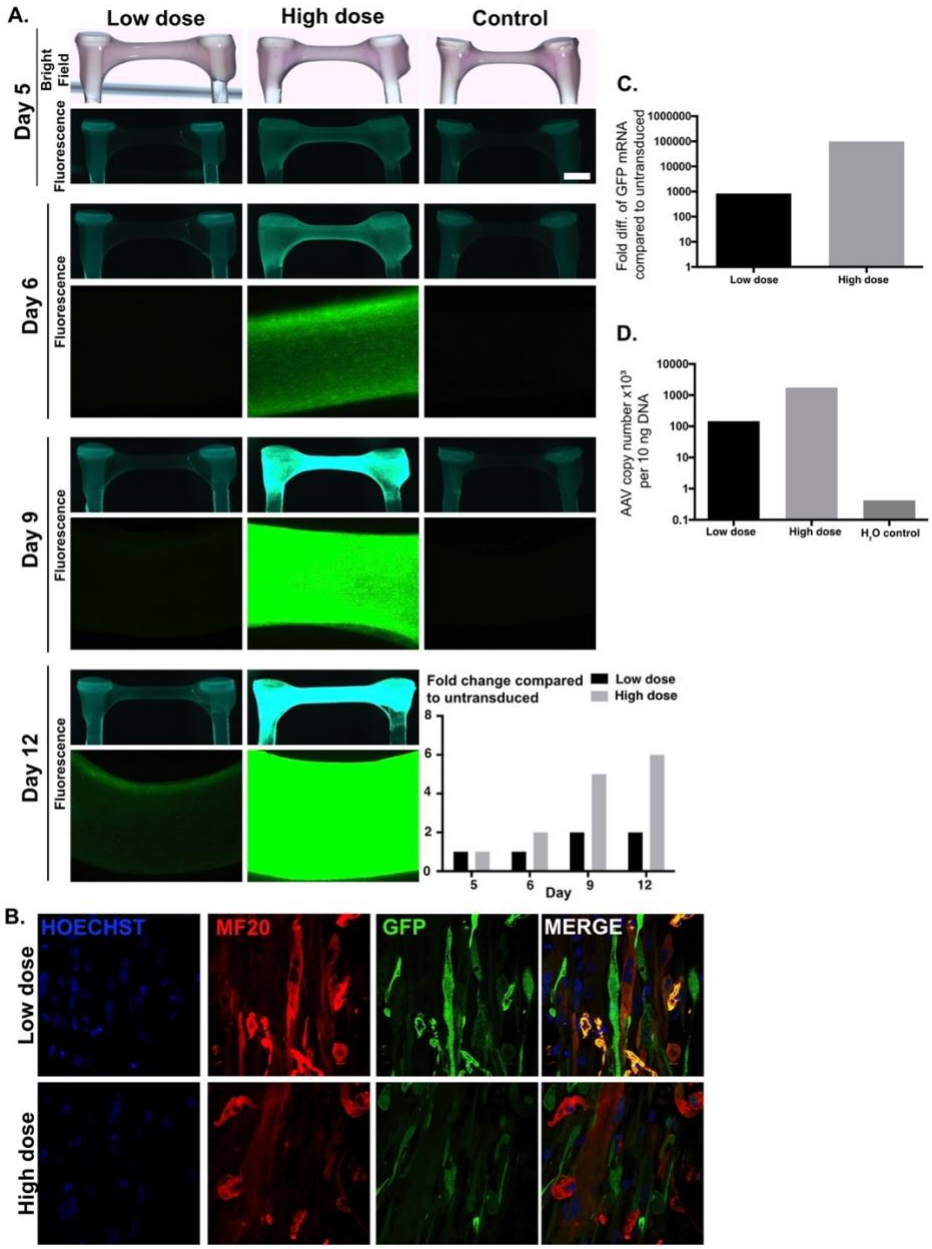

Figure 4-54. *Single lineage 3D muscle constructs transduced with high (HD) and low (LD) dose MOI of AAV9 containing a GFP plasmid under a CMV promoter.*

A) *Stereo microscope imaging of 3D muscle constructs showing brightfield image on initial day of imaging and GFP fluorescence of all imaging time points. The bottom panels on days 6, 9 and 12 show an inverted microscope fluorescence images of muscle constructs resting on the bottom of tissue-culture dishes. Bottom right graph shows fold change in intensity levels compared to un-transduced muscle constructs.* ***B)*** *Maximum intensity projection images of fixed and immunofluorescent stained muscle constructs after experiment termination. (Images are 1 representative of 7 technical repeats). HD is 100,000 MOI and LD is 1,000 MOI.* ***C)*** *qRT-PCR analysis GFP. Expression (dCt) is normalised against GAPDH, and relative expression (ddCt) is normalised against untransduced control. Results presented as fold change.* ***D)*** *qRT-PCR analysis for AAV genome. Results are presented as genome copy number per total amount of DNA extracted in ng. Conception and initial experimental design with Prof. F.S. Tedesco and Dr S. Sarcar. Muscle generation and live imaging with Dr. S. Sarcar.*

As a proof of concept that the platform can be used for assessing AAV tropism. AAVs 2, 5, 8 and 9 were used to transduce cells within bilineage artificial skeletal muscle organoid containing motor neurons and skeletal muscle cells. The AAVs contain a plasmid with Green Fluorescence Protein (GFP) under the drive of a CMV promoter and the titer used is 50,000 MOI. A minimum of four muscle constructs were used per condition for a total of 23 constructs (this includes uninfected controls). The 3D muscle constructs were transduced on day 5 of differentiation and imaged on a stereo and inverted microscopes at multiple timepoints; **Figure 4-55A** outlines the experimental procedure followed. Live stereo and inverted microscope images show some GFP+ cells can be observed in AAV2, 8 and 9-infected constructs from day 3. From day 6, levels of GFP expression increase progressively over time up to day 14 in **Figure 4-55B** and

quantified in **Figure 4-55C.** The maximum projection images obtained by confocally imaging samples 14 days post-transduction, show the overlay of GFP signal and Titin and/or SMI32 which indicates that the AAVs transduced myogenic cells and motor neurons within the constructs as seem in **Figure 4-55D**, the figure also shows the ratio of myogenic cells and motor neurons in the 3D artificial muscle constructs (quantified in **Figure 4-55E**). All the AAV serotypes infected the myotubes with AAV 2, 8 and 9 capable of transducing up to 17% of motor neurons as well as seen in **Figure 4-55F.**

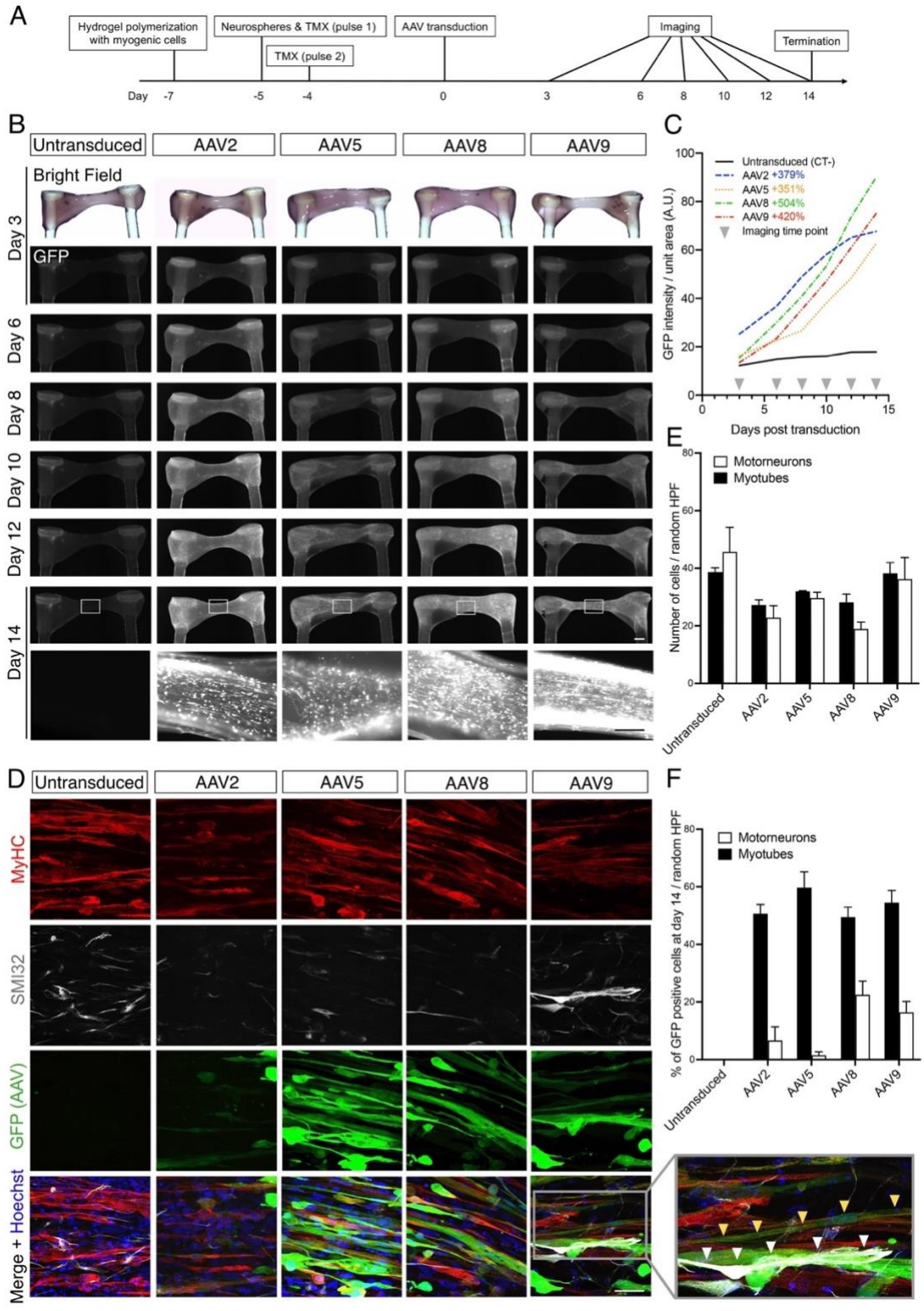

Figure 4-55. *Bilineage 3D muscle constructs transduced with various AAV serotypes at a dose of 50,000 MOI.*

A) *Schematic diagram summarising the timeline of the experiment including key steps; making the 3D muscle constructs gel using HIDEMs on day -7, addition of the neurospheres and beginning of HIDEMs and neurospheres' terminal differentiation using tamoxifen and neurosphere bi-lineage differentiation media on day -5, administration of a second tamoxifen pulse on day -4, transduction with AAVs on day 0, brightfield and fluorescence imaging using stereo microscope of 3D muscle constructs and inverted microscope fluorescence imaging, and finally termination on day 14 of the experiment after which fixing and immunofluorescence staining is carried out.* ***B)*** *Stereo microscope and inverted microscope imaging of 3D muscle constructs un-transduced and transduced with AAVs 2, 5, 8 and 9 at different time points.* ***C)*** *Graph showing change in fluorescence intensity in stereo microscope imaged 3D muscle over time.* ***D)*** *Maximum intensity projection images of fixed 3D muscle constructs on day 14 and immunofluorescent stained with muscle marker (MyHC), neural marker (SMI32), GFP and Hoechst for nuclei. Blow up image at bottom right with yellow arrows highlighting GFP+ MyHC+ fibre, and white arrows highlighting GFP+ SMI32+ neural cells.* ***E)*** *Graph showing absolute numbers of myotubes and motorneurons per random high-powered field (HPF) of 25X magnification maximum intensity projection of immunofluorescent stained fixed 3D muscle constructs on day 14 (n=4 for control and AAV2; n= 5 for AAV5, AAV8 and AA5 9).* ***F)*** *Graph showing percentage of GFP+ myotubes and motorneurons per random high-powered field (HPF) of 60X magnification maximum intensity projection of immunofluorescent stained fixed 3D muscle constructs on day 14 (n=10 for control and AAV 8; n=11 for AAV2; n=12 for AAV5 and AAV9). Transduction of multilineage artificial muscles with multiple AAV serotypes is a collaborative project conceived and led by Prof. F.S. Tedesco. I conducted the experimental design, live imaging, data analysis and presentation of the experiment. Luca Pinton conducted confocal imaging and contributed to data analysis. Elisa Negroni made the 3D muscle construct and carried out AAV transduction.*

Taken together, the results reported in section 4.5.2 demonstrate the use of the more physiologically relevant 3D skeletal muscle constructs for more advanced applications. This includes a functional outcome: calcium transients and the capability of the platform to screen for dosage and AAV tropism in bilineage constructs, laying the foundations for more advanced investigations using the platform.

5 DISCUSSION

5.1 GENERATION OF DYSTROPHIN REPORTER HUMAN IPS CELL LINE USING CRISPR/CAS9

5.1.1 BACKGROUND

In this study human iPS cells and their derivatives were used to develop a faithful *in vitro* dystrophin-reporting platform to model Duchenne muscular dystrophy. The rationale behind starting with iPS cells is their theoretically unlimited expansion potential, ability to derive all cell types of the adult form, and the lack of ethical constraints, discussed in detail in 1.4.2.2.1.

However, the main critical aspect pertaining to iPS cell-derived disease models is the maturation and aging of the obtained differentiated cells and the degree of faithfulness as models for human diseases, especially late-onset diseases (e.g., Parkinson's Disease) and diseases of progressive nature such as some muscular dystrophies, including DMD. Therefore, I aim to extend the use of these reporter cells into a *quasi vivo* environment, known as 3D skeletal muscle constructs, as discussed in subsequent sections.

In order to generate an effective dystrophin reporter, a multiple reporter cassette that includes a NanoLuciferase luminescence sequence, an mCherry fluorescence sequence and a HaloTag interchangeable labelling tag was designed and inserted before the dystrophin stop codon at the end of exon 79.

The resulting cell lines would enable real-time visualisation and localisation of dystrophin using the mCherry reporter, quantitative detection of dystrophin in live and fixed cells using the NanoLuminescence reporter, and allows interaction studies and more accurate detection by western blotting using the HaloTag reporter.

To date, two other groups have utilised similar approach, albeit using only a luminescence or fluorescence reporter, inserted in the same locus, at the end of exon 79 just before the stop codon, in a mouse model, (Amoasii *et al.*, 2019) and (Petkova *et al.*, 2016), respectively. Despite the clear relevance of this approach, using mouse models to model DMD have increasingly been less favourable due to their limitations in accurately replicating human disease progression (Hay *et al.*, 2014). In addition to the fact that generating gene-edited animal models is a laborious, expensive and ethically questionable procedure. Studies on disease patho-mechanisms, development and therapy require tens of replicates, and potentially isogenic wild type and mutant pairs, making the use of animal models increasingly unfeasible in light of the current regulations of animal research. In contrast, generating gene edited reporter iPS cell lines, their upkeep and use is straight forward, relatively inexpensive and ethically widely accepted.

Amoasii and colleagues generated wild type reporter and its isogenic dystrophic mutant reporter mouse model with exon 50 deletion and demonstrated the viability of using the bioluminescence reporter in live monitoring of restoration of dystrophin protein production in affected skeletal muscles upon administration of AAV-mediated CRISPR/Cas9 gene editing to induce exon 51 skipping. The

mouse model enables live non-invasive monitoring of dystrophin expression restoration *in vivo*, which is a very powerful tool in assessing therapy penetration and distribution of restored dystrophin expression at an organism level. However, the system does not allow for quantitative readouts at the cellular level given the three dimensional nature of the model and the limitations of imaging systems. This is in contrast to our human iPS cell reporter lines which provide quantitative readouts, in monolayer cultures, and even in 3D muscle constructs, enabling more comprehensive and higher throughput experiments and research.

Petkova and colleagues generated a wild type dystrophin reporter with an EGFP fluorescence tag and demonstrate the advantages of using a tag at this locus in robustly detecting dystrophin using immunofluorescence staining in fixed skeletal muscle, brain and retina samples. They also demonstrate reliable detection of dystrophin in live isolated muscle fibres *ex vivo*. This model provides clear advantages over common mouse models such as the *mdx* mouse, provided an isogenic dystrophic mice is generated.

A major advantage of using both mouse models is the facility of investigating dystrophin expression and therapy development in various affected tissues, not just skeletal muscle cells, and in a more physiologically relevant environment than *in vitro* organoids and models. Despite the lack of such complexity in iPS cell derived models, recent advancements in such systems, including mono layer cultures, 3D cultures and organ-on-chip systems, are rendering them increasingly relevant and more faithful.

It is worth noting that, apart from the aforementioned advantages of using human iPS cell-derived cells for modelling human disease, both mouse models include only one of two of the reporter components in our system, providing less versatility outcomes and have limited potential use compared to the system reported here.

5.1.2 CRISPR/Cas9 Gene Editing Using Triple-electroporation Method

In the initial trials using the triple electroporation method to edit five cell lines, a very low editing efficiency was obtained, and in the edited clones, many had off target inserts, which disqualified them from being used in further experiments. The low efficiency can be attributed to multiple causes including the fact that dystrophin is not robustly expressed in iPS cells. This is affected by the epigenetic status of the cell and causing the chromatin to be tightly wound in target locus, hindering the Cas9 nuclease from accessing the target region (Jensen *et al.*, 2017; Uusi-Mäkelä *et al.*, 2018; Verkuijl and Rots, 2019). Horlbeck and colleagues demonstrated that CRISPR/Cas9 efficiency varies drastically in accordance with the accessibility of the target locus, and that chemicals altering chromatin conformation can alter gene editing efficiency (Horlbeck *et al.*, 2016). however, this is unlikely to be the case in our experiments given the high cutting efficiency obtained in the subsequent optimisation experiments.

Another possibility is that, since our strategy relies on Cas9 being transcribed and translated within the cell as it is delivered as a plasmid, it is

difficult to predict whether the expression of Cas9 protein will be synchronous with the sustained presence of the sgRNAs and the DMD template, or whether they would be degraded by the cell at this stage. It has been shown that a higher gene-editing efficiency is obtained when sgRNAs are delivered into cells that have been nucleofected with the Cas9 expressing plasmid 24 hours earlier, compared to the standard approach, a collective single delivery (Kouranova *et al.*, 2016). This, is coupled with the fact that a plasmid expressing Cas9 protein in the cell increases the chances of off-target inserts and effects (Kouranova *et al.*, 2016), makes it a less than ideal approach.

Finally, we noticed that identical linker sequences upstream and downstream of the HaloTag sequence in this reporter were used, which could lead to spontaneous rearrangements, including deletions, inversions and excisions. This might explain multiple shorter-than-expected bands in the BSD probe southern blot. Although this issue caused delays in this critical step of the project, this opportunity was used to produce a backup, shorter version of the cassette, without HaloTag and mCherry sequences.

5.1.3 CRISPR/Cas9 Gene Editing Using RNP Complex Method

To address these challenges, I started optimising the CRISPR/Cas9 gene editing procedure, first by adopting a more recently developed approach in which the sgRNA is pre-incubated with the Cas9 protein to form an RNP complex (along with a linearised template for HDR) and using nucleofection rather than

electroporation. The results confirm that the new optimised approach is more efficient than the old one.

It is worth noting that the term “nucleofection” has been coined by the manufacturer of the Amaxa system, but it is essentially a form of electroporation in which proprietary electroporation protocols and buffers are used. Nucleofection has been demonstrated to have better efficiency in human embryonic and adult stem cells (Lakshmipathy *et al.*, 2004; Hohenstein *et al.*, 2008; Fong *et al.*, 2011). The results confirm that while electroporation leads to a higher number of clone outgrowth, nucleofection still gave rise to one successfully edited clone, similar to electroporation, from a lower number of isolated clones, demonstrating higher efficiency.

Following these optimisation steps, I carried out the RNP based CRISPR/Cas9 experiment on the NCRM1 iPS cells, a well-established and well-studied wild type control line. As expected, the conditions which were nucleofected with the shorter version of the cassette had more clonal outgrowth, which is likely due to reduced DNA toxicity compared to the full-length cassette (Ribeiro *et al.*, 2012). Conditions where the shorter cassette template was used also gave rise to a higher number of edited clones, confirmed through PCR, further suggesting that HDR with a shorter template is more likely to occur than with a long template. This finding is consistent with what has been reported in the literature, Li and colleagues demonstrated that in addition to the effect of homology arms in a template on the efficiency of HDR, the length of the inserted sequence has an inverse correlation with efficiency (K. Li *et al.*, 2014).

Initially, my aim was to use wild type control and dystrophic human iPS cell lines to generate reporter lines. However, I sought out an alternative strategy after the initial gene editing trials in 4.1.1 proved the difficulty in obtaining edited clones and the laborious steps needed for the new (RNP) approach in 4.1.3. Therefore, a different approach was adopted, generating dystrophin reporter clones from one control line and then inducing a mutation (deleting exon 52) to generate its isogenic dystrophic equivalent, which is a more robust strategy as the disease modelling and therapy developments fields have been increasingly adopting isogenic studies as standard for monogenic disease modelling to reduce line to line (or subject to subject) variation and obtain more reliable data.

Pairs of genetically isogenic cell lines, which differ by a single genetic variation, are considered powerful tools to study gene function in different contexts, and specifically for monoallelic diseases (DeWeirdt *et al.*, 2020). Gene editing to produce isogenic pairs is traditionally laborious, however, advances in gene editing technologies, specifically CRISPR/Cas9 mediated technologies, has significantly reduced the barriers to adopting the isogenic approach in research. Multiple research efforts have employed this strategy, detailed in 1.5.4. The impact of background genetic variation has been highlighted by Xu and colleagues, in which they employed principal component analysis of microarray gene expression data to show that corrected and diseased (Huntington's disease) genetically isogenic iPS cell-derived GABAergic neurons clustered together and far from wild type control cells (Xu *et al.*, 2017). Multiple groups have employed this approach using human cells for DMD disease modelling. Sarka

and colleagues demonstrated that dystrophic iPS cell-derived cardiac cells display mechanical impairment, increased heart rate variability, bradycardia and blunted β-adrenergic response compared to their genetically isogenic control cells (Jelinkova *et al.*, 2020). Soblechero-Martín and colleagues used wild type immortalised human myoblasts to generate exon 52 deleted isogenic cells and used them in assessing utrophin overexpression after drug treatment in differentiated myotubes (Soblechero-Martín *et al.*, 2021). Al Tanoury and colleagues used wild type iPS cell line NCRM1 to generate DMD exon 52 deleted isogenic control, then demonstrated that prednisolone treatment on myogenic cultures dramatically rescues force contraction, fusion and branching defects in DMD iPS cell-derived cultures compared to isogenic wild type control, this challenges the prevalent view that the drug's effects are caused by its anti-inflammatory properties, but rather through its direct action on myofibres (Al Tanoury *et al.*, 2021).

The choice of using an exon 52 deletion mutation for this study is due to two reasons. First, it is one of the most commonly deleted exons in DMD patients and occurs in the mutation "hotspot" of dystrophin (Bladen *et al.*, 2015). Secondly, it is amenable to both exon 51 and exon 53 skipping using antisense oligonucleotide therapies (such as Eteplirsen and Golodirsen, both of which are FDA approved) which can theoretically treat up to 21% of patients combined (Aartsma-Rus *et al.*, 2009).

In conclusion, the results in this section confirm the successful generation of dystrophin reporter human iPS cells harbouring NanoLuciferase, mCherry and

HaloTag sequences inserted in frame at the target locus by genome editing. To our knowledge it is the first time that such a complex multifunctional dystrophin reporting tool is successfully developed in human iPS cells, which offers unprecedented measures that enable more in-depth studies on the role of dystrophin and its isoforms developmentally and more insight on disease and therapy patho-mechanisms. The reporter lines with DMD exon 52 deletion, isogenic to the control reporter lines, will enable study of dystrophin restoration via mutation specific therapeutics and enable more accurate research outcomes. This lays the foundations for the creation of further isogenic dystrophic lines from our wild type reporter lines which will enable mutation comparing studies in a genetically isogenic setting.

5.2 Reporter Function in human iPS Cell Derived Skeletal Muscle Cells

Human iPS cells offer an unlimited source of patient-specific cells that can be differentiated into almost any cell type which is crucial for studying most genetic disorders, especially neuromuscular diseases. Efficient derivation of these cell types, and primarily cells of the myogenic lineage, is consequently crucial for *in vitro* and *in vivo* studies into disease biology and therapy screening.

Multiple strategies have been developed to derive myogenic, and subsequently fully differentiated skeletal muscle fibres, including transgene-based and transgene free strategies from human iPS cells (Jalal, Dastidar and Tedesco, 2021). I utilised the transgene-based method developed in our lab to derive skeletal muscle lines used as controls and for optimisation experiments in this research (Maffioletti *et al.*, 2015). More recent advancements have led to small molecule-based transgene-free approaches that generate expandable myogenic cells from human iPS cells (Awaya *et al.*, 2012; Borchin, Chen and Barberi, 2013; Caron *et al.*, 2016; Chal *et al.*, 2016; Xi *et al.*, 2017; Mazaleyrat *et al.*, 2020).

All results pertaining to DMD reporter iPS cell-derived lines presented in this thesis use the generated dystrophin reporter lines have been derived using the Genea/Myocea skeletal muscle differentiation kit based on the original publication of Caron and colleagues (Caron *et al.*, 2016). This protocol was chosen for its ease of use as it does not require any fluorescence activated cell sorting steps. In addition to that, the fact that it is commercially available, reduces

time and effort needed to formulate the media; and, more importantly, large-scale manufacturing entails multiple quality control steps and stringent manufacturing practices that reduce batch to batch variation, which is in contrast to home-made media, especially ones with multiple components and complex formulations.

The dystrophin gene gives rise to various isoforms in different tissues and during development. The results also show a consistent increase in expression of dystrophin in reporter myoblasts and differentiating myotubes over time. While iPS cell-derived myoblasts do not express Dp427 isoforms of DMD, the basal NanoLuciferase signal detected in myoblasts is likely to be of the Dp71 isoform. Dp71 is expressed ubiquitously in human cells including myogenic progenitor cells (Muntoni, Torelli and Ferlini, 2003; Tadayoni *et al.*, 2012; Chesshyre *et al.*, 2022). In addition to that, the isoform DP71ab has been found to be exclusively expressed in myoblasts (Farea *et al.*, 2020). More recently, it has been discovered that it enhances proliferation of human myoblast cells and confirmed that their proliferation is positively affected by Dp71ab in a dose-dependent manner (Farea *et al.*, 2022).

Recently Massouridès and colleagues have identified an embryonic dystrophin isoform, DP427e, which is expressed in hES and human iPS cells upon BMP4 treatment, a critical signalling molecule involved in multiple key cellular pathways (Massouridès *et al.*, 2015). Our reporter lines can be a powerful tool to study dystrophin isoforms, especially novel isoforms such as this one, utilising the HaloTag sequence in protein-protein interaction studies or in protein isolation using pulldown or immunoprecipitation and mass spectrometery.

The sensitivity of the nanoluciferase feature of the reporter is demonstrated in the results that show that exon 52 deleted clones exhibited a significantly higher expression of NanoLuciferase than wild type cells with no reporter but much lower than differentiated wild type reporter clone. Exon 52 deleted mutants can produce the ubiquitously expressed Dp71 as this isoform does not include exon 52 due to its promoter which lies in exon 61 (Muntoni, Torelli and Ferlini, 2003). This isoform was previously thought not to be expressed in human myogenic cells, based on *in vitro* studies, but further research proved it is expressed in myogenic and differentiated skeletal muscle cells (Tadayoni *et al.*, 2012). Collectively these results prove the reliability of the NanoLuciferase reporter in producing sensitive quantifiable data of dystrophin protein production in cells in a monolayer culture.

The mCherry function of the cassette has been validated and the results show robust detection and localisation of dystrophin using immunofluorescence. The live reporting function of mCherry, which is the main feature of using mCherry, has proven to be difficult to detect, as mCherry is commonly known to be a weak live reporter. However, optimisation steps have resulted in robust detection of it, and localisation and distribution can be visualised in live cells, which offers unprecedented potential in DMD studies in human cells *in vitro*.

The use of the HaloTag, and mCherry, feature of the cassette in western blotting shows the reliability of the system in detecting dystrophin isoforms, in contrast to the need to use multiple/an array of antibodies as commonly used and discussed earlier. Interestingly the HaloTag antibody shows a very crisp band at

Dp116 and a band at Dp40, which are expressed in the Schwann cells (Byers, Lidov and Kunkel, 1993), and in general human foetal tissues and schwannoma cells (Tinsley, Blake and Davies, 1993). Of notability, the *dystrophin* antibody was not very efficient in detecting the dystrophin bands and took over 13 minutes of exposure on the chemiluminescence machine to be visualised, in comparison to the mCherry and HaloTag antibody that took between 1-2 minutes of exposure.

One potential cause for observing dystrophin bands corresponding to Dp116 and Dp40 is the general problem of cell population purity often encountered with iPS cell derivation protocols. It is widely accepted that the biggest drawback of using small molecule-based differentiation protocols compared to transgene-based protocols in that they often result in less efficient and more variable cultures (Chal *et al.*, 2016). These protocols often generate heterogenous cultures of myogenic cells at different stages of commitment as well as non-myogenic populations (Kim *et al.*, 2017). Ectodermal cell populations are often the main contaminants of iPS cell-derived myogenic protocols as they share many developmental cues with the mesodermal lineage (Gouti *et al.*, 2014; Henrique *et al.*, 2015). For further confirmation of cell population purity, testing for the presence of Schwann cell contaminants using specific such as S100, p75NTR, Sox10, Sox2, GAP43, NCAM, Krox20, Oct6, MBP, and MPZ might be considered in future work (Liu *et al.*, 2015).

In conclusion, the results in this section confirm the functions of the three components of the cassette, NanoLuciferase, mCherry, and HaloTag; and demonstrate the clear advantages of using the system over conventional

methods on monolayer skeletal muscle cultures derived from various sources. Further investigations into the co-localisation of the mCherry signal with endogenous production of dystrophin protein using immunofluorescence staining is necessary, and the investigation of the bands obtained using western blotting for the HaloTag. In addition to that, identifying the expressed isoforms at the RNA level using qPCR or RNA seq would be desirable to further confirm and correlate to the results obtained from the reporter's NanoLuciferase feature.

Overall, the results demonstrate the quantitative sensitivity of the NanoLuciferase reporter, the potential of mCherry fluorescence in providing real time dystrophin tracking and the efficiency of the HaloTag at capturing dystrophin isoforms.

The use of these dystrophin reporter human iPS cells is not limited to skeletal muscles. Multiple tissue types and organs have been demonstrated to be affected by the lack of dystrophin in DMD patients (Duan *et al.*, 2021). Outstanding work includes using the reporter cell lines in deriving various tissues and validating its functionality in them.

One of the most important tissues implicated DMD, other than muscles, is the central nervous as patients with DMD often have cognitive impairment, depending on the type and location of mutation (Duan *et al.*, 2021). The Dp427, Dp140 and Dp71 isoforms of dystrophin- all of which would be conjugated to the reporter upon translation in the generated dystrophin reporting lines, are expressed in the central nervous system and have been proven to affect cognitive outcomes in DMD patients (Chesshyre *et al.*, 2022). Recent developments in

central nervous system cell derivation from iPS cells and especially in 3D cultures, such as the self-organising cerebral organoids developed by Lancaster and colleagues (Lancaster *et al.*, 2013), would allow for in-depth studies of dystrophin dynamics in development and disease.

5.3 HUMAN IPS CELL-DERIVED SKELETAL MYOTUBES IN 2D AND 3D CULTURES

A key application of the dystrophin reporter lines described in the previous two sections is their combination with dedicated scaffolds and biomaterials to establish a more physiologically relevant, 3D *quasi vivo* environment to model DMD. Combining these two platforms can deliver a next generation human dystrophin reporting muscle for personalised and precision medicine.

To further characterise our previously published 3D muscle platform (Maffioletti *et al.*, 2018) and to optimise imaging modalities for possible use to detect reported fluorescence, I performed light sheet microscopy. Light sheet microscopy enables the rapid imaging of thicker 3D samples and from multiple angles due to its stage design that enables the suspension of the sample. The software, combined with the motorised stage, allows for acquisition of precise spatial positioning data, used to reconstruct acquired image stacks into a 3D render.

One of the main advantages of the 3D muscle constructs and the use of fibrinogen in this constrained system is the uniaxial tension and directional cues signalling cell alignment along the axis of force (Jalal, Dastidar and Tedesco, 2021). This oriented myogenic differentiation is a critical feature of muscle tissue engineering. It has been demonstrated that culturing cells on substrates and surfaces that have physiological rigidity enhances muscle stem cell renewal (Gilbert *et al.*, 2010); it also greatly enhances myogenic differentiation and myotube maturation (Engler *et al.*, 2006; Jiwlawat *et al.*, 2019).

The acquired image stack does not go beyond a few myotubes in depth, around 100 μm. This might be explained by the fact the artificial constructs are quite dense and compact, with extracellular components contributing to opacity (June Hoan Kim *et al.*, 2018). In addition to the fact that skeletal muscle tissue is inherently opaque (Williams *et al.*, 2019). However, these hurdles are a result of limitations in imaging and microscopy technologies, which various groups are working on overcoming. In addition, methods for tissue clearing to enable deeper imaging penetration in skeletal muscle and artificial skeletal muscle constructs are constantly being developed (Decroix *et al.*, 2015; Williams *et al.*, 2019).

iPS cell-derivatives tend to exhibit a foetal-like identity which hinders generation of disease models for late onset disorders (Studer, Vera and Cornacchia, 2015), especially when differentiated in a monolayer (Xi *et al.*, 2020). This applies to muscular dystrophies, including DMD, considering their progressive nature. One relevant approach to overcome this is to provide a more *in vivo* niche-like environment for organoids with the structural, mechanical and electrical cues that more closely mimic the physiological environment and has been extensively studied in cardiac iPS cell-derived cultures (Cho *et al.*, 2021). This can be achieved through differentiating in a 3D environment, which the 3D muscle construct reported here is amongst the first systems developed for iPS cell-derived skeletal muscles (Maffioletti *et al.*, 2018).

To validate whether 3D cultures provide structural and mechanical cues that can lead to more differentiated and mature skeletal muscle fibres, dystrophin

and myosin expression patterns were investigated and compared between 3D and monolayer cultures.

Firstly, dystrophin is used as it is expressed during the later stages of myogenesis. Myosins are differentially expressed in the myogenic lineage during development, conferring different functional and metabolic properties to the muscle fibres (Schiaffino and Reggiani, 2011). In summary, myosins are grouped into embryonic, neonatal and adult isoform groups; their expression starts at the corresponding developmental age and adult myosins are the main types present in most adult muscle tissues (Schiaffino *et al.*, 2015). Data reported by Brown and colleagues shows that, at the mRNA level, developmental myosins are expressed and then downregulated in favour of the adult myosins, particularly in myogenic cells (Brown, Parr and Brameld, 2012).

Dystrophin expression levels were assessed in 3D muscle constructs and monolayer cultures over time between days 8 and 12 of terminal differentiation using qRT-PCR analysis. In summary, the results show consistent expression of the neonatal myosins in monolayer culture and little presence of the adult myosins. However, cells differentiated in 3D muscle constructs show increased levels of expression of the adult *MYH* isoforms, indicating that differentiation of myogenic cells in 3D gives rise to more mature skeletal muscle cells, consistent with what has been reported in the literature.

In conclusion, I have shown that human iPS cell-derived 3D muscle constructs display physiologically relevant morphology and demonstrated their maturation potential over monolayer cultures. This platform could bridge the gap

between human *in vitro* cell culture research and animal models, which often offer very poor prediction of the clinical outcomes in humans. In combination with dystrophin reporter lines, this innovative platform would offer an unprecedented opportunity to study disease and therapy mechanisms. The possibility of having multiple cell types incorporated into the 3D artificial skeletal muscles (isogenic at the cellular level) offers a novel chance to study disease modelling in other cell types in the tissue (apart from myotubes, which are the most studied cell type in muscular dystrophies) facilitating research into disorder-related phenotypes and disease mechanisms in other cell types. One such example would be endothelial cell defects, which have been identified in DMD cells (Miike *et al.*, 1987; Dabiré *et al.*, 2012; Palladino *et al.*, 2013). Studying them in this platform, co-cultured with myotubes, in a *quasi vivo* setting would provide unprecedented insight.

5.4 DYSTROPHIN REPORTER IPS CELLS FOR DMD THERAPY TESTING

Another key application of the dystrophin reporting iPS cells lines described in this thesis is their use in therapy development, given that they may enable qualitative and quantitative assessment of dystrophin restoration in specific mutation targetable with new genetic therapies. As a proof of principle AONs were used on monolayer cultures, utilising the reporter cassette's NanoLuciferase luminescence as a quantitative readout of dystrophin restoration (with a view that future work might further extend this strategy to obtaining readouts also from the other components of the reporter in both 2D and 3D cultures).

I first tested a PMO AON to skip exon 53, utilising the sequence of Golodirsen, to assess the potential of the reporter system in quantifying exon skipping *in vitro*. Having proven in principle the possibility of using the cassette in quantitively measuring exon skipping with PMO AONs, the next step was trialling next generation AON therapies: peptide conjugated PMO (PPMO). In the following experiments a PMO skipping exon 53 that is conjugated to a novel Pip protein developed by the Wood laboratory in Oxford University (in collaboration with Dr Y. Lomonosova and Prof M. Wood, Oxford University) was utilised.

Previous reports in the literature have demonstrated restoration of dystrophin expression in human iPS cell-derived skeletal muscle cells. Soblechero-Matrins and colleagues show the restoration of dystrophin expression using 2'MOE-phosphorotioate AON to skip *DMD* exon 51 and

delivered using lipofectamine. However, the percentage/rate of dystrophin restoration was not evaluated; the levels seem to be very minimal (around 10%) and was evaluated at the RNA level (Soblechero-Martín *et al.*, 2021). Shoji and colleagues demonstrated the restoration of dystrophin expression using a 2'MOE-phosphorotioate AON to skip *DMD* exon 45, showing skipping at both the RNA and protein level (by means of western blotting), although quantification of skipping was not calculated (Shoji *et al.*, 2015). While these studies are sufficient as proof-of-concept reports, the lack of a quantifiable measure of efficacy is a barrier to translation into pre-clinical and clinical trials. Assessing the skipping levels at the RNA and protein level is possible through amplicon sequencing and protein purification methods. These methods are laborious and inherently inefficient measures as they rely on the indirect quantification of RNA and protein isolates. The reporter iPS cell-derived system described here provides a reliable, easy to use platform to investigate these points.

In this chapter I demonstrated the quantitative feature of NanoLuciferase luminescence in measuring efficacy of skipping *in vitro.* The results show increase dose and temporal response to treatment and during differentiation in real time live experiments. The tested treatment regimen does not lead to full restoration of luminescence levels similar to that of the native gene in the isogenic reporter clones (as expected with PMO and PPMO AONs and in keeping with animal work and clinical trials (Popplewell *et al.*, 2009; Echigoya *et al.*, 2017; Lim *et al.*, 2019; Sheng *et al.*, 2020)). This could be due to efficiency in penetrating the cell membranes; however, the results also suggest otherwise, as PPMO

mediated exon skipping in wild type reporter clones significantly decreased luminescence levels, almost reaching levels expressed by exon 52 deleted reporter clones.

The experiments presented thus far employ a concentration range of 500 nM to 3 μM of AONs. The rationale behind the choice of concentrations relies in the previous studies employing AONs for DMD exon skipping (Aartsma-Rus *et al.*, 2004) and the assessment of cytotoxicity observed in cultures in relation to concentration, which prompted us to increase concentrations up to 3 μM. These concentration ranges are clinically relevant and acceptable ranges. For context, the maximal serum concentration levels (C_{max}) reported by the FDA after one hour of intravenous infusion ranged between 80 nm to 4.9 μM for Eteplirsen (https://www.accessdata.fda.gov/drugsatfda_docs/nda/2016/206488Orig1s000ClinPharmR.pdf), and 770 nm to 5 μM for Golodirsen (https://www.accessdata.fda.gov/drugsatfda_docs/nda/2019/211970Orig1s000ClinPharmR.pdf). These figures are for serum levels, their maximum levels are double the concentrations applied in the culture media reported in this thesis; however, the rationale is after crossing the endothelial layer *in vivo*, concentrations would drop to figures similar to the ones utilised throughout this thesis.

Potential experimental inefficiencies that are worth investigating include the delivery method utilised for PMO delivery. Endo-porter causes stress to the cells as discussed above and employing a liposome-based approach such as lipofectamine might circumvent stressing the cells and allow more channelling of

the PMO into the cytoplasm. Another issue is the potential erroneous handling of AONs, as these molecules are highly stable in various temperatures but are light sensitive.

Overall, the aforementioned findings demonstrate the potential use of the system in guiding research through providing more readouts, they suggest that high concentrations of AON (in the range of 1.5 μM and 3 μM) are optimal to proceed with further experiments, investigating administration of multiple dosages, different peptide chemistries and different AON targets.

Although the exon 52 deletion mutation used here is one of the most common and that antisense oligonucleotides used in skipping exon 53 is only second to exon 51 in the percentage of DMD patients that can benefit from, the AON skipping experiments presented in this thesis represent just up to 8% of the patient population. However, in the wider context of therapies for DMD, exon skipping strategies in general can target up to 80% of the patient population, readthrough technologies are applicable to 15% and CRISPR/Cas9 editing technologies are theoretically capable of correcting all mutations. The reporter lines described in this thesis can be utilised in screening for these approaches, through the creation of isogenic dystrophic lines similar to the approach reported here in deriving the exon 52 deleted lines.

Taken together, these results lay the foundation for use of the reporter cassette iPS cell line derivatives in screening AON-mediated exon skipping therapies for DMD, and in a wider context, all therapy approaches that target the

endogenous dystrophin sequence, including read through technologies and CRISPR/Cas9 technologies.

5.5 ADDITIONAL LINES OF INVESTIGATION

5.5.1 ALTERNATIVE EXPANSION MEDIA FOR IPS CELL-DERIVED MYOGENIC CELLS

As part of a lab wide effort, we investigated the potential of using various media for the expansion of myogenic progenitor cells we have derived using the Genea skeletal muscle kit, based on the Caron et al publication, which was discontinued by the company GENEA Biocells (*Skeletal muscle differentiation kit, SKM-KIT* | *AMSBIO*). I focused on trialling the use of the commercially available Megacell media. Our results showed that stage 2 myogenic cells were expandable in Megacell media, albeit with a reduced growth rate.

Interestingly, upon differentiation for 8 days using the standard protocol, the differentiation index for Megacell-grown myoblasts is higher than Myoblast media-grown myoblasts. That might be attributed to the fact that the Genea media is a very growth factor rich media, with growth factors such as EGF, HGF, IGF and FGF which are known to help maintain myogenic progenitor cells (Chargé and Rudnicki, 2004). In contrast, Megacell media's low concentrations of growth factors primes the cells for differentiation. Long-term expansion potential has not yet been investigated yet. The results shown, confirm that Megacell media is a viable alternative for growth and differentiation of Genea iPS cell-derived myoblasts.

5.5.2 SETTING UP FUNCTIONAL READOUTS FOR 3D MUSCLE CONSTRUCTS

Aiming to eventually combine both the reporter lines and the 3D artificial skeletal muscle platforms, I sought out to investigate potential functional readouts for 3D muscle constructs, starting with calcium transients. Reporter lines were not used for these experiments as both lines of investigation were done in parallel. The system was set up using human iPS cell-derived HIDEM based 3D muscle constructs and a confocal microscope Zeiss LSM880 with temperature and gas control.

Mature myotubes have a transient increase of intracellular calcium concentrations (known as calcium transients) upon chemical and/or electrical stimulation, which results in contractions (Berchtold, Brinkmeier and Müntener, 2000). This process, present in both types of striated muscles, cardiac and skeletal, is known as excitation-contraction coupling (Caillé, Ildefonse and Rougier, 1985). Caffeine is a pharmacological compound commonly used as a stimulant for this process in *in vitro* studies and is often used as surrogate evidence of the possession of the cells of an efficient excitation contraction coupling apparatus which is a feature of mature muscle cells.

Addition of caffeine at a 10mM concentration was used to stimulate sarcoplasmic reticulum calcium ions release via ryanodine receptors (Rousseau *et al.*, 1988; Thomas and Williams, 2012). The results show that muscles often only contract after the addition of multiple doses of caffeine and in some of the cells, a delayed response was also observed. This is possibly due to the extracellular matrix, and in particular fibrin covering the myotubes and delaying the uptake of caffeine by the cells. The removal of the 3D muscle constructs from

the posts and onto the bottom of the plate, removing the tension, might explain the presence of some non-responding cells as well as the responders.

The artificial skeletal muscle platform provides a tool to screen next generation therapies for DMD. One of these promising therapies is AAV delivered gene-editing tools (such as CRISPR/Cas9) to correct the dystrophin sequence. Using dystrophin-reporting cell lines will allow us to visualise in live cell culture (using mCherry fluorescence) and quantify (using NanoLuciferase luminescence) the restoration of dystrophin upon therapy administration and differentiation of isoforms using western blotting.

5.5.3 Transduction of 3D Muscle Constructs with Adeno Associated Viral Vectors

AAVs are particularly promising delivery vehicles for various treatment approaches for DMD, especially strategies to use gene editing CRISPR/Cas9 systems to correct and edit the endogenous mutations. Although many viral and non-viral delivery vehicles are being researched, AAVs are currently the forefront given that they have particularly low immunogenicity, a wide range of tropism, smaller size compared to other viral vectors, ability to sustain episomal transgene expression for a long time and that they can be delivered systemically (Gregorevic *et al.*, 2004; Howe *et al.*, 2008). AAVs are the leading vectors for gene delivery in humans and strategies using AAVs are currently under clinical evaluation for the treatment of a variety of genetic diseases, including degenerative neuromuscular disorders (Aguti, Malerba and Zhou, 2018). It has

been shown that AAV transduction in striated skeletal muscles can last for years (Rivera *et al.*, 2005) and AAV of serotypes 6, 8 and 9 have been demonstrated to successfully transduce cardiac and skeletal muscle *in vivo* in the *mdx* mouse model by several groups (Long *et al.*, 2016; Nelson *et al.*, 2016; Tabebordbar *et al.*, 2016; Bengtsson *et al.*, 2017). However, researchers have demonstrated the poor tropism of AAVs towards satellite cells (Arnett *et al.*, 2014; Tabebordbar *et al.*, 2016). This limits the extent of mutation correction efficiency and therefore the therapeutic potential of AAV mediated strategies relying on gene editing. Ideally, AAVs targeting both differentiated skeletal muscles cells and satellite cells would be optimal to obtain a higher efficiency therapy. Our 3D skeletal muscle model system could provide a reliable model for such research. Therefore, I sought out to investigate the potential of using iPS cell-derived artificial skeletal muscle constructs for screening AAVs as a proof of concept.

The first experiment, in which UCLi007 HIDEM-derived 3D muscle constructs were transduced with two different doses (1000 and 100,000 MOI) of AAV delivering a GFP plasmid under the drive of a CMV promoter, confirms that a dose-dependent response can be observed, measured through the intensity of the GFP fluorescence that is encoded by the plasmid delivered by the AAV. The temporal effect was also observed, increasing GFP intensity over time, from day 6 to day 12. The signal is confirmed to originate from myogenic cells through co-staining the transduced 3D muscle construct with a GFP antibody and MF20, in which signal overlapping.

The second experiment, bilineage 3D muscle constructs (derived from iPS cell-derived HIDEMs and neurospheres) were transduced with AAV2, AAV5, AAV8 and AAV9 which have different tropisms to skeletal muscle and spinal motor neuron cells (Zincarelli *et al.*, 2008; Muraine *et al.*, 2020; Wasala, Chen and Duan, 2020). Starting from day 3, some GFP+ cells were clearly observed in AAV2, and at a lower level in AAV8 and AAV9 transduced constructs. From day 6, levels of GFP expression in all constructs can be seen and progressively increase over time until day 14. Confocal imaging confirms that all AAVs transduced both myogenic cells and motor neurons in the construct. The ratios of the transduction of both cell types are quantified and consistent with previous studies (Zincarelli *et al.*, 2008; Muraine *et al.*, 2020; Wasala, Chen and Duan, 2020), all AAV serotypes transduce myotubes to various degrees and AAVs 2, 8 and 9 transduced up to 17% of motor neurons.

The results show that all AAV serotypes used have affinity to skeletal muscle cells, with the quantification of fluorescence trendline showing AAV 2 reaching saturation and starting to plateau by day 10 of imaging, while the other serotypes show sustained, almost exponential, expression of GFP. This is consistent with various studies confirming that AAV 2 has the lowest tropism to skeletal muscle cells (Zincarelli *et al.*, 2008; Muraine *et al.*, 2020). Interestingly, on day 3, the first day of imaging, the fluorescence levels observed for AAV 2 were higher than that of all other AAVs. This might be due to transduction dynamics of various AAVs and their tropism to motor neurons. Further in-depth studies using this 3D skeletal muscle platform would enable such research.

Taken together, the results detail potential outcome measures- imaging and calcium transients in 3D cultures, prove the relevance of 3D skeletal muscle constructs over monolayer cultures, and demonstrate the potential of using 3D muscle constructs in studying therapies' mechanisms in a more physiologically relevant *quasi vivo* setting.

6 CONCLUSION AND FUTURE PERSPECTIVES

In this thesis, I report the derivation of dystrophin reporter iPS cell lines, and their isogenic DMD exon 52 deleted pairs, that include NanoLuciferase luminescence, mCherry fluorescence and HaloTag interchangeable labelling protein reporters. I also detail the confirmation of function of all reporter cassette components, providing robust, quantitative and live outcomes. This work represents the first example of human cells that include a dystrophin reporter, and among the first examples of human iPS cell isogenic pairs of DMD and wild type control lines.

The significance of this work lies in the versatility of the reporter and the potential of iPS cells. Together, they enable studies that are isogenic at two levels, the genetic level, offering pairs of cells that have the same genetic background apart from the DMD mutation, and the cellular level, as iPS cells can be derived into different cell types. This is of particular relevance to studying DMD during development, DMD disease patho-mechanisms and the involvement of different cell types.

Some example applications include using the reporter iPS cell lines to generate muscle precursors using various protocols *in vitro*, identifying subpopulations such as satellite cells, which would enable studies into disease during development. Including other iPS cell-derived types in such a model would enable studies in disease patho-mechanisms which would be especially relevant if utilised in a more physiologically relevant environment, 3D muscle constructs.

Using isogenic dystrophic and wild type reporter iPS cells differentiated into different lineages and using them in co-culture applications, deriving skeletal muscle and neural cells for example, would provide better insight in tissue involvement in disease pathology research. Moreover, the developed reporter lines enable research into the often overlooked tissues such as the brain using advancements in iPS cell derived organoid systems and protocols.

The results of screening AON mediated exon skipping using the reporter lines, provide a proof of concept for the quantitative feature of the cassette and lay the foundation for investigating AON chemistry and conjugate screening and assessment of exon skipping efficiency. This is becoming increasingly relevant for the field as quantitative, real time readouts are necessary for further development and studies into AON-mediated exon skipping and other therapies' efficiency prior to animal and clinical trials, reducing the number of animals needed and increasing the chances of success of clinical trials.

In parallel, I investigated the potential of using 3D artificial muscle constructs as a platform to model DMD, demonstrating its relevance over monolayer cultures as it provides structural and mechanical cues that lead to better iPS cell-derived skeletal muscle maturation. Moreover, I report imaging strategies and functional readouts, such as calcium transients for 3D muscle constructs. In addition to providing a more physiologically relevant environment for disease modelling, using the isogenic reporter lines in a 3D muscle organoid enables functional quantitative studies such as contractions and calcium dynamics in response to electrical stimulation. This is particularly useful for

studies investigating rate and efficiency of restoration of function after therapy- with the facility of quantitively and qualitatively tracking dystrophin in the reporter cell lines.

In conclusion, this advanced platform, combining dystrophin reporting human iPS cell lines and 3D muscle constructs, will enable investigations into DMD disease mechanisms as well as studying therapies such as AONs or AAV mediated CRISPR/Cas9 gene editing in a more biologically and physiologically relevant *quasi vivo* environment with real time readouts.

7 REFERENCES

Aartsma-Rus, A. *et al.* (2004) 'Comparative analysis of antisense oligonucleotide analogs for targeted DMD exon 46 skipping in muscle cells', *Gene Therapy 2004 11:18*. Nature Publishing Group, 11(18), pp. 1391–1398. doi: 10.1038/sj.gt.3302313.

Aartsma-Rus, A. *et al.* (2009) 'Theoretic applicability of antisense-mediated exon skipping for Duchenne muscular dystrophy mutations', *Human Mutation*, 30(3), pp. 293–299. doi: 10.1002/humu.20918.

Abujarour, R. *et al.* (2014) 'Myogenic Differentiation of Muscular Dystrophy-Specific Induced Pluripotent Stem Cells for Use in Drug Discovery', *STEM CELLS Translational Medicine*, 3(2), pp. 149–160. doi: 10.5966/sctm.2013-0095.

Adli, M. (2018) 'The CRISPR tool kit for genome editing and beyond', *Nature Communications 2018 9:1*. Nature Publishing Group, 9(1), pp. 1–13. doi: 10.1038/s41467-018-04252-2.

Aguti, S., Malerba, A. and Zhou, H. (2018) 'The progress of AAV-mediated gene therapy in neuromuscular disorders', *Expert opinion on biological therapy*. Expert Opin Biol Ther, 18(6), pp. 681–693. doi: 10.1080/14712598.2018.1479739.

Allen, D. G., Whitehead, N. P. and Froehner, S. C. (2016) 'Absence of Dystrophin Disrupts Skeletal Muscle Signaling: Roles of Ca^{2+}, Reactive Oxygen Species, and Nitric Oxide in the Development of Muscular Dystrophy', *Physiological Reviews*. American Physiological Society Bethesda, MD, 96(1), pp. 253–305. doi: 10.1152/physrev.00007.2015.

Amenta, A. R. *et al.* (2011) 'Biglycan recruits utrophin to the sarcolemma and counters dystrophic pathology in mdx mice', *Proceedings of the National Academy of Sciences of the United States of America*, 108(2), pp. 762–767. doi: 10.1073/pnas.1013067108.

Amoasii, L. *et al.* (2018) 'Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy', *Science*, 362(6410), pp. 86–91. doi: 10.1126/science.aau1549.

Amoasii, L. *et al.* (2019) 'In vivo non-invasive monitoring of dystrophin correction in a new Duchenne muscular dystrophy reporter mouse', *Nature Communications*, 10(1). doi: 10.1038/s41467-019-12335-x.

Arend, M. C., Pereira, J. O. and Markoski, M. M. (2017) 'The CRISPR/Cas9 System and the Possibility of Genomic Edition for Cardiology', *Arquivos brasileiros de cardiologia*. Arq Bras Cardiol, 108(1), pp. 81–83. doi: 10.5935/ABC.20160200.

Arnett, A. L. *et al.* (2014) 'Adeno-associated viral vectors do not efficiently target muscle satellite cells', *Molecular Therapy - Methods & Clinical Development*, 1, p. 14038. doi: 10.1038/mtm.2014.38.

Arnold, L. *et al.* (2007) 'Inflammatory monocytes recruited after skeletal muscle injury switch into antiinflammatory macrophages to support myogenesis', *The Journal of experimental medicine*. J Exp Med, 204(5), pp. 1057–1069. doi: 10.1084/JEM.20070075.

Aung-Htut, M. T. *et al.* (2019) 'In Vitro Validation of Phosphorodiamidate Morpholino Oligomers', *Molecules*. Multidisciplinary Digital Publishing Institute (MDPI), 24(16). doi: 10.3390/MOLECULES24162922.

Awaya, T. *et al.* (2012) 'Selective Development of Myogenic Mesenchymal Cells from Human Embryonic and Induced Pluripotent Stem Cells', *PLoS ONE*. Edited by A. Asakura, 7(12), p. e51638. doi: 10.1371/journal.pone.0051638.

Bakooshli, M. A. *et al.* (2019) 'A 3d culture model of innervated human skeletal muscle enables studies of the adult neuromuscular junction', *eLife*. eLife Sciences Publications Ltd, 8. doi: 10.7554/eLife.44530.

Bar, S. *et al.* (1990) 'A novel product of the Duchenne muscular dystrophy gene which greatly differs from the known isoforms in its structure and tissue distribution', *The Biochemical journal*. Biochem J, 272(2), pp. 557–560. doi: 10.1042/BJ2720557.

Barrangou, R. *et al.* (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science (New York, N.Y.)*. Science, 315(5819), pp. 1709–1712. doi: 10.1126/SCIENCE.1138140.

Belfort, M. and Roberts, R. J. (1997) 'Homing endonucleases: keeping the house in order.', *Nucleic Acids Research*. Oxford University Press, 25(17), p. 3379. doi: 10.1093/NAR/25.17.3379.

Bellin, M. *et al.* (2012) 'Induced pluripotent stem cells: the new patient?', *Nature Reviews Molecular Cell Biology*, 13(11), pp. 713–726. doi: 10.1038/nrm3448.

Bellmann, J. *et al.* (2019) 'A customizable microfluidic platform for medium-throughput modeling of neuromuscular circuits', *Biomaterials*, 225. doi: 10.1016/j.biomaterials.2019.119537.

Benedetti, S., Hoshiya, H. and Tedesco, F. S. (2013) 'Repair or replace? Exploiting novel gene and cell therapy strategies for muscular dystrophies', *FEBS Journal*. John Wiley & Sons, Ltd (10.1111), 280(17), pp. 4263–4280. doi: 10.1111/febs.12178.

Bengtsson, N. E. *et al.* (2017) 'Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy', *Nature Communications 2017 8:1*. Nature Publishing Group, 8(1), pp. 1–10. doi: 10.1038/ncomms14454.

Berardo, A., DiMauro, S. and Hirano, M. (2010) 'A diagnostic algorithm for metabolic myopathies', *Current neurology and neuroscience reports*. Curr Neurol Neurosci Rep, 10(2), pp. 118–126. doi: 10.1007/S11910-010-0096-4.

Berchtold, M. W., Brinkmeier, H. and Müntener, M. (2000) 'Calcium ion in

skeletal muscle: its crucial role for muscle function, plasticity, and disease', *Physiological reviews*. Physiol Rev, 80(3), pp. 1215–1265. doi: 10.1152/PHYSREV.2000.80.3.1215.

Betts, J. G. *et al.* (2017) 'Anatomy & Physiology', in *Anatomy & Physiology*. Texas: OpenStax- Rice University, p. 408. Available at: https://openstax.org/details/books/anatomy-and-physiology.

Beumer, K. J. *et al.* (2008) 'Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases', *Proceedings of the National Academy of Sciences of the United States of America*, 105(50), pp. 19821–19826. doi: 10.1073/PNAS.0810475105/ASSET/8B233FD1-E4AF-46C0-AEBB-8F5A3A88DAF8/ASSETS/GRAPHIC/ZPQ9990858450004.JPEG.

Birling, M. C. *et al.* (2017) 'Efficient and rapid generation of large genomic variants in rats and mice using CRISMERE', *Scientific Reports 2017 7:1*. Nature Publishing Group, 7(1), pp. 1–11. doi: 10.1038/srep43331.

Bladen, C. L. *et al.* (2015) 'The TREAT-NMD DMD global database: Analysis of more than 7,000 duchenne muscular dystrophy mutations', *Human Mutation*. John Wiley and Sons Inc., 36(4), pp. 395–402. doi: 10.1002/HUMU.22758.

Blau, H. M., Webster, C. and Pavlath, G. K. (1983) 'Defective myoblasts identified in Duchenne muscular dystrophy.', *Proceedings of the National Academy of Sciences*, 80(15), pp. 4856–4860. doi: 10.1073/pnas.80.15.4856.

Boch, J. *et al.* (2009) 'Breaking the code of DNA binding specificity of TAL-type III effectors', *Science (New York, N.Y.)*. Science, 326(5959), pp. 1509–1512. doi: 10.1126/SCIENCE.1178811.

Bolotin, A. *et al.* (2005) 'Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin', *Microbiology (Reading, England)*. Microbiology (Reading), 151(Pt 8), pp. 2551–2561. doi: 10.1099/MIC.0.28048-0.

Borchin, B., Chen, J. and Barberi, T. (2013) 'Derivation and FACS-Mediated Purification of PAX3+/PAX7+ Skeletal Muscle Precursors from Human Pluripotent Stem Cells', *Stem Cell Reports*, 1(6), pp. 620–631. doi: 10.1016/j.stemcr.2013.10.007.

Boroviak, K. *et al.* (2016) 'Chromosome engineering in zygotes with CRISPR/Cas9', *Genesis*. John Wiley and Sons Inc., 54(2), pp. 78–85. doi: 10.1002/DVG.22915.

Bowles, D. E. *et al.* (2012) 'Phase 1 Gene Therapy for Duchenne Muscular Dystrophy Using a Translational Optimized AAV Vector', *Molecular Therapy*, 20(2), pp. 443–455. doi: 10.1038/mt.2011.237.

Brassard, J. A. and Lutolf, M. P. (2019) 'Engineering Stem Cell Self-organization to Build Better Organoids', *Cell Stem Cell*. Cell Press, pp. 860–876. doi: 10.1016/j.stem.2019.05.005.

Brouns, S. J. J. *et al.* (2008) ‘Small CRISPR RNAs guide antiviral defense in prokaryotes’, *Science (New York, N.Y.)*. Science, 321(5891), pp. 960–964. doi: 10.1126/SCIENCE.1159689.

Broutier, L. *et al.* (2016) ‘Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation’, *Nature Protocols*. Nature Publishing Group, 11(9), pp. 1724–1743. doi: 10.1038/nprot.2016.097.

Brown, D. M., Parr, T. and Brameld, J. M. (2012) ‘Myosin heavy chain mRNA isoforms are expressed in two distinct cohorts during C2C12 myogenesis’, *Journal of Muscle Research and Cell Motility*. Springer, 32(6), pp. 383–390. doi: 10.1007/S10974-011-9267-4/FIGURES/3.

Buckingham, M. *et al.* (2003) ‘The formation of skeletal muscle: from somite to limb’, *Journal of Anatomy*, 202(1), pp. 59–68. doi: 10.1046/j.1469-7580.2003.00139.x.

Buckingham, M. (2006) ‘Myogenic progenitor cells and skeletal myogenesis in vertebrates’, *Current opinion in genetics & development*. Curr Opin Genet Dev, 16(5), pp. 525–532. doi: 10.1016/J.GDE.2006.08.008.

Buckingham, M. and Relaix, F. (2015) ‘PAX3 and PAX7 as upstream regulators of myogenesis’, *Seminars in cell & developmental biology*. Semin Cell Dev Biol, 44, pp. 115–125. doi: 10.1016/J.SEMCDB.2015.09.017.

Burch, P. M. *et al.* (2015) ‘Muscle-Derived Proteins as Serum Biomarkers for Monitoring Disease Progression in Three Forms of Muscular Dystrophy’, *Journal of Neuromuscular Diseases*. IOS Press, 2(3), pp. 241–255. doi: 10.3233/JND-140066.

Bushby, K. *et al.* (2010) ‘Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management’, *The Lancet Neurology*, 9(1), pp. 77–93. doi: 10.1016/S1474-4422(09)70271-6.

Bushby, K. *et al.* (2014) ‘Ataluren treatment of patients with nonsense mutation dystrophinopathy’, *Muscle & Nerve*, 50(4), pp. 477–487. doi: 10.1002/mus.24332.

Byers, T. J., Lidov, H. G. W. and Kunkel, L. M. (1993) ‘An alternative dystrophin transcript specific to peripheral nerve’, *Nature genetics*. Nat Genet, 4(1), pp. 77–81. doi: 10.1038/NG0593-77.

Caillé, J., Ildefonse, M. and Rougier, O. (1985) ‘Excitation-contraction coupling in skeletal muscle’, *Progress in Biophysics and Molecular Biology*. Pergamon, 46(3), pp. 185–239. doi: 10.1016/0079-6107(85)90009-4.

Capel, A. J. *et al.* (2019) ‘Scalable 3D Printed Molds for Human Tissue Engineered Skeletal Muscle’, *Frontiers in Bioengineering and Biotechnology*. Frontiers, 7, p. 20. doi: 10.3389/fbioe.2019.00020.

CAPERS, C. R. (1960) ‘Multinucleation of skeletal muscle in vitro.’, *The*

Journal of biophysical and biochemical cytology, 7(3), pp. 559–66. doi: 10.1083/jcb.7.3.559.

Cardarelli, F. *et al.* (2016) 'The intracellular trafficking mechanism of Lipofectamine-based transfection reagents and its implication for gene delivery', *Scientific Reports 2016 6:1*. Nature Publishing Group, 6(1), pp. 1–8. doi: 10.1038/srep25879.

Caron, L. *et al.* (2016) 'A Human Pluripotent Stem Cell Model of Facioscapulohumeral Muscular Dystrophy-Affected Skeletal Muscles', *STEM CELLS Translational Medicine*, 5(9), pp. 1145–1161. doi: 10.5966/sctm.2015-0224.

Carroll, K. J. *et al.* (2016) 'A mouse model for adult cardiac-specific gene deletion with CRISPR/Cas9', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 113(2), pp. 338–343. doi: 10.1073/PNAS.1523918113.

Caudal, D. *et al.* (2020) 'Characterization of brain dystrophins absence and impact in dystrophin-deficient Dmdmdx rat model', *PLoS ONE*. Public Library of Science, 15(3). doi: 10.1371/journal.pone.0230083.

Cencic, R. *et al.* (2014) 'Protospacer Adjacent Motif (PAM)-Distal Sequences Engage CRISPR Cas9 DNA Target Cleavage', *PLOS ONE*. Public Library of Science, 9(10), p. e109213. doi: 10.1371/JOURNAL.PONE.0109213.

Cermak, T. *et al.* (2011) 'Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting', *Nucleic acids research*. Nucleic Acids Res, 39(12). doi: 10.1093/NAR/GKR218.

Chal, J. *et al.* (2015) 'Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy', *Nature Biotechnology*. Nature Publishing Group, 33(9), pp. 962–969. doi: 10.1038/nbt.3297.

Chal, J. *et al.* (2016) 'Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro', *Nature Protocols*, 11(10), pp. 1833–1850. doi: 10.1038/nprot.2016.110.

Chamberlain, J. S. *et al.* (2007) 'Dystrophin-deficient *mdx* mice display a reduced life span and are susceptible to spontaneous rhabdomyosarcoma', *The FASEB Journal*, 21(9), pp. 2195–2204. doi: 10.1096/fj.06-7353com.

Chargé, S. B. P. and Rudnicki, M. A. (2004) 'Cellular and Molecular Regulation of Muscle Regeneration', *Physiological Reviews*, 84(1), pp. 209–238. doi: 10.1152/physrev.00019.2003.

Chau, M. H. K. *et al.* (2021) 'Trio-Based Low-Pass Genome Sequencing Reveals Characteristics and Significance of Rare Copy Number Variants in Prenatal Diagnosis', *Frontiers in Genetics*. Frontiers Media S.A., 12, p. 1735. doi: 10.3389/fgene.2021.742325.

Chazaud, B. *et al.* (2003) 'Satellite cells attract monocytes and use macrophages as a support to escape apoptosis and enhance muscle growth',

The Journal of cell biology. J Cell Biol, 163(5), pp. 1133–1143. doi: 10.1083/JCB.200212046.

Chemello, F., Bassel-Duby, R. and Olson, E. N. (2020) 'Correction of muscular dystrophies by CRISPR gene editing', *The Journal of clinical investigation*. J Clin Invest, 130(6), pp. 2766–2776. doi: 10.1172/JCI136873.

Chen, X., Zaro, J. L. and Shen, W. C. (2013) 'Fusion Protein Linkers: Property, Design and Functionality', *Advanced drug delivery reviews*. NIH Public Access, 65(10), p. 1357. doi: 10.1016/J.ADDR.2012.09.039.

Chesshyre, M. *et al.* (2022) 'Investigating the role of dystrophin isoform deficiency in motor function in Duchenne muscular dystrophy', *Journal of Cachexia, Sarcopenia and Muscle*. Springer Nature, 13(2), pp. 1360–1372. doi: 10.1002/JCSM.12914.

Chevalier, B. S. *et al.* (2002) 'Design, Activity, and Structure of a Highly Specific Artificial Endonuclease', *Molecular Cell*. Elsevier, 10(4), pp. 895–905. doi: 10.1016/S1097-2765(02)00690-1.

Chevalier, B. S. and Stoddard, B. L. (2001) 'Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility', *Nucleic Acids Research*. Oxford University Press, 29(18), p. 3757. doi: 10.1093/NAR/29.18.3757.

Cho, S. *et al.* (2021) 'Reconstructing the heart using iPSCs: Engineering strategies and applications', *Journal of Molecular and Cellular Cardiology*. Academic Press, 157, pp. 56–65. doi: 10.1016/J.YJMCC.2021.04.006.

Choi, I. Y. *et al.* (2016) 'Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model', *Cell Reports*, 15(10), pp. 2301–2312. doi: 10.1016/j.celrep.2016.05.016.

Christ, B. and Ordahl, C. P. (1995) 'Early stages of chick somite development', *Anatomy and embryology*. Anat Embryol (Berl), 191(5), pp. 381–396. doi: 10.1007/BF00304424.

Chylinski, K., Le Rhun, A. and Charpentier, E. (2013) 'The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems', *RNA biology*. RNA Biol, 10(5), pp. 726–737. doi: 10.4161/RNA.24321.

Ciafaloni, E. *et al.* (2009) 'Delayed diagnosis in duchenne muscular dystrophy: data from the Muscular Dystrophy Surveillance, Tracking, and Research Network (MD STARnet)', *The Journal of pediatrics*. J Pediatr, 155(3), pp. 380–385. doi: 10.1016/J.JPEDS.2009.02.007.

Cirak, S. *et al.* (2011) 'Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study', *The Lancet*, 378(9791), pp. 595–605. doi: 10.1016/S0140-6736(11)60756-3.

Colella, P., Ronzitti, G. and Mingozzi, F. (2018) 'Emerging Issues in AAV-

Mediated In Vivo Gene Therapy.', *Molecular therapy. Methods & clinical development*. American Society of Gene & Cell Therapy, 8, pp. 87–104. doi: 10.1016/j.omtm.2017.11.007.

Cong, L. *et al.* (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science (New York, N.Y.)*. Science, 339(6121), pp. 819–823. doi: 10.1126/SCIENCE.1231143.

Cooper, G. M. (2000) *The cell : a molecular approach*. ASM Press.

Cooper, L. A., Stringer, A. M. and Wade, J. T. (2018) 'Determining the Specificity of Cascade Binding, Interference, and Primed Adaptation In Vivo in the Escherichia coli Type I-E CRISPR-Cas System', *mBio*. mBio, 9(2). doi: 10.1128/MBIO.02100-17.

Cooper, R. N. *et al.* (1999) 'In vivo satellite cell activation via Myf5 and MyoD in regenerating mouse skeletal muscle', *Journal of cell science*. J Cell Sci, 112 (Pt 17)(17), pp. 2895–2901. doi: 10.1242/JCS.112.17.2895.

Cossu, G. *et al.* (2015) ' Intra-arterial transplantation of HLA -matched donor mesoangioblasts in Duchenne muscular dystrophy ', *EMBO Molecular Medicine*. EMBO, 7(12), pp. 1513–1528. doi: 10.15252/emmm.201505636.

Cost, G. J. *et al.* (2010) 'BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells', *Biotechnology and bioengineering*. Biotechnol Bioeng, 105(2), pp. 330–340. doi: 10.1002/BIT.22541.

Cotton, S., Voudouris, N. J. and Greenwood, K. M. (2001) 'Intelligence and Duchenne muscular dystrophy: full-scale, verbal, and performance intelligence quotients', *Developmental medicine and child neurology*. Dev Med Child Neurol, 43(7), p. 497. doi: 10.1017/S0012162201000913.

Dabiré, H. *et al.* (2012) 'Vascular endothelial dysfunction in Duchenne muscular dystrophy is restored by bradykinin through upregulation of eNOS and nNOS', *Basic research in cardiology*. Basic Res Cardiol, 107(1). doi: 10.1007/S00395-011-0240-6.

Dalby, B. *et al.* (2004) 'Advanced transfection with Lipofectamine 2000 reagent: primary neurons, siRNA, and high-throughput applications', *Methods (San Diego, Calif.)*. Methods, 33(2), pp. 95–103. doi: 10.1016/J.YMETH.2003.11.023.

Daoud, F. *et al.* (2009) 'Analysis of Dp71 contribution in the severity of mental retardation through comparison of Duchenne and Becker patients differing by mutation consequences on Dp71 expression', *Human molecular genetics*. Hum Mol Genet, 18(20), pp. 3779–3794. doi: 10.1093/HMG/DDP320.

Deconinck, N. and Dan, B. (2007) 'Pathophysiology of Duchenne Muscular Dystrophy: Current Hypotheses', *Pediatric Neurology*, 36(1), pp. 1–7. doi: 10.1016/j.pediatrneurol.2006.09.016.

Decroix, L. *et al.* (2015) 'Tissue clearing for confocal imaging of native and bio-artificial skeletal muscle', *Biotechnic and Histochemistry*. Taylor and Francis

Ltd., 90(6), pp. 424–431. doi: 10.3109/10520295.2015.1019564.

Deltcheva, E. *et al.* (2011) 'CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III', *Nature 2011 471:7340*. Nature Publishing Group, 471(7340), pp. 602–607. doi: 10.1038/nature09886.

Deng, D. *et al.* (2012) 'Structural basis for sequence-specific recognition of DNA by TAL effectors', *Science (New York, N.Y.)*. Science, 335(6069), pp. 720–723. doi: 10.1126/SCIENCE.1215670.

DeWeirdt, P. C. *et al.* (2020) 'Genetic screens in isogenic mammalian cell lines without single cell cloning', *Nature Communications 2020 11:1*. Nature Publishing Group, 11(1), pp. 1–15. doi: 10.1038/s41467-020-14620-6.

Doench, J. G. *et al.* (2014) 'Rational design of highly active sgRNAs for CRISPR-Cas9–mediated gene inactivation', *Nature Biotechnology*. Nature Publishing Group, 32(12), pp. 1262–1267. doi: 10.1038/nbt.3026.

Doench, J. G. *et al.* (2016) 'Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9', *Nature Biotechnology 2015 34:2*. Nature Publishing Group, 34(2), pp. 184–191. doi: 10.1038/nbt.3437.

Doorenweerd, N. *et al.* (2017) 'Timing and localization of human dystrophin isoform expression provide insights into the cognitive phenotype of Duchenne muscular dystrophy', *Scientific Reports 2017 7:1*. Nature Publishing Group, 7(1), pp. 1–12. doi: 10.1038/s41598-017-12981-5.

Dow, L. E. (2015) 'Modeling disease in vivo with CRISPR/Cas9', *Trends in molecular medicine*. NIH Public Access, 21(10), p. 609. doi: 10.1016/J.MOLMED.2015.07.006.

Duan, D. *et al.* (2015) 'Early loss of ambulation is not a representative clinical feature in Duchenne muscular dystrophy dogs: remarks on the article of Barthelemy et al.', *Disease Models & Mechanisms*, 8(3), pp. 193–194. doi: 10.1242/dmm.019216.

Duan, D. (2018) 'Micro-Dystrophin Gene Therapy Goes Systemic in Duchenne Muscular Dystrophy Patients', *https://home.liebertpub.com/hum*. Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA , 29(7), pp. 733–736. doi: 10.1089/HUM.2018.012.

Duan, D. *et al.* (2021) 'Duchenne muscular dystrophy', *Nature Reviews Disease Primers 2021 7:1*. Nature Publishing Group, 7(1), pp. 1–19. doi: 10.1038/s41572-021-00248-3.

Duan, X., Gimble, F. S. and Quiocho, F. A. (1997) 'Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity', *Cell*. Cell, 89(4), pp. 555–564. doi: 10.1016/S0092-8674(00)80237-8.

Dumont, N. A. *et al.* (2015) 'Dystrophin expression in muscle stem cells regulates their polarity and asymmetric division.', *Nature medicine*. NIH Public Access, 21(12), pp. 1455–63. doi: 10.1038/nm.3990.

Echigoya, Y. *et al.* (2017) ‘Effects of systemic multiexon skipping with peptide-conjugated morpholinos in the heart of a dog model of Duchenne muscular dystrophy’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(16), pp. 4213–4218. doi: 10.1073/PNAS.1613203114/SUPPL_FILE/PNAS.201613203SI.PDF.

Eiraku, M. and Sasai, Y. (2012) ‘Self-formation of layered neural structures in three-dimensional culture of ES cells’, *Current Opinion in Neurobiology*. Elsevier Current Trends, pp. 768–777. doi: 10.1016/j.conb.2012.02.005.

Emery, A. E. (2002) ‘The muscular dystrophies’, *The Lancet*, 359(9307), pp. 687–695. doi: 10.1016/S0140-6736(02)07815-7.

England, S. B. *et al.* (1990) ‘Very mild muscular dystrophy associated with the deletion of 46% of dystrophin’, *Nature*, 343(6254), pp. 180–182. doi: 10.1038/343180a0.

Engler, A. J. *et al.* (2006) ‘Matrix Elasticity Directs Stem Cell Lineage Specification’, *Cell*. Cell Press, 126(4), pp. 677–689. doi: 10.1016/J.CELL.2006.06.044.

Evans, M. J. and Kaufman, M. H. (1981) ‘Establishment in culture of pluripotential cells from mouse embryos’, *Nature*. Nature Publishing Group, 292(5819), pp. 154–156. doi: 10.1038/292154a0.

Exeter, D. and Connell, D. (2010) ‘Skeletal Muscle: Functional Anatomy and Pathophysiology’, *Seminars in Musculoskeletal Radiology*, 14(02), pp. 097–105. doi: 10.1055/s-0030-1253154.

Fairclough, R. J., Wood, M. J. and Davies, K. E. (2013) ‘Therapy for Duchenne muscular dystrophy: renewed optimism from genetic approaches’, *Nature Reviews Genetics*, 14(6), pp. 373–378. doi: 10.1038/nrg3460.

Falzarano, M. S., Passarelli, C. and Ferlini, A. (2014) ‘Nanoparticle delivery of antisense oligonucleotides and their application in the exon skipping strategy for Duchenne muscular dystrophy’, *Nucleic acid therapeutics*. Nucleic Acid Ther, 24(1), pp. 87–100. doi: 10.1089/NAT.2013.0450.

Farea, M. *et al.* (2020) ‘Dystrophin Dp71ab is monoclonally expressed in human satellite cells and enhances proliferation of myoblast cells’, *Scientific reports*. Sci Rep, 10(1). doi: 10.1038/S41598-020-74157-Y.

Farea, M. *et al.* (2022) ‘Human Dystrophin Dp71ab Enhances the Proliferation of Myoblasts Across Species But Not Human Nonmyoblast Cells’, *Frontiers in Cell and Developmental Biology*. Frontiers, 0, p. 873. doi: 10.3389/FCELL.2022.877612.

Fatehullah, A., Tan, S. H. and Barker, N. (2016) ‘Organoids as an in vitro model of human development and disease’, *Nature Cell Biology*, 18(3), pp. 246–254. doi: 10.1038/ncb3312.

Faustino Martins, J. M. *et al.* (2020) ‘Self-Organizing 3D Human Trunk Neuromuscular Organoids’, *Cell Stem Cell*. Cell Press, 26(2), pp. 172-186.e6.

doi: 10.1016/j.stem.2019.12.007.

Felisari, G. *et al.* (2000) 'Loss of Dp140 dystrophin isoform and intellectual impairment in Duchenne dystrophy', *Neurology*. Neurology, 55(4), pp. 559–564. doi: 10.1212/WNL.55.4.559.

Ferrari, G. *et al.* (2021) 'DLL4 and PDGF-BB regulate migration of human iPSC-derived skeletal myogenic progenitors', *bioRxiv*. Cold Spring Harbor Laboratory, p. 2021.02.28.431778. doi: 10.1101/2021.02.28.431778.

Flynn, R. *et al.* (2015) 'CRISPR-mediated genotypic and phenotypic correction of a chronic granulomatous disease mutation in human iPS cells', *Experimental hematology*. Exp Hematol, 43(10), pp. 838-848.e3. doi: 10.1016/J.EXPHEM.2015.06.002.

Foley, J. E. *et al.* (2009) 'Targeted mutagenesis in zebrafish using customized zinc-finger nucleases', *Nature protocols*. Nat Protoc, 4(12), pp. 1855–1868. doi: 10.1038/NPROT.2009.209.

Fong, H. *et al.* (2011) 'Nucleofection of human embryonic stem cells', *Methods in molecular biology (Clifton, N.J.)*. Methods Mol Biol, 767, pp. 333–341. doi: 10.1007/978-1-61779-201-4_24.

Freedman, B. S. *et al.* (2015) 'Modelling kidney disease with CRISPR-mutant kidney organoids derived from human pluripotent epiblast spheroids', *Nature Communications*. Nature Publishing Group, 6. doi: 10.1038/NCOMMS9715.

Frock, R. L. *et al.* (2015) 'Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases', *Nature biotechnology*. Nat Biotechnol, 33(2), pp. 179–188. doi: 10.1038/NBT.3101.

Frontera, W. R. and Ochala, J. (2015) 'Skeletal Muscle: A Brief Review of Structure and Function', *Calcified Tissue International*, 96(3), pp. 183–195. doi: 10.1007/s00223-014-9915-y.

Fukada, S. *et al.* (2013) 'Isolation, characterization, and molecular regulation of muscle stem cells', *Frontiers in Physiology*. Frontiers, 4, p. 317. doi: 10.3389/fphys.2013.00317.

Funanage, V. L., Smith, S. M. and Minnich, M. A. (1992) 'Entactin promotes adhesion and long-term maintenance of cultured regenerated skeletal myotubes', *Journal of Cellular Physiology*. J Cell Physiol, 150(2), pp. 251–257. doi: 10.1002/jcp.1041500205.

Garneau, J. E. *et al.* (2010) 'The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA', *Nature 2010 468:7320*. Nature Publishing Group, 468(7320), pp. 67–71. doi: 10.1038/nature09523.

Gentil, C. *et al.* (2016) 'Dystrophin Threshold Level Necessary for Normalization of Neuronal Nitric Oxide Synthase, Inducible Nitric Oxide Synthase, and Ryanodine Receptor-Calcium Release Channel Type 1 Nitrosylation in Golden Retriever Muscular Dystrophy Dystrophinopathy', *Human*

gene therapy. Hum Gene Ther, 27(9), pp. 712–726. doi: 10.1089/HUM.2016.041.

Gerli, M. F. M. *et al.* (2019) 'Combined Notch and PDGF Signaling Enhances Migration and Expression of Stem Cell Markers while Inducing Perivascular Cell Features in Muscle Satellite Cells', *Stem Cell Reports*. Cell Press, 12(3), pp. 461–473. doi: 10.1016/J.STEMCR.2019.01.007.

Geurts, A. M. *et al.* (2009) 'Knockout rats via embryo microinjection of zinc-finger nucleases', *Science (New York, N.Y.)*. Science, 325(5939), p. 433. doi: 10.1126/SCIENCE.1172447.

Giacomelli, E. *et al.* (2017) 'Three-dimensional cardiac microtissues composed of cardiomyocytes and endothelial cells co-differentiated from human pluripotent stem cells', *Development*, 144(6), pp. 1008–1017. doi: 10.1242/dev.143438.

Gilbert, P. M. *et al.* (2010) 'Substrate elasticity regulates skeletal muscle stem cell self-renewal in culture', *Science*. American Association for the Advancement of Science, 329(5995), pp. 1078–1081. doi: 10.1126/SCIENCE.1191035/SUPPL_FILE/GILBERT_SOM.PDF.

Godfrey, C. *et al.* (2015) 'How much dystrophin is enough: the physiological consequences of different levels of dystrophin in the mdx mouse', *Human molecular genetics*. Hum Mol Genet, 24(15), pp. 4225–4237. doi: 10.1093/HMG/DDV155.

Goldstein, R. A. (2017) 'Skeletal Muscle Injury Biomarkers: Assay Qualification Efforts and Translation to the Clinic', *Toxicologic Pathology*. SAGE Publications Inc., 45(7), pp. 943–951. doi: 10.1177/0192623317738927.

Gong, S. *et al.* (2018) 'DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity', *Cell reports*. Cell Rep, 22(2), pp. 359–371. doi: 10.1016/J.CELREP.2017.12.041.

Górecki, D. C. *et al.* (1992) 'Expression of four alternative dystrophin transcripts in brain regions regulated by different promoters', *Human Molecular Genetics*, 1(7), pp. 505–510. doi: 10.1093/hmg/1.7.505.

Goudenege, S. *et al.* (2012) 'Myoblasts Derived From Normal hESCs and Dystrophic hiPSCs Efficiently Fuse With Existing Muscle Fibers Following Transplantation', *Molecular Therapy*, 20(11), pp. 2153–2167. doi: 10.1038/mt.2012.188.

Gouti, M. *et al.* (2014) 'In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity', *PLOS Biology*. Public Library of Science, 12(8), p. e1001937. doi: 10.1371/JOURNAL.PBIO.1001937.

Gregorevic, P. *et al.* (2004) 'Systemic delivery of genes to striated muscles using adeno-associated viral vectors', *Nature medicine*. Nat Med, 10(8), pp. 828–834. doi: 10.1038/NM1085.

Groenen, P. M. A. *et al.* (1993) 'Nature of DNA polymorphism in the direct

repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method', *Molecular microbiology*. Mol Microbiol, 10(5), pp. 1057–1065. doi: 10.1111/J.1365-2958.1993.TB00976.X.

Gu, J. *et al.* (2010) 'DNA-PKcs Regulates a Single-stranded DNA Endonuclease Activity of Artemis', *DNA repair*. NIH Public Access, 9(4), p. 429. doi: 10.1016/J.DNAREP.2010.01.001.

Guncay, A. and Yokota, T. (2015) 'Antisense oligonucleotide drugs for Duchenne muscular dystrophy: how far have we come and what does the future hold?', *Future Medicinal Chemistry*, 7(13), pp. 1631–1635. doi: 10.4155/fmc.15.116.

Gupta, R. M. *et al.* (2016) 'Genome-Edited Human Pluripotent Stem Cell-Derived Macrophages as a Model of Reverse Cholesterol Transport--Brief Report', *Arteriosclerosis, thrombosis, and vascular biology*. Arterioscler Thromb Vasc Biol, 36(1), pp. 15–18. doi: 10.1161/ATVBAHA.115.305956.

Gupta, R. M. and Musunuru, K. (2014) 'Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9', *The Journal of Clinical Investigation*. American Society for Clinical Investigation, 124(10), pp. 4154–4161. doi: 10.1172/JCI72992.

Gussoni, E. *et al.* (1992) 'Normal dystrophin transcripts detected in Duchenne muscular dystrophy patients after myoblast transplantation', *Nature*, 356(6368), pp. 435–438. doi: 10.1038/356435a0.

Hartig, R. *et al.* (1998) 'Active nuclear import of single-stranded oligonucleotides and their complexes with non-karyophilic macromolecules', *Biology of the Cell*. John Wiley & Sons, Ltd, 90(5), pp. 407–426. doi: 10.1111/J.1768-322X.1998.TB01050.X.

Hathout, Y. *et al.* (2016) 'Clinical utility of serum biomarkers in Duchenne muscular dystrophy', *Clinical Proteomics*. BioMed Central Ltd. doi: 10.1186/s12014-016-9109-x.

Hay, M. *et al.* (2014) 'Clinical development success rates for investigational drugs', *Nature Biotechnology*. Nat Biotechnol, 32(1), pp. 40–51. doi: 10.1038/nbt.2786.

Heath, P. J. *et al.* (1997) 'The structure of I-CreI, a group I intron-encoded homing endonuclease', *Nature structural biology*. Nat Struct Biol, 4(6), pp. 468–476. doi: 10.1038/NSB0697-468.

Heckl, D. *et al.* (2014) 'Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing', *Nature biotechnology*. NIH Public Access, 32(9), p. 941. doi: 10.1038/NBT.2951.

Heler, R. *et al.* (2015) 'Cas9 specifies functional viral targets during CRISPR–Cas adaptation', *Nature 2015 519:7542*. Nature Publishing Group, 519(7542), pp. 199–202. doi: 10.1038/nature14245.

Henrique, D. *et al.* (2015) ‘Neuromesodermal progenitors and the making of the spinal cord’, *Development (Cambridge, England)*. Development, 142(17), pp. 2864–2875. doi: 10.1242/DEV.119768.

Hicks, M. R. *et al.* (2018) ‘ERBB3 and NGFR mark a distinct skeletal muscle progenitor cell in human development and hPSCs’, *Nature Cell Biology*, 20(1), pp. 46–57. doi: 10.1038/s41556-017-0010-2.

Hildyard, J. C. W. *et al.* (2018) ‘Determination of qPCR Reference Genes Suitable for Normalizing Gene Expression in a Canine Model of Duchenne Muscular Dystrophy’, *Journal of neuromuscular diseases*. J Neuromuscul Dis, 5(2), pp. 177–191. doi: 10.3233/JND-170267.

Hoffman, E. P. *et al.* (1988) ‘Characterization of Dystrophin in Muscle-Biopsy Specimens from Patients with Duchenne’s or Becker’s Muscular Dystrophy’, *New England Journal of Medicine*, 318(21), pp. 1363–1368. doi: 10.1056/NEJM198805263182104.

Hoffman, E. P., Kunkel, L. M. and Brown, R. H. (1988) ‘Proteolytic fragment or new gene product?’, *Nature*. Nature, 336(6196), p. 210. doi: 10.1038/336210A0.

Hohenstein, K. A. *et al.* (2008) ‘Nucleofection mediates high-efficiency stable gene knockdown and transgene expression in human embryonic stem cells’, *Stem cells (Dayton, Ohio)*. Stem Cells, 26(6), pp. 1436–1443. doi: 10.1634/STEMCELLS.2007-0857.

Holder, E., Maeda, M. and Bies, R. D. (1996) ‘Expression and regulation of the dystrophin Purkinje promoter in human skeletal muscle, heart, and brain’, *Human genetics*. Hum Genet, 97(2), pp. 232–239. doi: 10.1007/BF02265272.

Hopkins, P. M. (2006) ‘Skeletal muscle physiology’, *Continuing Education in Anaesthesia Critical Care & Pain*. Narnia, 6(1), pp. 1–6. doi: 10.1093/bjaceaccp/mki062.

Horii, T. *et al.* (2013) ‘Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system’, *International journal of molecular sciences*. Int J Mol Sci, 14(10), pp. 19774–19781. doi: 10.3390/IJMS141019774.

Horlbeck, M. A. *et al.* (2016) ‘Nucleosomes impede cas9 access to DNA in vivo and in vitro’, *eLife*. eLife Sciences Publications Ltd, 5(MARCH2016). doi: 10.7554/ELIFE.12677.

Horst, D. *et al.* (2006) ‘Comparative expression analysis of Pax3 and Pax7 during mouse myogenesis’, *The International Journal of Developmental Biology*, 50(1), pp. 47–54. doi: 10.1387/ijdb.052111dh.

Howe, S. J. *et al.* (2008) ‘Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients’, *The Journal of clinical investigation*. J Clin Invest, 118(9), pp. 3143–3150. doi: 10.1172/JCI35798.

Hoy, S. M. (2017) ‘Nusinersen: First Global Approval’, *Drugs*. Drugs, 77(4), pp. 473–479. doi: 10.1007/S40265-017-0711-7.

Hsu, P. D. *et al.* (2013) ‘DNA targeting specificity of RNA-guided Cas9 nucleases’, *Nature Biotechnology*. Nature Publishing Group, 31(9), pp. 827–832. doi: 10.1038/nbt.2647.

Huch, M. *et al.* (2015) ‘Long-term culture of genome-stable bipotent stem cells from adult human liver’, *Cell*. Cell Press, 160(1–2), pp. 299–312. doi: 10.1016/j.cell.2014.11.050.

Hudziak, R. M. *et al.* (1996) ‘Resistance of morpholino phosphorodiamidate oligomers to enzymatic degradation’, *Antisense & nucleic acid drug development*. Antisense Nucleic Acid Drug Dev, 6(4), pp. 267–272. doi: 10.1089/OLI.1.1996.6.267.

Hug, N., Longman, D. and Cáceres, J. F. (2016) ‘Mechanism and regulation of the nonsense-mediated decay pathway’, *Nucleic acids research*. Nucleic Acids Res, 44(4), pp. 1483–1495. doi: 10.1093/NAR/GKW010.

Huxley, H. E. (1969) ‘The Mechanism of Muscular Contraction’, *Science*, 164(3886), pp. 1356–1366. doi: 10.1126/science.164.3886.1356.

Ira, G. *et al.* (2004) ‘DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1’, *Nature*, 431(7011), pp. 1011–1017. doi: 10.1038/NATURE02964.

Ishikawa, T. *et al.* (2016) ‘Genetic and pharmacological correction of aberrant dopamine synthesis using patient iPSCs with BH4 metabolism disorders’, *Human molecular genetics*. Hum Mol Genet, 25(23), pp. 5188–5197. doi: 10.1093/HMG/DDW339.

Ishino, Y. *et al.* (1987) ‘Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product’, *Journal of bacteriology*. J Bacteriol, 169(12), pp. 5429–5433. doi: 10.1128/JB.169.12.5429-5433.1987.

Ivanova, G. D. *et al.* (2008) ‘PNA-peptide conjugates as intracellular gene control agents’, *Nucleic acids symposium series (2004)*. Nucleic Acids Symp Ser (Oxf), (52), pp. 31–32. doi: 10.1093/NASS/NRN016.

Jalal, S., Dastidar, S. and Tedesco, F. S. (2021) ‘Advanced models of human skeletal muscle differentiation, development and disease: Three-dimensional cultures, organoids and beyond’, *Current opinion in cell biology*. Curr Opin Cell Biol, 73, pp. 92–104. doi: 10.1016/J.CEB.2021.06.004.

Jang, D. E. *et al.* (2018) ‘Multiple sgRNAs with overlapping sequences enhance CRISPR/Cas9-mediated knock-in efficiency’, *Experimental & Molecular Medicine*, 50(4), p. 16. doi: 10.1038/s12276-018-0037-x.

Jansen, Rund *et al.* (2002) ‘Identification of a novel family of sequence repeats among prokaryotes’, *Omics : a journal of integrative biology*. OMICS, 6(1), pp. 23–33. doi: 10.1089/15362310252780816.

Jansen, Ruud *et al.* (2002) 'Identification of genes that are associated with DNA repeats in prokaryotes', *Molecular microbiology*. Mol Microbiol, 43(6), pp. 1565–1575. doi: 10.1046/J.1365-2958.2002.02839.X.

Janssen, I. *et al.* (2000) 'Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr', *Journal of Applied Physiology*, 89(1), pp. 81–88. doi: 10.1152/jappl.2000.89.1.81.

Jearawiriyapaisarn, N. *et al.* (2008) 'Sustained Dystrophin Expression Induced by Peptide-conjugated Morpholino Oligomers in the Muscles of mdx Mice', *Molecular therapy : the journal of the American Society of Gene Therapy*. NIH Public Access, 16(9), p. 1624. doi: 10.1038/MT.2008.120.

Jelinkova, S. *et al.* (2020) 'DMD Pluripotent Stem Cell Derived Cardiac Cells Recapitulate in vitro Human Cardiac Pathophysiology', *Frontiers in Bioengineering and Biotechnology*. Frontiers Media S.A., 8, p. 535. doi: 10.3389/FBIOE.2020.00535/FULL.

Jensen, K. T. *et al.* (2017) 'Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency', *FEBS letters*. FEBS Lett, 591(13), pp. 1892–1901. doi: 10.1002/1873-3468.12707.

Jiang, F. *et al.* (2016) 'Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage', *Science (New York, N.Y.)*. NIH Public Access, 351(6275), p. 867. doi: 10.1126/SCIENCE.AAD8282.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science (New York, N.Y.)*. Science, 337(6096), pp. 816–821. doi: 10.1126/SCIENCE.1225829.

Jinek, M. *et al.* (2014) 'Structures of Cas9 endonucleases reveal RNA-mediated conformational activation', *Science (New York, N.Y.)*. Science, 343(6176). doi: 10.1126/SCIENCE.1247997.

Jiwlawat, N. *et al.* (2019) 'Micropatterned substrates with physiological stiffness promote cell maturation and Pompe disease phenotype in human induced pluripotent stem cell-derived skeletal myocytes', *Biotechnology and Bioengineering*. John Wiley & Sons, Ltd, 116(9), pp. 2377–2392. doi: 10.1002/BIT.27075.

Jurica, M. S. and Stoddard, B. L. (1999) 'Homing endonucleases: structure, function and evolution', *Cellular and molecular life sciences : CMLS*. Cell Mol Life Sci, 55(10), pp. 1304–1326. doi: 10.1007/S000180050372.

Karpati, G. *et al.* (1993) 'Myoblast transfer in duchenne muscular dystrophy', *Annals of Neurology*. John Wiley & Sons, Ltd, 34(1), pp. 8–17. doi: 10.1002/ana.410340105.

Kassar-Duchossoy, L. *et al.* (2005) 'Pax3/Pax7 mark a novel population of primitive myogenic cells during development', *Genes & Development*. Cold Spring Harbor Laboratory Press, 19(12), pp. 1426–1431. doi: 10.1101/GAD.345505.

Kazuki, Y. *et al.* (2010) ‘Complete Genetic Correction of iPS Cells From Duchenne Muscular Dystrophy’, *Molecular Therapy*, 18(2), pp. 386–393. doi: 10.1038/mt.2009.274.

Khan, Faheem Ahmed *et al.* (2016) ‘CRISPR/Cas9 therapeutics: a cure for cancer and other genetic diseases’, *Oncotarget*. Oncotarget, 7(32), pp. 52541–52552. doi: 10.18632/ONCOTARGET.9646.

Khedr, M. * *et al.* (2022) ‘3D Human iPS Cell-derived Bioengineered Skeletal Muscles for Tissue, Disease and Therapy Modelling’, *(Under review for Nature Protocols)*.

Khodabukus, A. *et al.* (2020) ‘Tissue-engineered human myobundle system as a platform for evaluation of skeletal muscle injury biomarkers’, *Toxicological Sciences*. Oxford University Press, 176(1), pp. 124–136. doi: 10.1093/toxsci/kfaa049.

Kim, J. *et al.* (2017) ‘Expansion and Purification Are Critical for the Therapeutic Application of Pluripotent Stem Cell-Derived Myogenic Progenitors’, *Stem Cell Reports*. Cell Press, 9(1), pp. 12–22. doi: 10.1016/J.STEMCR.2017.04.022.

Kim, Ji Hyun *et al.* (2018) ‘3D Bioprinted Human Skeletal Muscle Constructs for Muscle Function Restoration’, *SciEnTiFic RepoRts* |, 8, p. 12307. doi: 10.1038/s41598-018-29968-5.

Kim, June Hoan *et al.* (2018) ‘Optimizing tissue-clearing conditions based on analysis of the critical factors affecting tissue-clearing procedures’, *Scientific Reports 2018 8:1*. Nature Publishing Group, 8(1), pp. 1–11. doi: 10.1038/s41598-018-31153-7.

Kim, N. *et al.* (2020) ‘Prediction of the sequence-specific cleavage activity of Cas9 variants’, *Nature Biotechnology 2020 38:11*. Nature Publishing Group, 38(11), pp. 1328–1336. doi: 10.1038/s41587-020-0537-9.

Kim, Y. G. *et al.* (1998) ‘Chimeric restriction enzyme: Gal4 fusion to FokI cleavage domain’, *Biological chemistry*. Biol Chem, 379(4–5), pp. 489–496. doi: 10.1515/BCHM.1998.379.4-5.489.

Kim, Y. G., Cha, J. and Chandrasegaran, S. (1996) ‘Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 93(3), p. 1156. doi: 10.1073/PNAS.93.3.1156.

Kim, Y. G. and Chandrasegaran, S. (1994) ‘Chimeric restriction endonuclease.’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 91(3), p. 883. doi: 10.1073/PNAS.91.3.883.

Kleinjan, D. J. and Coutinho, P. (2009) ‘Cis-ruption mechanisms: disruption of cis-regulatory control as a cause of human genetic disease’, *Briefings in functional genomics & proteomics*. Oxford University Press, 8(4), pp.

317–32. doi: 10.1093/BFGP/ELP022.

Kleopa, K. A. *et al.* (2006) ‘Naturally occurring utrophin correlates with disease severity in Duchenne muscular dystrophy’, *Human Molecular Genetics*. Oxford Academic, 15(10), pp. 1623–1628. doi: 10.1093/HMG.

Kobinger, G. P. *et al.* (2003) ‘Correction of the Dystrophic Phenotype by *In Vivo* Targeting of Muscle Progenitor Cells’, *Human Gene Therapy*, 14(15), pp. 1441–1449. doi: 10.1089/104303403769211655.

Kodippili, K. *et al.* (2014) ‘Characterization of 65 Epitope-Specific Dystrophin Monoclonal Antibodies in Canine and Murine Models of Duchenne Muscular Dystrophy by Immunostaining and Western Blot’, *PLoS ONE*. Public Library of Science, 9(2), p. 88280. doi: 10.1371/JOURNAL.PONE.0088280.

Kondash, M. E. *et al.* (2020) ‘Glucose Uptake and Insulin Response in Tissue-engineered Human Skeletal Muscle’, *Tissue Engineering and Regenerative Medicine*. Korean Tissue Engineering and Regenerative Medicine Society, 17(6), pp. 801–813. doi: 10.1007/s13770-020-00242-y.

Koo, T. *et al.* (2011) ‘Long-term functional adeno-associated virus-microdystrophin expression in the dystrophic CXMDj dog’, *The Journal of Gene Medicine*, 13(9), pp. 497–506. doi: 10.1002/jgm.1602.

Koo, Y., Jung, D. kyo and Bae, E. (2012) ‘Crystal Structure of Streptococcus pyogenes Csn2 Reveals Calcium-Dependent Conformational Changes in Its Tertiary and Quaternary Structure’, *PLOS ONE*. Public Library of Science, 7(3), p. e33401. doi: 10.1371/JOURNAL.PONE.0033401.

Koonin, E. V., Makarova, K. S. and Zhang, F. (2017) ‘Diversity, classification and evolution of CRISPR-Cas systems’, *Current opinion in microbiology*. Curr Opin Microbiol, 37, pp. 67–78. doi: 10.1016/J.MIB.2017.05.008.

Kouprina, N. *et al.* (2014) ‘Human artificial chromosome-based gene delivery vectors for biomedicine and biotechnology’, *Expert Opinion on Drug Delivery*, 11(4), pp. 517–535. doi: 10.1517/17425247.2014.882314.

Kouranova, E. *et al.* (2016) ‘CRISPRs for Optimal Targeting: Delivery of CRISPR Components as DNA, RNA, and Protein into Cultured Cells and Single-Cell Embryos.’, *Human gene therapy*. Mary Ann Liebert, Inc., 27(6), pp. 464–75. doi: 10.1089/hum.2016.009.

Kraft, K. *et al.* (2015) ‘Deletions, inversions, duplications: Engineering of structural variants using CRISPR/Cas in mice’, *Cell Reports*. Elsevier B.V., 10(5), pp. 833–839. doi: 10.1016/J.CELREP.2015.01.016.

Lakshmipathy, U. *et al.* (2004) ‘Efficient transfection of embryonic and adult stem cells’, *Stem cells (Dayton, Ohio)*. Stem Cells, 22(4), pp. 531–543. doi: 10.1634/STEMCELLS.22-4-531.

Lam, L. T. *et al.* (2010) ‘Exon-specific dystrophin antibodies for studies of duchenne muscular dystrophy’, *Translational Neuroscience*, 1(3), pp. 233–237.

doi: 10.2478/V10134-010-0034-7.

Lancaster, M. A. *et al.* (2013) ‘Cerebral organoids model human brain development and microcephaly’, *Nature*. Nature Publishing Group, 501(7467), pp. 373–379. doi: 10.1038/nature12517.

Lancaster, M. A. and Knoblich, J. A. (2014) ‘Organogenesis in a dish: Modeling development and disease using organoid technologies’, *Science*, 345(6194), pp. 1247125–1247125. doi: 10.1126/science.1247125.

Laoharawee, K. *et al.* (2018) ‘Dose-Dependent Prevention of Metabolic and Neurologic Disease in Murine MPS II by ZFN-Mediated In Vivo Genome Editing’, *Molecular therapy : the journal of the American Society of Gene Therapy*. Mol Ther, 26(4), pp. 1127–1136. doi: 10.1016/J.YMTHE.2018.03.002.

Lapidos, K. A., Kakkar, R. and Mcnally, E. M. (2004) ‘The Dystrophin Glycoprotein Complex Signaling Strength and Integrity for the Sarcolemma’, *Circulation Research*, 94, pp. 1023–1031. doi: 10.1161/01.RES.0000126574.61061.25.

Larcher, T. *et al.* (2014) ‘Characterization of dystrophin deficient rats: A new model for duchenne muscular dystrophy’, *PLoS ONE*. Public Library of Science, 9(10). doi: 10.1371/journal.pone.0110371.

Latour, Y. L. *et al.* (2019) ‘Human GLB1 knockout cerebral organoids: A model system for testing AAV9-mediated GLB1 gene therapy for reducing GM1 ganglioside storage in GM1 gangliosidosis’, *Molecular genetics and metabolism reports*. Mol Genet Metab Rep, 21. doi: 10.1016/J.YMGMR.2019.100513.

Lee, J. K. *et al.* (2018) ‘Directed evolution of CRISPR-Cas9 to increase its specificity’, *Nature Communications 2018 9:1*. Nature Publishing Group, 9(1), pp. 1–10. doi: 10.1038/s41467-018-05477-x.

Lehto, T. *et al.* (2016) ‘Peptides for nucleic acid delivery’, *Advanced drug delivery reviews*. Adv Drug Deliv Rev, 106(Pt A), pp. 172–182. doi: 10.1016/J.ADDR.2016.06.008.

Lessard, S. *et al.* (2017) ‘Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(52), pp. E11257–E11266. doi: 10.1073/PNAS.1714640114/SUPPL_FILE/PNAS.1714640114.SD09.XLSX.

Levin, A. A. (2019) ‘Treating Disease at the RNA Level with Oligonucleotides’, *The New England journal of medicine*. N Engl J Med, 380(1), pp. 57–70. doi: 10.1056/NEJMRA1705346.

Li, G. *et al.* (2018) ‘Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts’, *The international journal of biochemistry & cell biology*. Int J Biochem Cell Biol, 99, pp. 154–160. doi: 10.1016/J.BIOCEL.2018.04.011.

Li, H. L. *et al.* (2015) ‘Precise correction of the dystrophin gene in

duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9', *Stem Cell Reports*. Cell Press, 4(1), pp. 143–154. doi: 10.1016/j.stemcr.2014.10.013.

Li, K. *et al.* (2014) 'Optimization of Genome Engineering Approaches with the CRISPR/Cas9 System', *PLOS ONE*. Public Library of Science, 9(8), p. e105779. doi: 10.1371/JOURNAL.PONE.0105779.

Li, L., Wu, L. P. and Chandrasegaran, S. (1992) 'Functional domains in Fok I restriction endonuclease', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 89(10), pp. 4275–4279. doi: 10.1073/PNAS.89.10.4275.

Li, M. *et al.* (2014) 'The catalytic subunit of DNA-dependent protein kinase is required for cellular resistance to oxidative stress independent of DNA double-strand break repair', *Free radical biology & medicine*. Free Radic Biol Med, 76, pp. 278–285. doi: 10.1016/J.FREERADBIOMED.2014.08.019.

Li, P. *et al.* (2015) 'Efficient generation of genetically distinct pigs in a single pregnancy using multiplexed single-guide RNA and carbohydrate selection', *Xenotransplantation*. Blackwell Publishing Inc., 22(1), pp. 20–31. doi: 10.1111/XEN.12131.

Li, Y. *et al.* (2017) 'Induction of Expansion and Folding in Human Cerebral Organoids', *Cell Stem Cell*. Cell Press, 20(3), pp. 385-396.e3. doi: 10.1016/J.STEM.2016.11.017.

Lieber, M. R. (2010) 'The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway', *Annual review of biochemistry*. NIH Public Access, 79, p. 181. doi: 10.1146/ANNUREV.BIOCHEM.052308.093131.

Lim, K. R. Q. *et al.* (2019) 'Efficacy of Multi-exon Skipping Treatment in Duchenne Muscular Dystrophy Dog Model Neonates', *Molecular Therapy*. American Society of Gene & Cell Therapy, 27(1), p. 76. doi: 10.1016/J.YMTHE.2018.10.011.

Lim, K. R. Q., Yoon, C. and Yokota, T. (2018) 'Applications of CRISPR/Cas9 for the Treatment of Duchenne Muscular Dystrophy', *Journal of Personalized Medicine 2018, Vol. 8, Page 38*. Multidisciplinary Digital Publishing Institute, 8(4), p. 38. doi: 10.3390/JPM8040038.

Limbo, O. *et al.* (2007) 'Ctp1 is a Cell Cycle-Regulated Protein that Functions with Mre11 Complex to Control Double-Strand Break Repair by Homologous Recombination', *Molecular cell*. NIH Public Access, 28(1), p. 134. doi: 10.1016/J.MOLCEL.2007.09.009.

Liu, Z. *et al.* (2015) 'Specific Marker Expression and Cell State of Schwann Cells during Culture In Vitro', *PLoS ONE*. Public Library of Science, 10(4). doi: 10.1371/JOURNAL.PONE.0123278.

Long, C. *et al.* (2014) 'Prevention of muscular dystrophy in mice by

CRISPR/Cas9-mediated editing of germline DNA', *Science (New York, N.Y.)*. Science, 345(6201), pp. 1184–1188. doi: 10.1126/SCIENCE.1254445.

Long, C. *et al.* (2016) 'Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy', *Science (New York, N.Y.)*. NIH Public Access, 351(6271), p. 400. doi: 10.1126/SCIENCE.AAD5725.

Loperfido, M. *et al.* (2015) 'Pluripotent Stem Cells for Gene Therapy of Degenerative Muscle Diseases.', *Current gene therapy*, 15(4), pp. 364–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26122099 (Accessed: 10 July 2019).

Lord, C. J. and Ashworth, A. (2017) 'PARP inhibitors: Synthetic lethality in the clinic', *Science*. American Association for the Advancement of Science, 355(6330), pp. 1152–1158. doi: 10.1126/SCIENCE.AAM7344/ASSET/B1C2C000-4AAF-41D9-A934-CD4097D85A2A/ASSETS/GRAPHIC/355_1152_F2.JPEG.

Lorenz, P. *et al.* (2000) 'Nucleocytoplasmic shuttling: a novel in vivo property of antisense phosphorothioate oligodeoxynucleotides', *Nucleic Acids Research*. Oxford University Press, 28(2), p. 582. doi: 10.1093/NAR/28.2.582.

Love, D. R. *et al.* (1989) 'An autosomal transcript in skeletal muscle with homology to dystrophin', *Nature*. Nature, 339(6219), pp. 55–58. doi: 10.1038/339055a0.

Madden, L. *et al.* (2015) 'Bioengineered human myobundles mimic clinical responses of skeletal muscle to drugs', *eLife*. eLife Sciences Publications Ltd, 2015(4). doi: 10.7554/eLife.04885.

Maffioletti, S. M. *et al.* (2015) 'Efficient derivation and inducible differentiation of expandable skeletal myogenic cells from human ES and patient-specific iPS cells', *Nature Protocols*, 10(7), pp. 941–958. doi: 10.1038/nprot.2015.057.

Maffioletti, S. M. *et al.* (2018) 'Three-Dimensional Human iPSC-Derived Artificial Skeletal Muscles Model Muscular Dystrophies and Enable Multilineage Tissue Engineering', *Cell Reports*, 23(3), pp. 899–908. doi: 10.1016/j.celrep.2018.03.091.

Magli, A. *et al.* (2017) 'PAX7 Targets, CD54, Integrin α9β1, and SDC2, Allow Isolation of Human ESC/iPSC-Derived Myogenic Progenitors', *Cell Reports*, 19(13), pp. 2867–2877. doi: 10.1016/j.celrep.2017.06.005.

Makarova, K. S. *et al.* (2015) 'An updated evolutionary classification of CRISPR–Cas systems', *Nature Reviews Microbiology 2015 13:11*. Nature Publishing Group, 13(11), pp. 722–736. doi: 10.1038/nrmicro3569.

Mali, P. *et al.* (2013) 'RNA-guided human genome engineering via Cas9', *Science (New York, N.Y.)*. Science, 339(6121), pp. 823–826. doi: 10.1126/SCIENCE.1232033.

Mamchaoui, K. *et al.* (2011) ‘Immortalized pathological human myoblasts: Towards a universal tool for the study of neuromuscular disorders’, *Skeletal Muscle*. BioMed Central, 1(1), pp. 1–11. doi: 10.1186/2044-5040-1-34/FIGURES/4.

Man, N. T. and Morris, G. E. (1993) ‘Use of epitope libraries to identify exon-specific monoclonal antibodies for characterization of altered dystrophins in muscular dystrophy.’, *American Journal of Human Genetics*. Elsevier, 52(6), p. 1057. Available at: /pmc/articles/PMC1682265/?report=abstract (Accessed: 23 August 2022).

Mani, M. *et al.* (2005) ‘Binding of two zinc finger nuclease monomers to two specific sites is required for effective double-strand DNA cleavage’, *Biochemical and biophysical research communications*. NIH Public Access, 334(4), p. 1191. doi: 10.1016/J.BBRC.2005.07.021.

Mann, C. J. *et al.* (2011) ‘Aberrant repair and fibrosis development in skeletal muscle’, *Skeletal Muscle*, 1(1), p. 21. doi: 10.1186/2044-5040-1-21.

Manzur, A. Y., Kinali, M. and Muntoni, F. (2008) ‘Update on the management of Duchenne muscular dystrophy’, *Archives of Disease in Childhood*, 93(11), pp. 986–990. doi: 10.1136/adc.2007.118141.

Marczenke, M. *et al.* (2017) ‘Cardiac subtype-specific modeling of Kv1.5 ion channel deficiency using human pluripotent stem cells’, *Frontiers in Physiology*. Frontiers Media S.A., 8(JUL), p. 469. doi: 10.3389/FPHYS.2017.00469/BIBTEX.

Mareschal, S. *et al.* (2021) ‘Challenging conventional karyotyping by next-generation karyotyping in 281 intensively treated patients with AML’, *Blood Advances*. The American Society of Hematology, 5(4), p. 1003. doi: 10.1182/BLOODADVANCES.2020002517.

Marraffini, L. A. and Sontheimer, E. J. (2008) ‘CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA’, *Science (New York, N.Y.)*. Science, 322(5909), pp. 1843–1845. doi: 10.1126/SCIENCE.1165771.

Martin, G. R. (1981) ‘Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.’, *Proceedings of the National Academy of Sciences*, 78(12), pp. 7634–7638. doi: 10.1073/pnas.78.12.7634.

Massouridès, E. *et al.* (2015) ‘Dp412e: A novel human embryonic dystrophin isoform induced by BMP4 in early differentiated cells’, *Skeletal Muscle*, 5(1). doi: 10.1186/s13395-015-0062-6.

Matthews, E. *et al.* (2016) ‘Corticosteroids for the treatment of Duchenne muscular dystrophy’, *Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD003725.pub4.

Mauro, A. (1961) ‘Satellite cell of skeletal muscle fibers.’, *The Journal of biophysical and biochemical cytology*. Rockefeller University Press, 9(2), pp.

493–495. doi: 10.1083/jcb.9.2.493.

Mazaleyrat, K. *et al.* (2020) ‘Multilineage Differentiation for Formation of Innervated Skeletal Muscle Fibers from Healthy and Diseased Human Pluripotent Stem Cells’, *Cells*. Multidisciplinary Digital Publishing Institute (MDPI), 9(6). doi: 10.3390/CELLS9061531.

McClearn, D. and Noden, D. M. (1988) ‘Ontogeny of architectural complexity in embryonic quail visceral arch muscles’, *American Journal of Anatomy*, 183(4), pp. 277–293. doi: 10.1002/aja.1001830402.

McDonald, C. M. *et al.* (2017) ‘Ataluren in patients with nonsense mutation Duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial’, *The Lancet*, 390(10101), pp. 1489–1498. doi: 10.1016/S0140-6736(17)31611-2.

McGreevy, J. W. *et al.* (2015) ‘Animal models of Duchenne muscular dystrophy: from basic mechanisms to gene therapy.’, *Disease models & mechanisms*. Company of Biologists, 8(3), pp. 195–213. doi: 10.1242/dmm.018424.

Mendell, J. R. *et al.* (2010) ‘Dystrophin Immunity in Duchenne’s Muscular Dystrophy’, *New England Journal of Medicine*, 363(15), pp. 1429–1437. doi: 10.1056/NEJMoa1000228.

Mendell, J. R. *et al.* (2020) ‘Assessment of Systemic Delivery of rAAVrh74.MHCK7.micro-dystrophin in Children With Duchenne Muscular Dystrophy: A Nonrandomized Controlled Trial’, *JAMA Neurology*. American Medical Association, 77(9), pp. 1122–1131. doi: 10.1001/JAMANEUROL.2020.1484.

Mercado, M. L. *et al.* (2006) ‘Biglycan regulates the expression and sarcolemmal localization of dystrobrevin, syntrophin, and nNOS’, *The FASEB Journal*. Wiley, 20(10), pp. 1724–1726. doi: 10.1096/fj.05-5124fje.

Mercuri, E. and Muntoni, F. (2013) ‘Muscular dystrophies’, *The Lancet*. Elsevier Ltd, 381(9869), pp. 845–860. doi: 10.1016/S0140-6736(12)61897-2.

Miike, T. *et al.* (1987) ‘Vascular endothelial cell injury and platelet embolism in Duchenne muscular dystrophy at the preclinical stage’, *Journal of the neurological sciences*. J Neurol Sci, 82(1–3), pp. 67–80. doi: 10.1016/0022-510X(87)90007-4.

Millette, K. and Georgia, S. (2017) ‘Gene Editing and Human Pluripotent Stem Cells: Tools for Advancing Diabetes Disease Modeling and Beta-Cell Development’, *Current diabetes reports*. Curr Diab Rep, 17(11). doi: 10.1007/S11892-017-0947-3.

Mills, R. J. *et al.* (2019) ‘Development of a human skeletal micro muscle platform with pacing capabilities’, *Biomaterials*. Elsevier Ltd, 198, pp. 217–227. doi: 10.1016/j.biomaterials.2018.11.030.

Min, Y. L., Bassel-Duby, R. and Olson, E. N. (2019) ‘CRISPR Correction

of Duchenne Muscular Dystrophy', *Annual review of medicine*. Annu Rev Med, 70, pp. 239–255. doi: 10.1146/ANNUREV-MED-081117-010451.

Moizard, M. P. *et al.* (2000) 'Severe cognitive impairment in DMD: obvious clinical indication for Dp71 isoform point mutation screening', *European journal of human genetics : EJHG*. Eur J Hum Genet, 8(7), pp. 552–556. doi: 10.1038/SJ.EJHG.5200488.

Mojica, F. J. M. *et al.* (1995) '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', *Molecular microbiology*. Mol Microbiol, 17(1), pp. 85–93. doi: 10.1111/J.1365-2958.1995.MMI_17010085.X.

Mojica, F. J. M. *et al.* (2000) 'Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria', *Molecular microbiology*. Mol Microbiol, 36(1), pp. 244–246. doi: 10.1046/J.1365-2958.2000.01838.X.

Mojica, F. J. M. *et al.* (2005) 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements', *Journal of molecular evolution*. J Mol Evol, 60(2), pp. 174–182. doi: 10.1007/S00239-004-0046-3.

Mojica, F. J. M., Juez, G. and Rodriguez-Valera, F. (1993) 'Transcription at different salinities of Haloferax mediterranei sequences adjacent to partially modified PstI sites', *Molecular microbiology*. Mol Microbiol, 9(3), pp. 613–621. doi: 10.1111/J.1365-2958.1993.TB01721.X.

Monaco, A. P. *et al.* (1988) 'An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus', *Genomics*. Genomics, 2(1), pp. 90–95. doi: 10.1016/0888-7543(88)90113-9.

Monroy, J. A. *et al.* (2012) 'What is the role of titin in active muscle?', *Exercise and sport sciences reviews*. Exerc Sport Sci Rev, 40(2), pp. 73–78. doi: 10.1097/JES.0B013E31824580C6.

Moon, H. *et al.* (2017) 'Effects of PTCs on nonsense-mediated mRNA decay are dependent on PTC location', *Oncology letters*. Oncol Lett, 13(3), pp. 1944–1948. doi: 10.3892/OL.2017.5627.

Moretti, A. *et al.* (2020) 'Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy', *Nature medicine*. Nat Med, 26(2), pp. 207–214. doi: 10.1038/S41591-019-0738-2.

Morgan, J. E. and Partridge, T. A. (1992) 'Challenges: Cell transplantation and gene therapy in muscular dystrophy', *BioEssays*. John Wiley & Sons, Ltd, 14(9), pp. 641–645. doi: 10.1002/bies.950140913.

Morgan, M. A. *et al.* (2020) 'Use of Cell and Genome Modification Technologies to Generate Improved "Off-the-Shelf" CAR T and CAR NK Cells', *Frontiers in Immunology*. Frontiers Media S.A., 11, p. 1965. doi:

10.3389/FIMMU.2020.01965/BIBTEX.

Morizane, R. *et al.* (2015) 'Nephron organoids derived from human pluripotent stem cells model kidney development and injury', *Nature Biotechnology*. Nature Publishing Group, 33(11), pp. 1193–1200. doi: 10.1038/nbt.3392.

Müller, F. J. *et al.* (2011) 'A bioinformatic assay for pluripotency in human cells', *Nature methods*. NIH Public Access, 8(4), p. 315. doi: 10.1038/NMETH.1580.

Muñoz, I. G. *et al.* (2011) 'Molecular basis of engineered meganuclease targeting of the endogenous human RAG1 locus', *Nucleic Acids Research*. Oxford University Press, 39(2), p. 729. doi: 10.1093/NAR/GKQ801.

Muntoni, F., Torelli, S. and Ferlini, A. (2003) 'Dystrophin and mutations: one gene, several proteins, multiple phenotypes', *The Lancet Neurology*, 2(12), pp. 731–740. doi: 10.1016/S1474-4422(03)00585-4.

Muraine, L. *et al.* (2020) 'Transduction Efficiency of Adeno-Associated Virus Serotypes After Local Injection in Mouse and Human Skeletal Muscle', *Human gene therapy*. Hum Gene Ther, 31(3–4), pp. 233–240. doi: 10.1089/HUM.2019.173.

Murakami, N. *et al.* (2017) 'Proteasome impairment in neural cells derived from HMSN-P patient iPSCs', *Molecular Brain*. BioMed Central Ltd., 10(1), pp. 1–10. doi: 10.1186/S13041-017-0286-Y/FIGURES/3.

Murphy, M. and Kardon, G. (2011) 'Origin of Vertebrate Limb Muscle', in *Current topics in developmental biology*, pp. 1–32. doi: 10.1016/B978-0-12-385940-2.00001-2.

Mussolino, C. and Cathomen, T. (2012) 'TALE nucleases: tailored genome engineering made easy', *Current opinion in biotechnology*. Curr Opin Biotechnol, 23(5), pp. 644–650. doi: 10.1016/J.COPBIO.2012.01.013.

Nakamura, K. *et al.* (2015) 'Generation of muscular dystrophy model rats with a CRISPR/Cas system', *Scientific Reports*. Nature Publishing Group, 4(1), pp. 1–6. doi: 10.1038/srep05635.

Nakata, A., Amemura, M. and Makino, K. (1989) 'Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome', *Journal of bacteriology*. J Bacteriol, 171(6), pp. 3553–3556. doi: 10.1128/JB.171.6.3553-3556.1989.

Nelson, C. E. *et al.* (2016) 'In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy', *Science (New York, N.Y.)*. Science, 351(6271), pp. 403–407. doi: 10.1126/SCIENCE.AAD5143.

Neri, M. *et al.* (2007) 'Dystrophin levels as low as 30% are sufficient to avoid muscular dystrophy in the human', *Neuromuscular disorders : NMD*. Neuromuscul Disord, 17(11–12), pp. 913–918. doi: 10.1016/J.NMD.2007.07.005.

Nguyen, Q. and Yokota, T. (2017) 'Immortalized muscle cell model to test the exon skipping efficacy for duchenne muscular dystrophy', *Journal of Personalized Medicine*. doi: 10.3390/jpm7040013.

Nguyen, T. M. *et al.* (1991) 'Localization of the DMDL gene-encoded dystrophin-related protein using a panel of nineteen monoclonal antibodies: presence at neuromuscular junctions, in the sarcolemma of dystrophic skeletal muscle, in vascular and other smooth muscles, and in proliferating brain cell lines.', *Journal of Cell Biology*. The Rockefeller University Press, 115(6), pp. 1695–1700. doi: 10.1083/JCB.115.6.1695.

Nguyen, T. M. *et al.* (1992) 'Monoclonal antibodies for dystrophin analysis. Epitope mapping and improved binding to SDS-treated muscle sections', *Biochemical Journal*. Portland Press, 288(2), pp. 663–668. doi: 10.1042/BJ2880663.

Nishimasu, H. *et al.* (2014) 'Crystal structure of Cas9 in complex with guide RNA and target DNA', *Cell*. Elsevier B.V., 156(5), pp. 935–949. doi: 10.1016/J.CELL.2014.02.001/ATTACHMENT/EA7E5665-5D2D-420A-83FC-9EDE1785C477/MMC1.PDF.

Noden, D. M. and Francis-West, P. (2006) 'The differentiation and morphogenesis of craniofacial muscles', *Developmental dynamics : an official publication of the American Association of Anatomists*. Dev Dyn, 235(5), pp. 1194–1218. doi: 10.1002/DVDY.20697.

North, K. N. *et al.* (2014) 'Approach to the diagnosis of congenital myopathies', *Neuromuscular disorders : NMD*. Neuromuscul Disord, 24(2), pp. 97–116. doi: 10.1016/J.NMD.2013.11.003.

Osaki, T., Uzel, S. G. M. and Kamm, R. D. (2018) 'Microphysiological 3D model of amyotrophic lateral sclerosis (ALS) from human iPS-derived muscle cells and optogenetic motor neurons', *Science Advances*, 4, p. 5847. Available at: http://advances.sciencemag.org/ (Accessed: 11 February 2021).

Ostrovidov, S. *et al.* (2019) '3D Bioprinting in Skeletal Muscle Tissue Engineering', *Small*. John Wiley & Sons, Ltd, 15(24), p. 1805530. doi: 10.1002/smll.201805530.

Otto, A., Schmidt, C. and Patel, K. (2006) 'Pax3 and Pax7 expression and regulation in the avian embryo.', *Anatomy and embryology*, 211(4), pp. 293–310. doi: 10.1007/s00429-006-0083-3.

De Paepe, B. and De Bleecker, J. L. (2013) 'Cytokines and chemokines as regulators of skeletal muscle inflammation: presenting the case of Duchenne muscular dystrophy.', *Mediators of inflammation*. Hindawi, 2013, p. 540370. doi: 10.1155/2013/540370.

Palermo, G. *et al.* (2018) 'Key role of the REC lobe during CRISPR–Cas9 activation by "sensing", "regulating", and "locking" the catalytic HNH domain', *Quarterly reviews of biophysics*. NIH Public Access, 51. doi: 10.1017/S0033583518000070.

Palladino, M. *et al.* (2013) ‘Angiogenic impairment of the vascular endothelium: a novel mechanism and potential therapeutic target in muscular dystrophy’, *Arteriosclerosis, thrombosis, and vascular biology*. Arterioscler Thromb Vasc Biol, 33(12), pp. 2867–2876. doi: 10.1161/ATVBAHA.112.301172.

Pâques, F. and Duchateau, P. (2007) ‘Meganucleases and DNA Double-Strand Break-Induced Recombination: Perspectives for Gene Therapy’, *Current Gene Therapy*, 7, pp. 49–66.

Partridge, T. A. *et al.* (1989) ‘Conversion of mdx myofibres from dystrophin-negative to -positive by injection of normal myoblasts’, *Nature*. Nature, 337(6203), pp. 176–179. doi: 10.1038/337176a0.

Pasca, A. M. *et al.* (2015) ‘Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture’, *Nature Methods*. Nature Publishing Group, 12(7), pp. 671–678. doi: 10.1038/nmeth.3415.

Pavletich, N. P. and Pabo, C. O. (1991) ‘Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A’, *Science (New York, N.Y.)*. Science, 252(5007), pp. 809–817. doi: 10.1126/SCIENCE.2028256.

Pawelczak, K. S. *et al.* (2018) ‘Modulating DNA Repair Pathways to Improve Precision Genome Engineering’, *ACS chemical biology*. ACS Chem Biol, 13(2), pp. 389–396. doi: 10.1021/ACSCHEMBIO.7B00777.

Perez, E. E. *et al.* (2008) ‘Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases’, *Nature biotechnology*. Nat Biotechnol, 26(7), pp. 808–816. doi: 10.1038/NBT1410.

Périé, S. *et al.* (2014) ‘Autologous myoblast transplantation for oculopharyngeal muscular dystrophy: A phase I/IIa clinical study’, *Molecular Therapy*. Nature Publishing Group, 22(1), pp. 219–225. doi: 10.1038/mt.2013.155.

Petkova, M. V. *et al.* (2016) ‘Characterization of a DmdEGFPreporter mouse as a tool to investigate dystrophin expression’, *Skeletal Muscle*, 6(1), p. 25. doi: 10.1186/S13395-016-0095-5.

Petrof, B. J. *et al.* (1993) ‘Dystrophin protects the sarcolemma from stresses developed during muscle contraction.’, *Proceedings of the National Academy of Sciences*, 90(8), pp. 3710–3714. doi: 10.1073/pnas.90.8.3710.

Popplewell, L. J. *et al.* (2009) ‘Design of Phosphorodiamidate Morpholino Oligomers (PMOs) for the Induction of Exon Skipping of the Human DMD Gene’, *Molecular Therapy: the Journal of the American Society of Gene Therapy*. American Society of Gene & Cell Therapy, 17(3), p. 554. doi: 10.1038/MT.2008.287.

Porteus, M. (2010) ‘Creating Zinc Finger Nucleases to Manipulate the Genome in a Site-Specific Manner Using a Modular-Assembly Approach’, *Cold Spring Harbor Protocols*. Cold Spring Harbor Laboratory Press, 2010(12), p. pdb.top93. doi: 10.1101/PDB.TOP93.

Pourcel, C., Salvignol, G. and Vergnaud, G. (2005) ‘CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies’, *Microbiology (Reading, England)*. Microbiology (Reading), 151(Pt 3), pp. 653–663. doi: 10.1099/MIC.0.27437-0.

Pul, Ü. *et al.* (2010) ‘Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS’, *Molecular microbiology*. Mol Microbiol, 75(6), pp. 1495–1512. doi: 10.1111/J.1365-2958.2010.07073.X.

Radhakrishnan, S. K. and Lees-Miller, S. P. (2017) ‘DNA requirements for interaction of the C-terminal region of Ku80 with the DNA-dependent protein kinase catalytic subunit (DNA-PKcs)’, *DNA repair*. DNA Repair (Amst), 57, pp. 17–28. doi: 10.1016/J.DNAREP.2017.06.001.

Ramirez, C. L. *et al.* (2008) ‘Unexpected failure rates for modular assembly of engineered zinc fingers’, *Nature methods*. Nat Methods, 5(5), pp. 374–375. doi: 10.1038/NMETH0508-374.

Ran, Y. *et al.* (2018) ‘Zinc finger nuclease-mediated precision genome editing of an endogenous gene in hexaploid bread wheat (Triticum aestivum) using a DNA repair template’, *Plant biotechnology journal*. Plant Biotechnol J, 16(12), pp. 2088–2101. doi: 10.1111/PBI.12941.

Rando, T. A. (2001) ‘The dystrophin-glycoprotein complex, cellular signaling, and the regulation of cell survival in the muscular dystrophies.’, *Muscle & nerve*, 24(12), pp. 1575–94. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11745966 (Accessed: 10 July 2019).

Rani, A. Q. M. *et al.* (2020) ‘Intronic Alternative Polyadenylation in the Middle of the DMD Gene Produces Half-Size N-Terminal Dystrophin with a Potential Implication of ECG Abnormalities of DMD Patients’, *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI), 21(10), p. 3555. doi: 10.3390/IJMS21103555.

Rao, L. *et al.* (2018) ‘Engineering human pluripotent stem cells into a functional skeletal muscle tissue’, *Nature Communications*. Nature Publishing Group, 9(1), p. 126. doi: 10.1038/s41467-017-02636-4.

Relaix, F. *et al.* (2005) ‘A Pax3/Pax7-dependent population of skeletal muscle progenitor cells’, *Nature*. Nature, 435(7044), pp. 948–953. doi: 10.1038/NATURE03594.

Relaix, F. *et al.* (2006) ‘Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells.’, *The Journal of cell biology*. Rockefeller University Press, 172(1), pp. 91–102. doi: 10.1083/jcb.200508044.

Ribeiro, S. *et al.* (2012) ‘Plasmid DNA size does affect nonviral gene delivery efficiency in stem cells’, *Cellular reprogramming*. Cell Reprogram, 14(2), pp. 130–137. doi: 10.1089/CELL.2011.0093.

Ricotti, V. *et al.* (2016) ‘Neurodevelopmental, emotional, and behavioural

problems in Duchenne muscular dystrophy in relation to underlying dystrophin gene mutations', *Developmental medicine and child neurology*. Dev Med Child Neurol, 58(1), pp. 77–84. doi: 10.1111/DMCN.12922.

Ricotti, V., Roberts, R. G. and Muntoni, F. (2011) 'Dystrophin and the brain', *Developmental medicine and child neurology*. Dev Med Child Neurol, 53(1), pp. 12–12. doi: 10.1111/J.1469-8749.2010.03836.X.

Rivera, V. M. *et al.* (2005) 'Long-term pharmacologically regulated expression of erythropoietin in primates following AAV-mediated gene transfer', *Blood*. Blood, 105(4), pp. 1424–1430. doi: 10.1182/BLOOD-2004-06-2501.

Rodino-Klapac, L. R. *et al.* (2010) 'Persistent Expression of FLAG-tagged Micro dystrophin in Nonhuman Primates Following Intramuscular and Vascular Delivery', *Molecular Therapy*, 18(1), pp. 109–117. doi: 10.1038/mt.2009.254.

Roper, J. *et al.* (2018) 'Colonoscopy-based colorectal cancer modeling in mice with CRISPR-Cas9 genome editing and organoid transplantation', *Nature protocols*. Nat Protoc, 13(2), pp. 217–234. doi: 10.1038/NPROT.2017.136.

Rosen, L. E. *et al.* (2006) 'Homing endonuclease I-CreI derivatives with novel DNA target specificities', *Nucleic Acids Research*. Oxford University Press, 34(17), p. 4791. doi: 10.1093/NAR/GKL645.

Rousseau, E. *et al.* (1988) 'Activation of the Ca2+ release channel of skeletal muscle sarcoplasmic reticulum by caffeine and related compounds', *Archives of biochemistry and biophysics*. Arch Biochem Biophys, 267(1), pp. 75–86. doi: 10.1016/0003-9861(88)90010-0.

Rui, Y., Wilson, D. R. and Green, J. J. (2019) 'Non-Viral Delivery To Enable Genome Editing.', *Trends in biotechnology*. Elsevier, 37(3), pp. 281–293. doi: 10.1016/j.tibtech.2018.08.010.

Russo, F. B. *et al.* (2015) 'Induced pluripotent stem cells for modeling neurological disorders', *World Journal of Transplantation*, 5(4), p. 209. doi: 10.5500/wjt.v5.i4.209.

Rybakova, I. N. *et al.* (2006) 'Dystrophin and utrophin bind actin through distinct modes of contact', *Journal of Biological Chemistry*. J Biol Chem, 281(15), pp. 9996–10001. doi: 10.1074/jbc.M513121200.

Sacco, A. *et al.* (2010) 'Short Telomeres and Stem Cell Exhaustion Model Duchenne Muscular Dystrophy in mdx/mTR Mice', *Cell*, 143(7), pp. 1059–1071. doi: 10.1016/j.cell.2010.11.039.

Santiago, Y. *et al.* (2008) 'Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases', *Proceedings of the National Academy of Sciences of the United States of America*, 105(15), pp. 5809–5814. doi: 10.1073/PNAS.0800940105/SUPPL_FILE/0800940105SI.PDF.

Sartori, A. A. *et al.* (2007) 'Human CtIP promotes DNA end resection', *Nature*. Nature, 450(7169), pp. 509–514. doi: 10.1038/NATURE06337.

Schiaffino, S. *et al.* (2015) ‘Developmental myosins: Expression patterns and functional significance’, *Skeletal Muscle*. BioMed Central Ltd., 5(1), pp. 1–14. doi: 10.1186/S13395-015-0046-6/FIGURES/1.

Schiaffino, S. and Reggiani, C. (2011) ‘Fiber types in mammalian skeletal muscles’, *Physiological reviews*. Physiol Rev, 91(4), pp. 1447–1531. doi: 10.1152/PHYSREV.00031.2010.

Schienda, J. *et al.* (2006) ‘Somitic origin of limb muscle satellite and side population cells’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 103(4), p. 945. doi: 10.1073/PNAS.0510164103.

Schmalbruch, H. and Hellhammer, U. (1976) ‘The number of satellite cells in normal human muscle’, *The Anatomical Record*. John Wiley & Sons, Ltd, 185(3), pp. 279–287. doi: 10.1002/ar.1091850303.

Schmittgen, T. D. and Livak, K. J. (2008) ‘Analyzing real-time PCR data by the comparative CT method’, *Nature Protocols 2008 3:6*. Nature Publishing Group, 3(6), pp. 1101–1108. doi: 10.1038/nprot.2008.73.

Schwander, M. *et al.* (2003) ‘Beta1 integrins regulate myoblast fusion and sarcomere assembly’, *Developmental cell*. Dev Cell, 4(5), pp. 673–685. doi: 10.1016/S1534-5807(03)00118-7.

Scott, J. M. *et al.* (2002) ‘Viral vectors for gene transfer of micro-, mini-, or full-length dystrophin.’, *Neuromuscular disorders : NMD*, 12 Suppl 1, pp. S23-9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12206791 (Accessed: 10 July 2019).

Seale, P. *et al.* (2000) ‘Pax7 is required for the specification of myogenic satellite cells’, *Cell*. Cell, 102(6), pp. 777–786. doi: 10.1016/S0092-8674(00)00066-0.

Seligman, L. M. *et al.* (2002) ‘Mutations altering the cleavage specificity of a homing endonuclease’, *Nucleic Acids Research*. Oxford University Press, 30(17), p. 3870. doi: 10.1093/NAR/GKF495.

Shah, S. A. *et al.* (2013) ‘Protospacer recognition motifs: mixed identities and functional diversity’, *RNA biology*. RNA Biol, 10(5), pp. 891–899. doi: 10.4161/RNA.23764.

Shao, Y. *et al.* (2014) ‘CRISPR/Cas-mediated genome editing in the rat via direct injection of one-cell embryos’, *Nature Protocols 2014 9:10*. Nature Publishing Group, 9(10), pp. 2493–2512. doi: 10.1038/nprot.2014.171.

Shelton, M. *et al.* (2014) ‘Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells.’, *Stem cell reports*. Elsevier, 3(3), pp. 516–29. doi: 10.1016/j.stemcr.2014.07.001.

Shen, B. *et al.* (2013) ‘Generation of gene-modified mice via Cas9/RNA-mediated gene targeting’, *Cell Research*. Nature Publishing Group, 23(5), p. 720. doi: 10.1038/CR.2013.46.

Shen, X. and Corey, D. R. (2018) ‘Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs’, *Nucleic Acids Research*. Oxford University Press, 46(4), p. 1584. doi: 10.1093/NAR/GKX1239.

Sheng, L. *et al.* (2020) ‘Comparison of the efficacy of MOE and PMO modifications of systemic antisense oligonucleotides in a severe SMA mouse model’, *Nucleic acids research*. Nucleic Acids Res, 48(6), pp. 2853–2865. doi: 10.1093/NAR/GKAA126.

Sheridan, S. D., Surampudi, V. and Rao, R. R. (2012) ‘Analysis of embryoid bodies derived from human induced pluripotent stem cells as a means to assess pluripotency’, *Stem cells international*. Stem Cells Int, 2012. doi: 10.1155/2012/738910.

Sherwood, L. (no date) *Human physiology : from cells to systems*.

Shibata, A. *et al.* (2011) ‘Factors determining DNA double-strand break repair pathway choice in G2 phase’, *The EMBO Journal*. John Wiley & Sons, Ltd, 30(6), pp. 1079–1092. doi: 10.1038/EMBOJ.2011.27.

Shimatsu, Y. *et al.* (2005) ‘Major clinical and histopathological characteristics of canine X-linked muscular dystrophy in Japan, CXMDJ.’, *Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology*, 24(2), pp. 145–54. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16550932 (Accessed: 10 July 2019).

Shivji, M. K. K. *et al.* (2009) ‘The BRC repeats of human BRCA2 differentially regulate RAD51 binding on single- versus double-stranded DNA to stimulate strand exchange’, *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 106(32), pp. 13254–13259. doi: 10.1073/PNAS.0906208106.

Shoji, E. *et al.* (2015) ‘Early pathogenesis of Duchenne muscular dystrophy modelled in patient-derived human induced pluripotent stem cells’, *Scientific Reports*. Nature Publishing Group, 5(1), pp. 1–13. doi: 10.1038/srep12831.

Sicinski, P. *et al.* (1989) ‘The molecular basis of muscular dystrophy in the mdx mouse: a point mutation’, *Science*, 244(4912), pp. 1578–1580. doi: 10.1126/science.2662404.

Siegel, A. L., Kuhlmann, P. K. and Cornelison, D. (2011) ‘Muscle satellite cell proliferation and association: new insights from myofiber time-lapse imaging’, *Skeletal Muscle*. BioMed Central, 1(1), p. 7. doi: 10.1186/2044-5040-1-7.

Silva, G. *et al.* (2011) ‘Meganucleases and Other Tools for Targeted Genome Engineering: Perspectives and Challenges for Gene Therapy’, *Current Gene Therapy*. Bentham Science Publishers, 11(1), p. 11. doi: 10.2174/156652311794520111.

Skeletal muscle differentiation kit (with control myoblasts), SKM-KIT | AMSBIO (no date). Available at: https://www.amsbio.com/skeletal-muscle-

differentiation-kit-with-control-myoblasts-skm-kit (Accessed: 25 February 2021).

Skuk, D. *et al.* (2007) ‘First test of a “high-density injection” protocol for myogenic cell transplantation throughout large volumes of muscles in a Duchenne muscular dystrophy patient: eighteen months follow-up’, *Neuromuscular Disorders*, 17(1), pp. 38–46. doi: 10.1016/j.nmd.2006.10.003.

Skuk, D. and Tremblay, J. P. (2015) ‘Cell therapy in muscular dystrophies: many promises in mice and dogs, few facts in patients’, *Expert Opinion on Biological Therapy*, 15(9), pp. 1307–1319. doi: 10.1517/14712598.2015.1057564.

Slaymaker, I. M. and Gaudelli, N. M. (2021) ‘Engineering Cas9 for human genome editing’, *Current Opinion in Structural Biology*. Elsevier Current Trends, 69, pp. 86–98. doi: 10.1016/J.SBI.2021.03.004.

Smith, J. *et al.* (2006) ‘A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences’, *Nucleic Acids Research*. Oxford University Press, 34(22), p. e149. doi: 10.1093/NAR/GKL720.

Smith, K. (2011) ‘Feline muscular dystrophy: parallels between cats and people’, *Veterinary Record*, 168(19), pp. 507–508. doi: 10.1136/vr.d2940.

Soblechero-Martín, P. *et al.* (2021) ‘Duchenne muscular dystrophy cell culture models created by CRISPR/Cas9 gene editing and their application in drug screening’, *Scientific Reports 2021 11:1*. Nature Publishing Group, 11(1), pp. 1–14. doi: 10.1038/s41598-021-97730-5.

Van Soolingen, D. *et al.* (1993) ‘Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of Mycobacterium tuberculosis.’, *Journal of Clinical Microbiology*. American Society for Microbiology (ASM), 31(8), p. 1987. doi: 10.1128/jcm.31.8.1987-1995.1993.

Spence, J. R. *et al.* (2011) ‘Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro’, *Nature*. Nature Publishing Group, 470(7332), pp. 105–110. doi: 10.1038/nature09691.

Stacpoole, S. R. *et al.* (2011) ‘Efficient derivation of NPCs, spinal motor neurons and midbrain dopaminergic neurons from hESCs at 3% oxygen’, *Nature protocols*. Nat Protoc, 6(8), pp. 1229–1240. doi: 10.1038/NPROT.2011.380.

Steele-Stallard, H. B. *et al.* (2018) ‘Modeling Skeletal Muscle Laminopathies Using Human Induced Pluripotent Stem Cells Carrying Pathogenic LMNA Mutations’, *Frontiers in Physiology*. Frontiers Media S.A., 9(OCT), p. 1332. doi: 10.3389/fphys.2018.01332.

Sternberg, S. H. *et al.* (2014) ‘DNA interrogation by the CRISPR RNA-guided endonuclease Cas9’, *Nature*. Nature, 507(7490), pp. 62–67. doi: 10.1038/NATURE13011.

Stracker, T. H. and Petrini, J. H. J. (2011) ‘The MRE11 complex: starting from the ends’, *Nature reviews. Molecular cell biology*. Nat Rev Mol Cell Biol, 12(2), pp. 90–103. doi: 10.1038/NRM3047.

Studer, L., Vera, E. and Cornacchia, D. (2015) ‘Programming and Reprogramming Cellular Age in the Era of Induced Pluripotency’, *Cell stem cell*. Cell Stem Cell, 16(6), pp. 591–600. doi: 10.1016/J.STEM.2015.05.004.

Summerton, J. *et al.* (1997) ‘Morpholino and phosphorothioate antisense oligomers compared in cell-free and in-cell systems’, *Antisense & nucleic acid drug development*. Antisense Nucleic Acid Drug Dev, 7(2), pp. 63–70. doi: 10.1089/OLI.1.1997.7.63.

Summerton, J. E. (2005) ‘Endo-Porter: a novel reagent for safe, effective delivery of substances into cells’, *Annals of the New York Academy of Sciences*. Ann N Y Acad Sci, 1058, pp. 62–75. doi: 10.1196/ANNALS.1359.012.

Summit Therapeutics (2020) *Proof of Concept Study to Assess Activity and Safety of SMT C1100 (Ezutromid) in Boys With Duchenne Muscular Dystrophy (DMD) - Full Text View - ClinicalTrials.gov*, *ClinicalTrials.gov*. Available at: https://www.clinicaltrials.gov/ct2/show/NCT02858362 (Accessed: 23 February 2021).

Szczelkun, M. D. *et al.* (2014) ‘Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes’, *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 111(27), pp. 9798–9803. doi: 10.1073/PNAS.1402597111/-/DCSUPPLEMENTAL/PNAS.1402597111.SAPP.PDF.

Tabebordbar, M. *et al.* (2016) ‘In vivo gene editing in dystrophic mouse muscle and muscle stem cells’, *Science*, 351(6271), pp. 407–411. doi: 10.1126/science.aad5177.

Tadayoni, R. *et al.* (2012) ‘Dystrophin Dp71: the smallest but multifunctional product of the Duchenne muscular dystrophy gene’, *Molecular neurobiology*. Mol Neurobiol, 45(1), pp. 43–60. doi: 10.1007/S12035-011-8218-9.

Taglietti, V. *et al.* (2022) ‘Duchenne muscular dystrophy trajectory in R-DMDdel52 preclinical rat model identifies COMP as biomarker of fibrosis’, *Acta neuropathologica communications*. Acta Neuropathol Commun, 10(1). doi: 10.1186/S40478-022-01355-2.

Takahashi, K. *et al.* (2007) ‘Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors’, *Cell*, 131(5), pp. 861–872. doi: 10.1016/j.cell.2007.11.019.

Takahashi, K. and Yamanaka, S. (2006) ‘Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors’, *Cell*, 126(4), pp. 663–676. doi: 10.1016/j.cell.2006.07.024.

Tanaka, A. *et al.* (2013) ‘Efficient and Reproducible Myogenic Differentiation from Human iPS Cells: Prospects for Modeling Miyoshi Myopathy In Vitro’, *PLoS ONE*. Edited by A. Asakura, 8(4), p. e61540. doi: 10.1371/journal.pone.0061540.

Al Tanoury, Z. *et al.* (2021) ‘Prednisolone rescues Duchenne muscular dystrophy phenotypes in human pluripotent stem cell-derived skeletal muscle in vitro’, *Proceedings of the National Academy of Sciences of the United States of America*, 118(28). doi: 10.1073/pnas.2022960118.

Tedesco, F. S. *et al.* (2012) ‘Transplantation of Genetically Corrected Human iPSC-Derived Progenitors in Mice with Limb-Girdle Muscular Dystrophy’, *Science Translational Medicine*, 4(140), pp. 140ra89-140ra89. doi: 10.1126/scitranslmed.3003541.

Tennyson, C. N., Klamut, H. J. and Worton, R. G. (1995) ‘The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced’, *Nature genetics*. Nat Genet, 9(2), pp. 184–190. doi: 10.1038/NG0295-184.

Thierry, A. and Dujon, B. (1992) ‘Nested chromosomal fragmentation in yeast using the meganuclease I-Sce I: a new method for physical mapping of eukaryotic genomes.’, *Nucleic Acids Research*. Oxford University Press, 20(21), p. 5625. doi: 10.1093/NAR/20.21.5625.

Thomas, N. L. and Williams, A. J. (2012) ‘Pharmacology of ryanodine receptors and Ca2+-induced Ca2+ release’, *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, 1(4), pp. 383–397. doi: 10.1002/WMTS.34.

Thomson, J. A. *et al.* (1998) ‘Embryonic Stem Cell Lines Derived from Human Blastocysts’, *Science*, 282(5391), pp. 1145–1147. doi: 10.1126/science.282.5391.1145.

Tinsley, J. M. *et al.* (2011) ‘Daily treatment with SMTC1100, a novel small molecule utrophin upregulator, dramatically reduces the dystrophic symptoms in the mdx mouse.’, *PloS one*. Public Library of Science, 6(5), p. e19189. doi: 10.1371/journal.pone.0019189.

Tinsley, J. M., Blake, D. J. and Davies, K. E. (1993) ‘Apo-dystrophin-3: a 2.2kb transcript from the DMD locus encoding the dystrophin glycoprotein binding site’. Oxford University Press Human Molecular Genetics, 2(5), pp. 521–524. Available at: https://academic.oup.com/hmg/article/2/5/521/636612 (Accessed: 4 June 2022).

TJ, L. *et al.* (2013) ‘Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia’, *The New England journal of medicine*. N Engl J Med, 368(22), pp. 2059–2074. doi: 10.1056/NEJMOA1301689.

Tomé, F. M. S. *et al.* (1994) ‘Expression of dystrophin-associated glycoproteins during human fetal muscle development: A preliminary immunocytochemical study’, *Neuromuscular Disorders*. Neuromuscul Disord, 4(4), pp. 343–348. doi: 10.1016/0960-8966(94)90070-1.

Tomimatsu, N. *et al.* (2014) ‘Phosphorylation of EXO1 by CDKs 1 and 2 regulates DNA end resection and repair pathway choice’, *Nature Communications 2014 5:1*. Nature Publishing Group, 5(1), pp. 1–10. doi: 10.1038/ncomms4561.

Tsai, S. Q. *et al.* (2014) 'GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases', *Nature Biotechnology 2014 33:2*. Nature Publishing Group, 33(2), pp. 187–197. doi: 10.1038/nbt.3117.

Tsai, S. Q. *et al.* (2017) 'CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets', *Nature methods*. Nat Methods, 14(6), pp. 607–614. doi: 10.1038/NMETH.4278.

Uusi-Mäkelä, M. I. E. *et al.* (2018) 'Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio)', *PLOS ONE*. Public Library of Science, 13(4), p. e0196238. doi: 10.1371/JOURNAL.PONE.0196238.

Vakulskas, C. A. *et al.* (2018) 'A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells', *Nature Medicine 2018 24:8*. Nature Publishing Group, 24(8), pp. 1216–1224. doi: 10.1038/s41591-018-0137-0.

Valdez, M. R. *et al.* (2000) 'Failure of Myf5 to support myogenic differentiation without myogenin, MyoD, and MRF4', *Developmental biology*. Dev Biol, 219(2), pp. 287–298. doi: 10.1006/DBIO.2000.9621.

Verhaart, I. E. C. and Aartsma-Rus, A. (2019) 'Therapeutic developments for Duchenne muscular dystrophy', *Nature Reviews Neurology 2019 15:7*. Nature Publishing Group, 15(7), pp. 373–386. doi: 10.1038/s41582-019-0203-3.

Verkuijl, S. A. and Rots, M. G. (2019) 'The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies', *Current opinion in biotechnology*. Curr Opin Biotechnol, 55, pp. 68–73. doi: 10.1016/J.COPBIO.2018.07.005.

Vila, O. F. *et al.* (2019) 'Quantification of human neuromuscular function through optogenetics', *Theranostics*, 9(5). doi: 10.7150/thno.25735.

Vilenchik, M. M. and Knudson, A. G. (2003) 'Endogenous DNA double-strand breaks: Production, fidelity of repair, and induction of cancer', *Proceedings of the National Academy of Sciences of the United States of America*, 100(22), pp. 12871–12876. doi: 10.1073/pnas.2135498100.

Vilquin, J. T. *et al.* (1994) 'Myoblast allotransplantation in mice: degree of success varies depending on the efficacy of various immunosuppressive treatments.', *Transplantation proceedings*, 26(6), pp. 3372–3. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7998179 (Accessed: 10 July 2019).

Waddington, S. N. *et al.* (2016) 'A Broad Overview and Review of CRISPR-Cas Technology and Stem Cells', *Current Stem Cell Reports*. Springer, 2(1), p. 9. doi: 10.1007/S40778-016-0037-5.

Wang, H. *et al.* (2013) 'One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering', *Cell*. Cell, 153(4), pp. 910–918. doi: 10.1016/J.CELL.2013.04.025.

Wang, L. *et al.* (2017) 'CRISPR/Cas9-mediated targeted gene correction in amyotrophic lateral sclerosis patient iPSCs', *Protein & cell*. Protein Cell, 8(5),

pp. 365–378. doi: 10.1007/S13238-017-0397-3.

Wang, Z. *et al.* (2012) ‘Successful Regional Delivery and Long-term Expression of a Dystrophin Gene in Canine Muscular Dystrophy: A Preclinical Model for Human Therapies’, *Molecular Therapy*, 20(8), pp. 1501–1507. doi: 10.1038/mt.2012.111.

Warner, L. E. *et al.* (2002) ‘Expression of Dp260 in muscle tethers the actin cytoskeleton to the dystrophin-glycoprotein complex and partially prevents dystrophy’, *Human molecular genetics*. Hum Mol Genet, 11(9), pp. 1095–1105. doi: 10.1093/HMG/11.9.1095.

Wasala, N. B., Chen, S. J. and Duan, D. (2020) ‘Duchenne muscular dystrophy animal models for high-throughput drug discovery and precision medicine’, *Expert Opinion on Drug Discovery*. Taylor and Francis Ltd, pp. 443–456. doi: 10.1080/17460441.2020.1718100.

Wassing, I. E. *et al.* (2021) ‘The RAD51 recombinase protects mitotic chromatin in human cells’, *Nature Communications 2021 12:1*. Nature Publishing Group, 12(1), pp. 1–17. doi: 10.1038/s41467-021-25643-y.

Welch, E. M. *et al.* (2007) ‘PTC124 targets genetic disorders caused by nonsense mutations’, *Nature*, 447(7140), pp. 87–91. doi: 10.1038/nature05756.

West, N. A. *et al.* (2013) ‘Patterns of Growth in Ambulatory Males with Duchenne Muscular Dystrophy’, *The Journal of Pediatrics*, 163(6), pp. 1759-1763.e1. doi: 10.1016/j.jpeds.2013.08.004.

Wiles, M. V. *et al.* (2015) ‘CRISPR-Cas9-mediated genome editing and guide RNA design’, *Mammalian genome : official journal of the International Mammalian Genome Society*. Mamm Genome, 26(9–10), pp. 501–510. doi: 10.1007/S00335-015-9565-Z.

Williams, G. J. *et al.* (2014) ‘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*. DNA Repair (Amst), 17, pp. 110–120. doi: 10.1016/J.DNAREP.2014.02.009.

Williams, M. P. I. *et al.* (2019) ‘A novel optical tissue clearing protocol for mouse skeletal muscle to visualize endplates in their tissue context’, *Frontiers in Cellular Neuroscience*. Frontiers Media S.A., 13, pp. 1–12. doi: 10.3389/FNCEL.2019.00049/BIBTEX.

Wu, B. *et al.* (2008) ‘Effective rescue of dystrophin improves cardiac function in dystrophin-deficient mice by a modified morpholino oligomer’, *Proceedings of the National Academy of Sciences of the United States of America*, 105(39), pp. 14814–14819. doi: 10.1073/PNAS.0805676105/SUPPL_FILE/0805676105SI.PDF.

Xi, H. *et al.* (2017) ‘In Vivo Human Somitogenesis Guides Somite Development from hPSCs’, *Cell Reports*, 18(6), pp. 1573–1585. doi: 10.1016/j.celrep.2017.01.040.

Xi, H. *et al.* (2020) ‘A Human Skeletal Muscle Atlas Identifies the Trajectories of Stem and Progenitor Cells across Development and from Human Pluripotent Stem Cells’, *Cell Stem Cell*. Cell Press, 27(1), pp. 158-176.e10. doi: 10.1016/j.stem.2020.04.017.

Xu, J., Du, Y. and Deng, H. (2015) ‘Direct Lineage Reprogramming: Strategies, Mechanisms, and Applications’, *Cell Stem Cell*, 16(2), pp. 119–134. doi: 10.1016/j.stem.2015.01.013.

Xu, S. *et al.* (2020) ‘Abnormalities in Brain and Muscle Microstructure and Neurochemistry of the DMD Rat Measured by in vivo Diffusion Tensor Imaging and High Resolution Localized 1H MRS’, *Frontiers in Neuroscience*. Frontiers Media S.A., 14. doi: 10.3389/fnins.2020.00739.

Xu, X. *et al.* (2017) ‘Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells’, *Stem cell reports*. Stem Cell Reports, 8(3), pp. 619–633. doi: 10.1016/J.STEMCR.2017.01.022.

XU, Z. yin *et al.* (2017) ‘Action modes of transcription activator-like effectors (TALEs) of Xanthomonas in plants’, *Journal of Integrative Agriculture*. Chinese Academy of Agricultural Sciences, 16(12), pp. 2736–2745. doi: 10.1016/S2095-3119(17)61750-7.

Yamanaka, S. (2009) ‘A Fresh Look at iPS Cells’, *Cell*, 137(1), pp. 13–17. doi: 10.1016/j.cell.2009.03.034.

Yang, H. T. *et al.* (2012) ‘Dystrophin Deficiency Compromises Force Production of the Extensor Carpi Ulnaris Muscle in the Canine Model of Duchenne Muscular Dystrophy’, *PLoS ONE*. Edited by J. M. Ervasti, 7(9), p. e44438. doi: 10.1371/journal.pone.0044438.

Yang, M. *et al.* (2018) ‘The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET’, *Cell reports*. Cell Rep, 22(2), pp. 372–382. doi: 10.1016/J.CELREP.2017.12.048.

Yasuno, T. *et al.* (2014) ‘Functional analysis of iPSC-derived myocytes from a patient with carnitine palmitoyltransferase II deficiency’, *Biochemical and Biophysical Research Communications*, 448(2), pp. 175–181. doi: 10.1016/j.bbrc.2014.04.084.

Yin, H. *et al.* (2011) ‘Pip5 transduction peptides direct high efficiency oligonucleotide-mediated dystrophin exon skipping in heart and phenotypic correction in mdx mice’, *Molecular therapy : the journal of the American Society of Gene Therapy*. Mol Ther, 19(7), pp. 1295–1303. doi: 10.1038/MT.2011.79.

Young, C. S., Pyle, A. D. and Spencer, M. J. (2019) ‘CRISPR for Neuromuscular Disorders: Gene Editing and Beyond’, *Physiology (Bethesda, Md.)*. Physiology (Bethesda), 34(5), pp. 341–353. doi: 10.1152/PHYSIOL.00012.2019.

Yu, H. H. *et al.* (2016) ‘Porcine Zygote Injection with Cas9/sgRNA Results

in DMD-Modified Pig with Muscle Dystrophy', *International journal of molecular sciences*. Int J Mol Sci, 17(10). doi: 10.3390/IJMS17101668.

Yue, Y. *et al.* (2008) 'A Single Intravenous Injection of Adeno-associated Virus Serotype-9 Leads to Whole Body Skeletal Muscle Transduction in Dogs', *Molecular Therapy*, 16(12), pp. 1944–1952. doi: 10.1038/mt.2008.207.

Zammit, P. S., Partridge, T. A. and Yablonka-Reuveni, Z. (2006) 'The skeletal muscle satellite cell: the stem cell that came in from the cold', *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*. J Histochem Cytochem, 54(11), pp. 1177–1191. doi: 10.1369/JHC.6R6995.2006.

Zatz, M. *et al.* (2014) 'Milder course in Duchenne patients with nonsense mutations and no muscle dystrophin', *Neuromuscular Disorders*, 24(11), pp. 986–989. doi: 10.1016/j.nmd.2014.06.003.

Zhang, L. *et al.* (2015) 'Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9', *PLoS ONE*. Public Library of Science, 10(3). doi: 10.1371/JOURNAL.PONE.0120396.

Zhang, M., Koishi, K. and McLennan, I. (1998) 'Skeletal muscle fibre types: Detection methods and embryonic determinants', *Histology and Histopathology*. Histol Histopathol, 13(1), pp. 201–207. doi: 10.14670/HH-13.201.

Zhang, Y. *et al.* (2017) 'CRISPR-Cpf1 correction of muscular dystrophy mutations in human cardiomyocytes and mice', *Science advances*. Sci Adv, 3(4). doi: 10.1126/SCIADV.1602814.

Zhang, Y. *et al.* (2018) 'Myoediting: Toward Prevention of Muscular Dystrophy by Therapeutic Genome Editing', *Physiological reviews*. Physiol Rev, 98(3), pp. 1205–1240. doi: 10.1152/PHYSREV.00046.2017.

Zhao, C. *et al.* (2014) 'Recombinase-mediated reprogramming and dystrophin gene addition in mdx mouse induced pluripotent stem cells', *PLoS ONE*. Public Library of Science, 9(4). doi: 10.1371/journal.pone.0096279.

Zincarelli, C. *et al.* (2008) 'Analysis of AAV Serotypes 1–9 Mediated Gene Expression and Tropism in Mice After Systemic Injection', *Molecular Therapy*. Cell Press, 16(6), pp. 1073–1080. doi: 10.1038/MT.2008.76.

Zuo, Z. and Liu, J. (2017) 'Structure and Dynamics of Cas9 HNH Domain Catalytic State', *Scientific Reports 2017 7:1*. Nature Publishing Group, 7(1), pp. 1–13. doi: 10.1038/s41598-017-17578-6.

8 SUPPLEMENTARY MATERIALS

8.1 PUBLICATIONS RELATING TO THE PHD

8.1.1 3D HUMAN INDUCED PLURIPOTENT STEM CELL-DERIVED BIOENGINEERED SKELETAL MUSCLES FOR TISSUE, DISEASE AND THERAPY MODELLING, 2022, NATURE PROTOCOLS

8.1.2 THREE-DIMENSIONAL HUMAN IPSC-DERIVED ARTIFICIAL SKELETAL MUSCLES MODEL MUSCULAR DYSTROPHIES AND ENABLE MULTILINEAGE TISSUE ENGINEERING, 2018, CELL REPORTS