University of Nottingham
UK | CHINA | MALAYSIA

Investigating Epitranscriptomic Modification of the *DMPK* Expansion RNA in DM1.

Submitted on the 2nd of March 2026 in fulfilment of the conditions for the degree MRes in Molecular Genetics.

Imogen Lily Rayer, Student ID: 20358213

Supervised by Prof. J David Brook & Rebecca Trueman

School of Life Sciences

Word Count: 23, 606

Table of Contents

Acknowledgements

I would like to sincerely thank my supervisor, David, for his invaluable guidance and support throughout this project. I am deeply grateful for all the time, advice, and encouragement he has given me over the past year.

My thanks also go to Tushar, the undisputed ‘western blot king,’ and PhD student Petter for their generous help in troubleshooting and optimising my western blots. A special mention to Petter as well for kindly sharing his R-script, which greatly improved the efficiency of my foci data analysis.

I am very grateful to Sarah for training me in cell culture techniques and for always offering a helping hand whenever it was needed. I would also like to thank the rest of the lab - Oliver, Daisy, Gareth, Jessie, Binta, Lucy, James, and Becky - for creating such a supportive and enjoyable working environment.

Finally, a massive thank you to my grandma, whose belief in me and generosity in funding this degree made it possible!

Abstract

Myotonic dystrophy type 1 (DM1) is a rare, progressive, multisystemic disorder caused by a (CTG)n expansion in the *DMPK* gene. The mutant mRNA forms hairpin structures that sequester muscleblind-like proteins, leading to the formation of ribonuclear foci and dysregulated RNA metabolism. Epitranscriptomic modifications such as 5-methylcytosine (m^5C) are emerging regulators of RNA metabolism and may influence DM1 pathology. Epitranscriptomic regulation is mediated by 'writers,' 'readers,' and 'erasers' that add, interpret, or remove RNA modifications. This study investigates the role of m^5C in DM1 by targeting the 'reader,' ALYREF, involved in mRNA export and m^5C recognition. This study also incorporates unpublished datasets on *ALYREF* knockdown and *NSUN2* knockout, providing a broader perspective on m^5C regulation in DM1.

DM1 patient-derived fibroblasts (KB-TeloMyoD) were transfected with either ALYREF-specific siRNAs, a non-targeting siRNA control or a transfection control. Cells were analysed by fluorescence *in situ* hybridisation (FISH). Foci area, intensity, and number per nucleus were quantified using the Zeiss CellDiscovery7 microscope, R-script, and GraphPad Prism. ALYREF knockdown was validated with western blotting.

Preliminary FISH results (n=4) show that ALYREF knockdown resulted in statistically significant but variable effects on RNA foci area, number, and intensity. NSUN2 knockout (n=2) produced an approximate two-fold reduction under similar conditions. These findings suggest that ALYREF and NSUN2, and by extension m^5C, may influence RNA foci formation and stability in DM1. Future work investigating TET2 will further define epitranscriptomic contributions to DM1 pathology and may highlight novel avenues for therapeutic intervention aimed at reducing RNA toxicity.

Abbreviations

ALYREF = ALY/REF Export Factor

DM1 = Myotonic Dystrophy Type 1

DM2 = Myotonic Dystrophy Type 2

DMPK = myotonic dystrophy protein kinase

FISH = Fluorescence *In Situ* Hybridisation

KD = knockdown

KO = knockout

m^5C = 5-methylcytosine

NSUN2 = NOP2/Sun RNA methyltransferase 2

REDs = repeat expansion disorders

TET2 = Ten-eleven translocation methylcytosine deoxygenase 2

TREX complex = transcription/export complex

1 Introduction

Myotonic dystrophy belongs to a group of rare genetic disorders known as muscular dystrophies, characterised by progressive muscle weakening and atrophy. These disorders result in significant impairments in mobility, dexterity, and overall motor function (Huml, 2015). The underlying cause behind most muscular dystrophies lies in genetic mutations that affect the regulation of proteins involved in muscle structure and function, ultimately leading to the degeneration and loss of muscle cells and tissue (see Table 1; Emery, 2002). These mutations can be transmitted as autosomal dominant, autosomal recessive, or X-linked recessive inheritance (Harper, 1989).

Despite advances in understanding the genetic mechanisms of these conditions, there is no cure. In 2016, the global prevalence of muscular dystrophies was estimated at 16 to 25 individuals per 100,000 (Theadom et al., 2015; Mah et al., 2016). Of the nine main types, Duchenne muscular dystrophy (DMD), an X-linked recessive disorder, is the most severe form of muscular dystrophy in childhood, affecting 19.8 live male births per 100,000 and reducing life expectancy to less than 30 years (see Table 1). On the other hand, myotonic dystrophy type 1 (DM1) affects 48 per 100,000 individuals and is considered the most prevalent form overall. (Sacconi et al., 2013; Huml, 2015; Crisafulli et al., 2020; Klug et al., 2020; Johnson et al., 2021). This emphasises the importance of researching this group of diseases.

Form of Muscular Dystrophy	**Gene(s) of Interest**	**Prevalence per 100,000**	**Age of Onset**	**References**
Becker (BDM)	reduction of *dystrophin*	1.6 (mostly males)	5 to 15 years	Hoffman et al., 1988; Salari et al., 2022.
Congenital (CMD)	~35 genes	<1	Birth to Early Infancy	Zambon & Muntoni, 2021.
Distal	>25 genes	<1	Mid-late Adulthood	Norwood et al., 2009; Milone & Liewluck, 2019
Duchenne (DMD)	absence of *dystrophin*	2.8 (general population) 19.80 (live male births)	Before 4 years	Hoffman et al., 1988; Crisafulli et al., 2020; Salari et al., 2022
Emery-Dreifuss (EDMD)	~9 genes	0.39	Childhood to Adolescence	Mah et al., 2016; Heller et al., 2020
Facioscapulo -humeral (FSHD)	*DUX4* *SMCHD1*	3.95	Adolescence to Adulthood	Van der Maarel et al., 2012; Mah et al., 2016
Limb-Girdle (LGMD)	~30 genes	1.63	Childhood to Adulthood	Mah et al., 2016; Taghizadeh et al., 2019
Myotonic Dystrophy Type 1 (DM1)	*DMPK*	48	Variable	Brook et al., 1992; Johnson et al., 2021
Myotonic Dystrophy Type 2 (DM2)	*CNBP*	2.29	Mid-adulthood	Liquori et al., 2001; Liao et al., 2022
Oculo-pharyngeal (OPMD)	*PABPN1*	1	~50 years	Richard et al., 1998; Malerba et al., 2017

Table 1: Overview of the main muscular dystrophies.

Muscular dystrophies are traditionally classified by age of onset, patterns of muscle weakness, histopathology, disease progression, systemic involvement and inheritance (Huml, 2015). Table 1 summarises this information for each of the nine main types of muscular dystrophy.

1.1 Overview of Myotonic Dystrophy

Myotonic dystrophy is a genetic disorder characterised by progressive muscle weakness (dystrophy), delayed muscle relaxation after voluntary contraction (myotonia), and multi-system involvement (Harper, 1989). It was the first autosomal dominant condition linked to a non-coding repeat expansion that is transcribed into RNA but not translated into protein (Norman et al., 1989; Brook et al., 1992). There are two forms of this disorder: DM1 and DM2. While both disorders are caused by unstable repeat expansions, they differ in the affected gene, repeat sequence, and genomic location (see Figure 1).

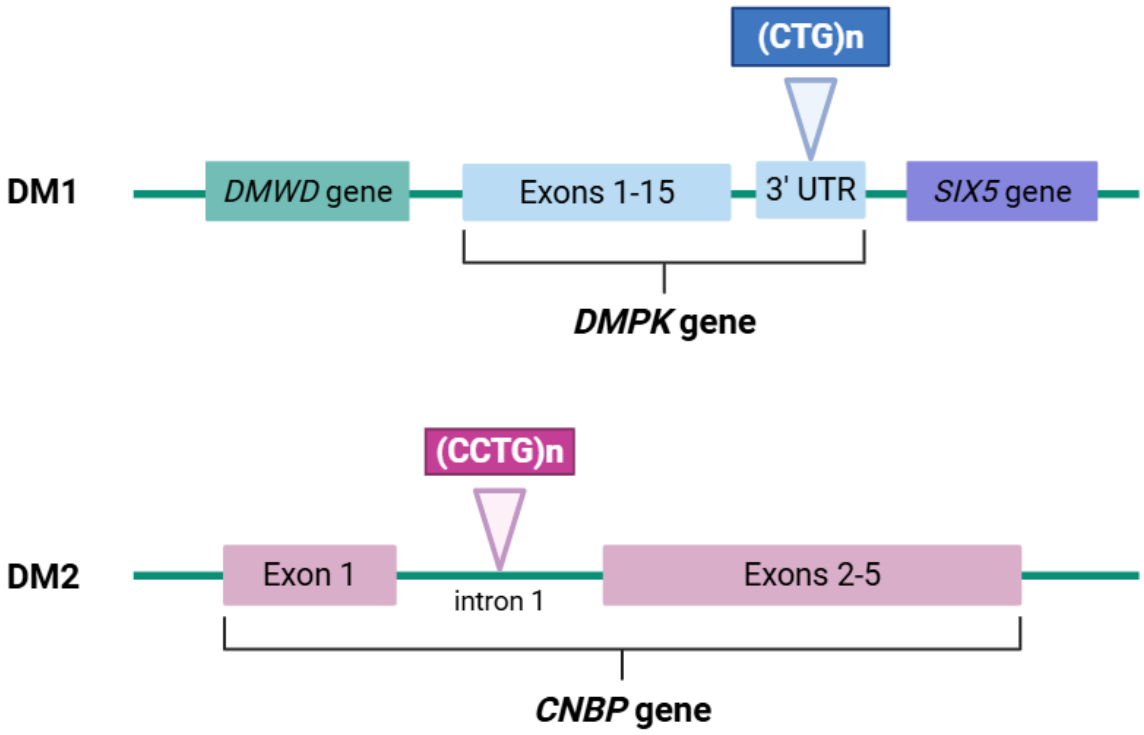

Figure 1. Genetics of DM1 and DM2.

The (CTG)n triplet repeat expansion in DM1 is in the 3' untranslated region (UTR) of chromosome 19q13.3. It affects the *DMPK* gene, the *SIX homeobox 5* (*SIX5*) gene and the *myotonic dystrophy WD repeat-containing* (*DMWD*) gene. These lesser-involved genes are presumed to contribute to some of the disease's clinical symptoms (see Figure 2), especially reduced expression of *SIX5,* which is associated with cataracts (Winchester et al., 1999). The (CCTG)n tetranucleotide repeat expansion in DM2 is found in intron 1 of the *CNBP* gene on chromosome 3q21.3. (Figure adapted from Visconti et al., 2021).

The DM1 mutation (CTG)n is in the 3’ untranslated region of the *myotonic dystrophy protein kinase* (*DMPK)* gene, spanning 14 exons over a 14 kb region and producing multiple mRNA isoforms across various tissues (Brook et al., 1992; Mahadevan et al., 1992; Novelli et al., 1993). The DM2 mutation (CCTG)n is in the first intron of the *Cellular Nucleic Acid Binding Protein* (*CNBP)* gene, formally known as ZNF9, which is comprised of 5 exons (see Figure 1; Ranum et al., 1998; Liquori et al., 2001).

These genetic differences are reflected in the clinical features. DM1 manifests significantly earlier and more severely, with pronounced muscle weakness, myotonia, and multi-system complications, whereas DM2 presents later in life, primarily affecting the proximal muscles but with fewer systemic symptoms (see Figure 2).

1.1.1 Clinical features and complications

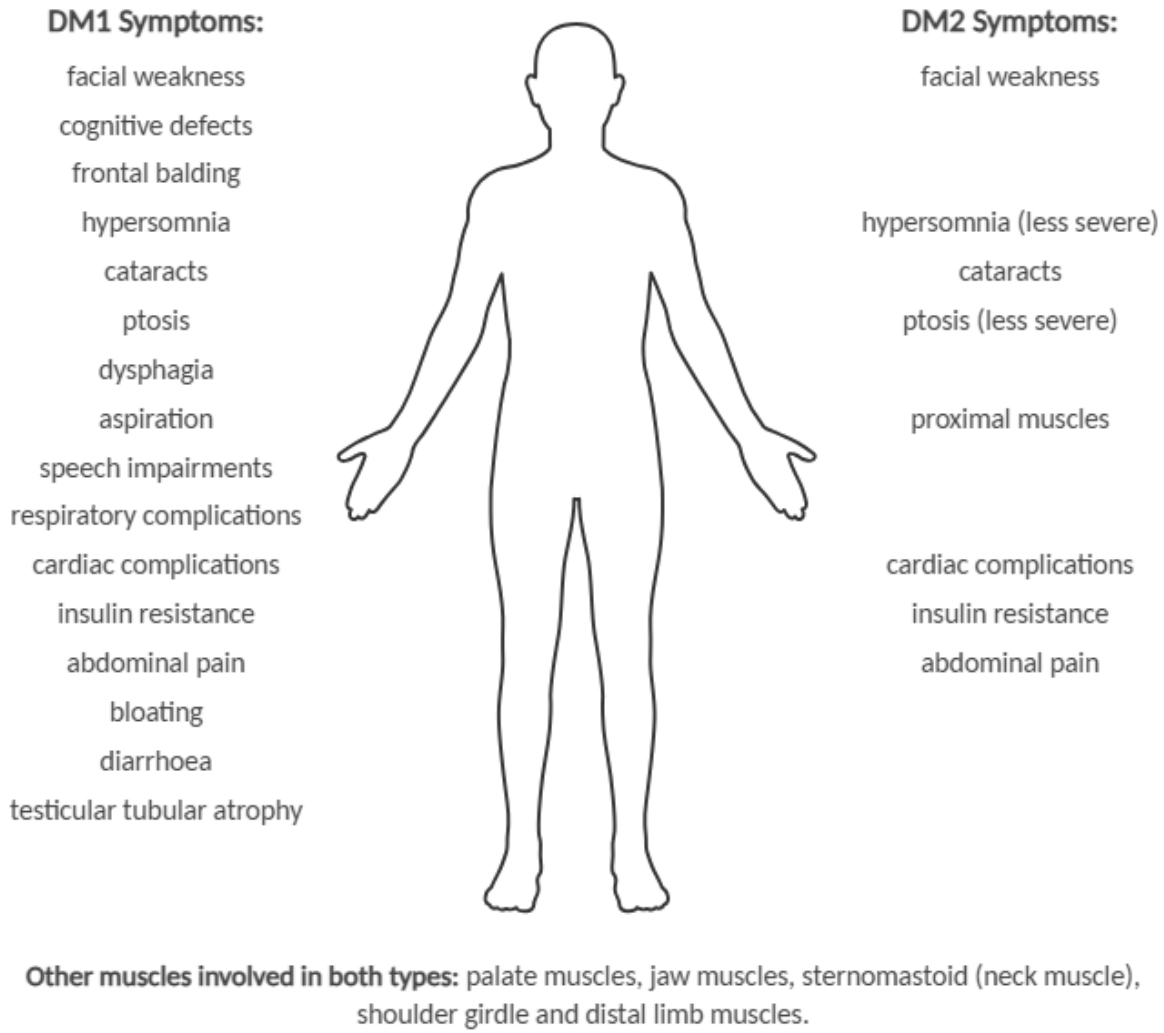

Figure 2. Symptoms of DM1 vs. DM2.

Myotonic dystrophy is a multisystemic disorder with extremely variable clinical manifestations, from asymptomatic adults to severely affected neonates. Whilst there are many overlapping symptoms, DM1 generally presents with greater severity. This diagram represents an individual with myotonic dystrophy regardless of gender or age. (Self-made diagram using BioRender.com, based upon information in Soltanzadeh, 2022 and Harper, 1989).

Respiratory dysfunction is the leading cause of mortality in myotonic dystrophy, accounting for 17.8% of DM1 deaths, followed by sudden cardiac death (13.7%) and pneumonia (12.9%) (Wahbi et al., 2019). It leads to restrictive ventilatory failure, recurrent aspiration, and ultimately, respiratory failure (Mathieu et al., 1999; Lo Mauro & Aliverti, 2016; Hawkins et al., 2019). This stems from progressive respiratory pump failure, which manifests as: (1) restrictive pulmonary function, (2) abnormal thoracoabdominal breathing patterns, (3) chronic CO_2 retention, (4) dyspnoea, (5) impaired breathing regulation due to altered feedback from respiratory muscle receptors, (6) ineffective coughing from bulbar and respiratory muscle weakness, and (7) sleep-related breathing disorder (Kiyan et al., 2010; Lo Mauro & Aliverti., 2016; Hawkins et al., 2019; Rossi et al., 2019). Clinically, these complications are reflected as excessive daytime sleepiness (63%), exertional dyspnoea (48.3%), chronic obstructive pulmonary disease (13.2%), and apnoea (12%) among DM1 patients (Rossi et al., 2019).

Cardiac complications are the second leading cause of unexpected morbidity and mortality in myotonic dystrophy (Groh et al., 2008; Wahbi et al., 2018). Up to 75% of adult DM1 patients are affected, with a lower but notable incidence in DM2 (Mahadevan et al., 2021). Common complications include arrhythmias (e.g., atrial fibrillation/flutter, ventricular tachycardia/fibrillation), conduction defects (e.g., atrioventricular block and bundle branch block), and left ventricular systolic dysfunction, which occurs in up to 40% of DM1 patients (Russo et al., 2020; Mahadevan et al., 2021). These issues arise early in disease progression, partly due to the misregulation of key cardiac transcripts.

These abnormalities arise early in disease progression and have been linked to downstream molecular consequences of RNA toxicity, which are discussed in detail in Section 1.2 and Table 2. For example, aberrant splicing of SCN5A, encoding the cardiac voltage-gated sodium channel Nav1.5, reduces cardiomyocyte excitability and contributes to conduction delays and arrhythmias in DM1 heart tissue (Freyermuth et al., 2016). Additionally, RNA toxicity-induced upregulation of NKX2-5, a transcription factor crucial for cardiac conduction system maintenance, may further

exacerbate conduction abnormalities (Yadava et al., 2008). Sudden cardiac death is the leading cardiovascular cause of death in DM1, accounting for 43% of cardiac deaths and 13.7% of total deaths (Wahbi et al., 2018).

Myotonic dystrophy can also lead to a range of ocular complications: ptosis (droopy eyelids), extraocular muscle weakness, low intraocular pressure, retinal damage and cataracts (Greenfield, 1911; Burian & Burns, 1967). Irrespective of age, patients can also exhibit various speech and language impairments. Mutant *DMPK* mRNA is expressed in cortical and subcortical neurons, suggesting that the CNS is a significant site of RNA toxicity in DM1. This accumulation of mutant RNA and subsequent compromise of MBNL1 correlates with the observed neurological and cognitive symptoms (Jiang et al., 2004).

1.1.2 Genetic and molecular characteristics

DM1, also known as Steinert's disease, was first identified in 1909 (Steinert, 1909). It affects 86% of patients with myotonic dystrophy, notably in Europe, China, India, and the Middle East (Wood et al., 2018).

The DMPK protein, which is central to the genetic basis of DM1, is a serine/threonine kinase, primarily expressed in skeletal and cardiac muscle and, to a lesser extent, in smooth muscle, with key roles in muscle physiology and cellular signalling (Bush et al., 2000; Lam et al., 2000). It can also be found in the spinal motor neurons of the brain and the retina and ciliary body of the eyes (Balasubramanyam et al., 1998; Winchester et al., 1999).

DMPK promotes muscle contractions by phosphorylating myosin phosphatase target subunit 1 (MYPT1), in a similar way to Rho-kinases, inhibiting myosin phosphatase activity, maintaining myosin light chain (MLC) phosphorylation, increasing Ca^{2+} sensitivity, and sustaining contraction (Kimura et al., 1996; Hartshorne et al., 1998; Murányi et al., 2001). DMPK also regulates membrane excitability by

phosphorylating phospholemman (PLM), reducing its membrane expression and Cl^- conductance, a mechanism linked to the myotonia observed in DM1 – see Figure 2 (Mounsey et al., 2000). MYPT1 inhibition by DMPK also leads to cytoskeletal reorganisation by increasing stress fibres and focal adhesions in non-muscle cells (Kimura et al., 1996; Murányi et al., 2001). It must be noted that the C-terminal domain of DMPK can cause autoinhibition, with truncation enhancing the phosphorylation of targets such as PLM and MYPT1 by up to 10-fold (Bush et al., 2000; Wansink et al., 2003).

On the other hand, DM2, also known as proximal myotonic myopathy (PROMM), was first clinically distinguished from DM1 in 1994 (Thornton et al., 1994). *CNBP* is highly conserved across eukaryotes and encodes a nucleic acid-binding protein with multiple CCHC-type zinc knuckle motifs and an RGG box (Armas et al., 2008). These domains enable *CNBP* to act as a G-quadruplex (G4) chaperone, facilitating transcriptional activation of genes such as c-MYC and TCOF1 in vertebrate models (Armas et al., 2008; Benhalvey et al., 2017). Through this mechanism, CNBP facilitates the transcriptional activation of genes such as c-MYC and TCOF1 in vertebrate models, underscoring its role as a G4 chaperone in development and disease (David et al., 2019; Rosas et al., 2024).

Evidence from model organisms demonstrates the developmental importance of *CNBP*. In Drosophila melanogaster, knockout (KO) of the homolog *no neck (non)* causes early embryonic lethality with craniofacial abnormalities, while complete *CNBP* deletion in mice results in severe forebrain and craniofacial defects and early embryonic death (Chen et al., 2003; Chen et al., 2007). Furthermore, CNBP haploinsufficiency ($CNBP^{+/-}$) in mice produces multiorgan abnormalities resembling DM2 pathology (Antonucci et al., 2013).

1.1.3 Diagnostic and clinical significance of repeat expansion size

The copy number of repeat expansions in *DMPK* and *CNBP* is a key diagnostic marker. In DM1, 5-37 (CTG)n repeats are considered within the normal non-pathogenic range, 38-49 are referred to as premutations, and ≥50 repeats are pathogenic. Whilst carriers within the premutation range are asymptomatic, the CTG repeats are still highly unstable and prone to expanding when passed on to the next generation.

The DM1 copy number can predict the subtype and age of onset, with larger expansions often associated with a more severe subtype and an earlier age of onset (Hunter et al., 1992; Cumming et al., 2019). However, accurately predicting disease course based solely on repeat length remains challenging due to individual variability over time and somatic instability (see Section 1.2.1).

In DM2, 10-26 (CCTG)n are considered normal, and >75 repeats are pathogenic - DM2 repeat expansion RNA is ~10-fold higher (Harley et al., 1993; Martorell et al., 1995; Mankodi et al., 2003; Udd & Krahe, 2012; Strachan & Lucassen, 2023). Unlike DM1, repeat length in DM2 does not appear to correlate with disease severity or the age of onset, making it less useful for subtyping (Day et al., 2003; Udd & Krahe, 2012).

These pathogenic repeat expansions are part of a wider class of repeat expansion disorders (REDs), as highlighted in Figure 3. They are caused by short-tandem repeats, which are among the most prevalent forms of genetic variation in humans, and make up 3% of the human genome (Lander et al., 2001; Shi et al., 2023). These loci are inherently polymorphic and exhibit length-dependent instability (Depienne & Mandel, 2021).

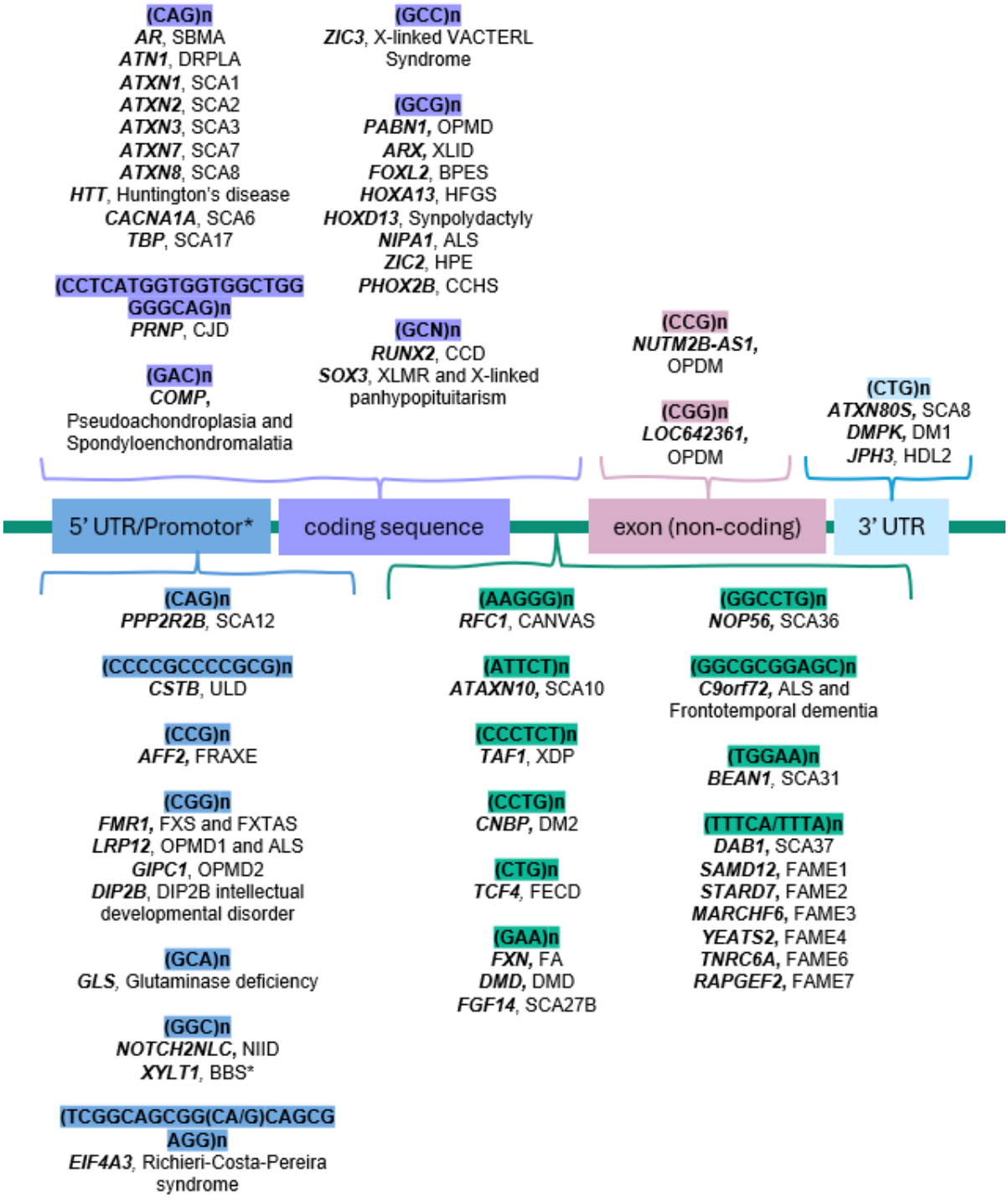

Figure 3: Summary of repeat expansion disorders.

As of 2025, 62 REDs have been identified (Chen et al., 2025). The majority are triplet repeat disorders (45 in total), but only three non-coding triplet repeat expansions have been described in 3' UTRs: SCA8, HDL2 and DM1 (Depienne & Mandel, 2021). The schematic illustrates the genomic distribution of REDS, highlighting repeat motifs, associated genes, and corresponding disorders (adapted from Ellerby, 2020; Depienne & Mandel, 2021; Shi et al., 2024; Chen et al., 2025).

1.2 Molecular mechanism of DM1 vs. DM2

The pathogenesis of myotonic dystrophy involves multiple stages, including transcription, aberrant nuclear retention of mutant transcripts, and their downstream processing or degradation (Xing et al., 2021). A central molecular feature of these disorders is the accumulation of expanded-repeat-containing RNAs in the nucleus, although the extent of nuclear retention can vary between DM1 and DM2. As shown in Figure 4, the mutant repeat expansions are transcribed into RNA and form stable, hairpin-like secondary structures (Klug et al., 2020). These GC-rich RNAs are predominantly retained in the nucleus, limiting their availability for cytoplasmic translation; however, a subset of expanded transcripts may escape nuclear retention, and additional pathogenic mechanisms beyond RNA sequestration have been reported (Yuan et al., 2007; Meola & Cardani, 2015).

The ubiquitously expressed splicing regulator proteins muscleblind-like 1 and 2 (MBNL1 and MBNL2) exhibit high affinity for expanded CUG-repeat RNA through recognition of repeated YGCY (Y = C or U) sequence motifs (the consensus MBNL binding site), which are bound by their four CCCH zinc-finger domains and presented at high density within the pathogenic hairpin secondary structures formed by the repeat expansions (Goers et al., 2010; Konieczny et al., 2014). Their binding results in the formation of RNA-protein complexes known as ribonuclear foci, which can be visualised microscopically using fluorescence *in situ* hybridisation (FISH) assays, providing direct evidence of nuclear retention of mutant transcripts in myotonic dystrophy (Swanson et al., 1993; Taneja et al., 1995; Michalowski et al., 1999).

MBNL proteins are key regulators of developmentally programmed alternative splicing, and their sequestration within ribonuclear foci (see Figure 4) leads to a functional depletion of MBNL activity. This loss of MBNL function is a hallmark of the molecular pathology of myotonic dystrophy and results in widespread misregulation of alternative splicing during postnatal development, contributing to disease progression (Mankodi et al., 2001; Xing et al., 2021).

While the downstream molecular pathology of myotonic dystrophy is driven by pathogenic repeat-containing RNA, the nature and severity of these mechanisms are fundamentally determined by the underlying genetic features of the disease, including repeat length, inheritance patterns, and repeat instability.

1.2.1 Genetic features and inheritance of DM1

DM1 follows an autosomal dominant inheritance pattern, with a 50% transmission risk to offspring of an affected individual (Shaw & Harper, 1989). Both sexes inherit the condition equally, although among asymptomatic carriers, the father is more frequently the transmitting parent (Pavićević et al., 2013). DM1 is strongly characterised by genetic anticipation, phenotypic variability, mosaicism and germline instability.

1.2.1.1 Genetic Anticipation

Anticipation refers to the progressively earlier onset and greater severity of a disease in successive generations (Harper, 1989). In DM1, the number of CTG repeats typically increases across generations due to meiotic instability, wherein larger expansions correlate with greater disease severity and earlier onset (Harley et al., 1993; Miller et al., 2000). A 1989 study found that in 98% of parent-child pairs, onset occurred over a decade earlier in the child, highlighting the impact of intergenerational repeat expansion, particularly from premutation carriers whose near-threshold repeat lengths often expand significantly in offspring (Höweler et al., 1989; Joosten et al., 2020).

1.2.1.2 Phenotypic Variability and Mosaicism

Phenotypic variability is the observed variability of a particular trait, and genetic mosaicism is the presence of two or more genetically distinct cell lineages within the same individual that are derived from a single fertilised egg (Thorpe et al., 2020). DM1 displays incomplete penetrance, with significant heterogeneity in symptoms among individuals and even within families, reflecting its high clinical and genetic variability (Tomé & Gourdon, 2020). This is because there are very high levels of

somatic mosaicism, driven by mitotic/age-related CTG repeat instability, which contributes to phenotypic variability (Monckton et al., 1995; Fortune et al., 2000). Expansions can vary widely between tissues and increase with age, leading to large differences in repeat size within the same individual (Harley et al., 1993; Martorell et al., 1995; Martorell et al., 2000).

1.2.1.3 Germline Instability

Germline instability also underlies intergenerational variability in DM1. Maternal transmission shows a strong bias toward large expansions during oogenesis, which explains the exclusive maternal inheritance of congenital DM1 (Harley et al., 1993). In contrast, non-congenital forms display greater variability during paternal transmission, with sperm cells often showing significant expansions or contractions of CTG repeat counts increasing by 100–1,000 units in some cases (Martorell et al., 2000). Additionally, the presence of rare repeat interruptions, e.g. CCG or CTC, appear to stabilise the repeat, particularly during paternal transmission, thereby reducing the likelihood of severe expansions and mitigating disease severity (Cumming et al., 2018).

1.2.2 Subtypes of DM1

The genetic properties of DM1, particularly CTG repeat length, intergenerational instability, and mosaicism, give rise to a spectrum of clinical subtypes that differ in age of onset, severity, and prognosis.

1.2.2.1 Congenital DM1 (CDM1)

CDM1, the most severe form, typically involves >1400 CTG repeats and is predominantly maternally inherited due to marked repeat expansions during female gametogenesis (Harper, 1975; Pavićević et al., 2013; Soltanzadeh, 2022). Prenatal signs include polyhydramnios, reduced foetal movement, and prematurity (Dunn & Dierker, 1973; Stokes et al., 2019). Postnatal features often include severe hypotonia, respiratory distress (with >50% requiring ventilation), and characteristic facial weakness (Dodge et al., 1966; Stokes et al., 2019; Hanoun et al., 2022).

Clubfoot occurs in ~50% of cases due to foetal hypotonia (Harper, 1989). Prognosis is poor, with 50% mortality by the mid-thirties (Pavićević et al., 2013).

1.2.2.2 Childhood/Juvenile

Emerging in 1–10-year-old children with 50–1000 CTG repeats, this form manifests as dysarthria, delayed motor milestones, facial weakness, and social difficulties (Koch et al., 1991; Soltanzadeh, 2022). Intellectual disabilities affect over 90% of patients, and ADHD and mood disorders are significantly more prevalent. Some studies report an unusually high incidence of psychotic symptoms (Jacobs et al., 2017; Ho et al., 2019).

1.2.2.3 Adult-onset

The most common form, adult-onset, occurs in individuals of 10–50 years and is associated with 50–1000 CTG repeats (Pavićević et al., 2013). While men frequently present with myotonia, severe muscle weakness, cardiac complications, and respiratory issues, women more often exhibit cataracts, dysphagia and gastrointestinal dysfunctions (Dogan et al., 2016; Hanoun et al., 2022). Life expectancy is reduced, with early symptom onset linked to poorer outcomes. Common causes of death include pneumonia and cardiac arrhythmias (de Die-Smulders et al., 1998; Alsaggaf et al., 2024).

1.2.2.4 Late onset

Typically diagnosed after age 50 with 50-100 CTG repeats, this form is mild. It features symptoms such as myotonia, cataracts, and frontal balding, or patients are asymptomatic (Hanoun et al., 2022; Soltanzadeh, 2022).

1.2.3 Ribonuclear foci

At the molecular level, variation in CTG and CCTG repeat length that underpins both genetic instability and clinical heterogeneity directly influences the behaviour of mutant transcripts, including their nuclear retention and propensity to form ribonuclear foci.

Ribonuclear foci are abnormal, predominantly nuclear aggregates of mutant transcripts containing expanded (CUG)n or (CCUG)n repeats. DM1 and DM2 ribonuclear foci differ in size, shape, and fluorescent intensity. The DM2 foci are rod-like, rather than spherical, as seen in DM1 (Wojciechowska & Krzyzosiak., 2011). DM2 foci are also larger and display an 8-13-fold greater fluorescence intensity, because DM2 (CCTG)n expansions are longer (Mankodi et al., 2003; Cardani et al., 2014). Notably, while DM1 foci are largely restricted to the nucleus, expanded (CCUG)n transcripts in DM2 can also accumulate in the cytoplasm, indicating incomplete nuclear retention and suggesting additional pathogenic mechanisms beyond nuclear RNA toxicity (Salisbury et a., 2009; Meola & Cardani, 2015).

The relationship between expansion size and RNA foci formation in myotonic dystrophy is complex. However, several studies have found correlations between the number of ribonuclear foci and (CTG)n expansion size (Botta et al., 2008; Ballester-Lopez et al., 2020; Núñez-Manchón et al., 2024).

Using RNA-FISH, Botta et al. (2008) showed that the number of ribonuclear foci in DM1 muscle tissue increased with greater (CTG)n repeat size. The foci appeared larger and more intensely fluorescent in the samples with longer repeats, and the proportion of nuclei lacking foci significantly declined. These findings indicate that CTG expansion size strongly influences nuclear retention and aggregation of DMPK transcripts, rather than the number of transcriptional events detected by RNA-FISH.

Ballester-Lopez et al. (2020) further supported this through 3D analysis, showing a strong correlation between (CTG)n repeat length and both the average number of nuclear foci per myoblast and the total number (cytoplasmic and nuclear) of foci per myoblast. RNA foci number correlated with *DMPK* expression (r = 0.967), while cytoplasmic foci correlated with an earlier disease onset (r = -0.818).

More recently, Núñez-Manchón et al. (2024) established three new immortalised DM1 patient-derived muscle cell lines with varying clinical presentations and expansion sizes. All cell lines exhibited RNA foci accumulation; however, the two with the longest expansions and most severe clinical phenotypes (JCC-DM1 and ADE-DM1) had significantly more foci per nucleus and displayed an ~2-fold increase in foci per nucleus. These higher expansion cell lines also had the highest proportion of cells with >10 foci per nucleus, reinforcing the relationship between CTG repeat expansion size and foci formation.

DM2 does not show a strong correlation between (CCUG)n expansion size and disease severity or the number of foci per nucleus (Udd & Krahe, 2012). However, the expression of non-coding CCUG-expanded RNA in *Drosophila* eyes has been shown to cause repeat-length-dependent toxicity (Yu et al., 2015). This toxicity correlated with the formation of RNA foci and sequestration of MBNL1, but unlike DM1, MBNL1 overexpression enhanced rather than rescued the phenotype, indicating a distinct pathogenic mechanism in DM2 (Yu et al., 2015). Consistent with this, MBNLs bind (CCUG)n repeats with a higher affinity than (CUG)n repeats, resulting in larger ribonuclear inclusions in DM2 patients (Kino et al., 2004; Warf & Berglund, 2007; Cardani et al., 2014).

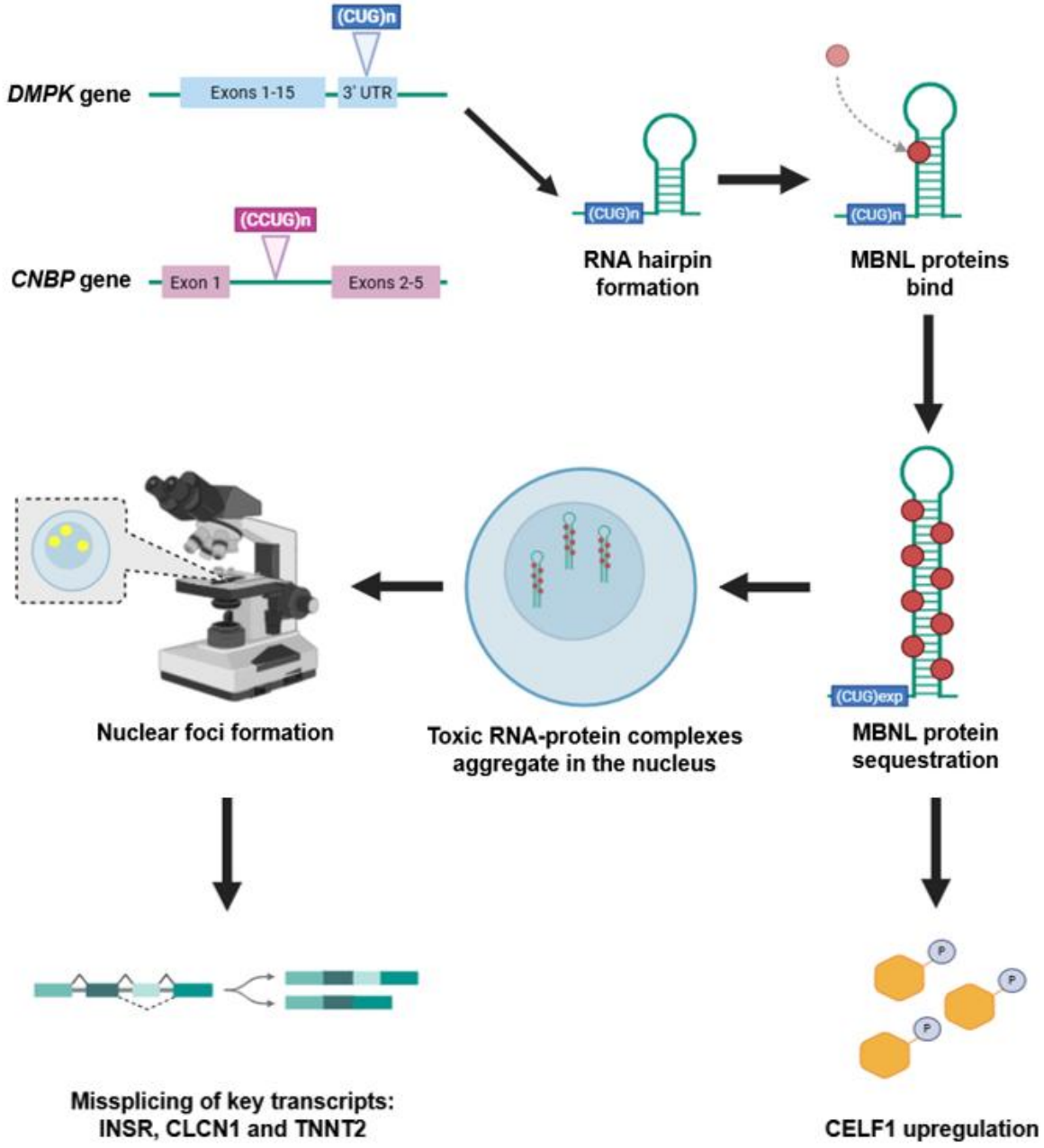

Figure 4. RNA hairpin and ribonuclear foci formation.

Expanded (CUG)n and (CCUG)n RNA repeats form hairpins, which bind MBNL1 and MBNL2 (red circles). This disrupts the alternative splicing of key pre-mRNA targets, including the Insulin Receptor (INSR), Chloride Channel 1 (CLCN1), and Cardiac Troponin T (TNNT2). MBNL sequestration also causes upregulation of CUGBP Elav-like family member 1 (*CELF1*) by PKC-mediated hyperphosphorylation. CELF1, a tissue-specific splicing regulator antagonistic to MBNL1, further alters the splicing of *TNNT2*, *INSR*, *CLCN1*, and *MBNL2* (López-Martínez et al., 2020; Xing et al., 2021). (Self-made diagram using Biorender.com; based upon information in Depienne & Mandel, 2021).

1.2.4 MBNL sequestration and splicing dysregulation

The formation of ribonuclear foci is not just a structural hallmark of myotonic dystrophy, but it also has profound functional consequences, most notably through the sequestration of MBNL proteins and the disruption of alternative splicing regulation.

MBNLs are RNA-binding proteins that regulate a developmental alternative splicing programme established during prenatal development, in which foetal-specific isoforms predominate (Lin et al., 2006; Kalsotra et al., 2010). During the early postnatal period, increasing MBNL expression drives the transition from foetal to adult splicing patterns by repressing embryonic isoforms and promoting adult exon usage (Wang et al., 2012; Lee et al., 2013). Disruption of this developmental splicing switch underlies the persistence of foetal isoforms observed in congenital DM1 (see Section 1.2.2.1).

MBNLs also regulate tissue-specific splicing during mammalian postnatal development, modulating exon inclusions and exclusions in target pre-mRNAs to increase proteome diversity (Cardani et al., 2006; Yuan et al., 2007; Yin et al., 2020). These processes, collectively termed alternative splicing, allow a single gene to produce multiple mRNA isoforms, contributing to cell differentiation, lineage specification, tissue identity, and organ development (Yin et al., 2020).

MBNL1 and 2 have partial compensatory roles, as demonstrated by the upregulation of MBNL2 following MBNL1 knockdown (KD) and widespread splicing defects observed in mice lacking both MBNLs (Wang et al., 2012; Lee et al., 2013). In support of this model, disruption of murine Mbnl1 produces a phenotype of myopathy (myotonia, defective regulation of alternative splicing) and lens cataracts that closely resemble the clinical symptoms of DM1 in humans (Kanadia et al., 2003).

As a consequence, the repeat expansions of DM1 cause MBNL-regulated transcripts to revert to foetal isoforms, impairing protein synthesis for a subset of muscle transcripts and leading to widespread pre-mRNA mis-splicing (Jiang et al., 2004; Xing et al., 2021). This mis-splicing alters cellular metabolism, increases catabolism and drives muscle atrophy (Botta et al., 2008; Strachan & Lucassen, 2023). Therefore, sequestration of MBNLs into ribonuclear foci underpins why myotonic dystrophy is widely considered a spliceopathy (Klug et al., 2020).

Category	Gene	Splicing defect	Type	References
Skeletal muscle contraction and excitability	*ALPK3*	Inclusion of exon 2	DM1	Nakamori et al., 2013
	ATP2A1 (SERCA1)*	Exclusion of exon 22	Both	Kimura et al., 2005; Yamashita et al., 2012
	*BIN1**	Exclusion of exon 11	DM1	Fugier et al., 2011; Carrascosa-Sàez et al., 2025
	CACNA1S	Exclusion of exon 29	Both	Tang et al., 2012; Carrascosa-Sàez et al., 2025
	CAMK2B	Mis-splicing of CamKIIβ	DM1	Nakamori et al., 2013; Falcetta et al., 2024
	CAPN3	Exclusion of exon 16	DM1	Lin et al., 2006; Yamashita et al., 2012
	*CLCN1**	Inclusion of intron 1 Inclusion of exon 7a	Both	Charlet et al., 2002; Mankodi et al., 2002; Lueck et al., 2007
	DTNA (DB2)	Exclusion of exons 11a and 12	DM1	Nakamori et al., 2013
	MTMR1	Inclusion of exons 2.1, 2.2 and 2.3	DM1	Buj-Bello et al., 2002

	MYOM1	Inclusion of exon 17a	DM1	Koebis et al., 2011
	NFIX	Inclusion of exon 7	DM1	Nakamori et al., 2013; Wagner et al., 2016
	*RYR1**	Exclusion of exon 70	DM1	Kimura et al., 2005
	SOS1	Exclusion of exon 25	DM1	López-Martínez et al., 2020
Cardiac conduction and contractility	*ANK2*	Inclusion of mis-spliced segments (miEs)	DM1	Sznajder et al., 2025
	PDLIM3	Inclusion of exon 4 Exclusion of exon 5	Both	Ohsawa et al., 2011; Yamashita et al., 2012
	RBFOX2	Inclusion of exon 40 (instead of exon 43)	DM1	Misra et al., 2020
	*SCN5A**	Inclusion of foetal exon 6a (not adult exon 6b)	DM1	Freyermuth et al., 2016
	*TNNT2**	Inclusion of foetal exon 5 in adult heart and skeletal muscle	Both	Bosè et al., 2019
	*TTN**	Inclusion of Zr4, Zr5 and Mex5 Exclusion of exon 45	DM1	Lin et al., 2006; Yamashita et al., 2012
Metabolic regulation	*INSR**	Exclusion of exon 11	Both	Savkur et al., 2001; Savkur et al., 2004
	PKM2	Inclusion of exons 9 and 10	DM1	Gao & Cooper, 2013
Neurological/ CNS	*MAPT**	Exclusion of exons 2, 3 and 10	DM1	Goedert et al., 1989; Goodwin et al., 2015
	*NMDAR1** *(GRIN1)*	Inclusion of exon 5	DM1	Tsien et al., 1996; Jiang et al., 2004; Goodwin et al., 2015
RNA processing and Splicing factors	*CELF1*	Stabilised by PKC-mediated hyperphosphorylation → altered splicing of multiple targets	DM1	Ho et al., 2005; Berger & Ladd, 2012

	CLASP1	Exclusion of exon 19	DM1	Wagner et al., 2016; El Boujnouni et al., 2025
	DMPK	Isoform shift (expanded CTG RNA sequesters MBNLs → formation of ribonuclear foci)	DM1	Kanadia et al., 2003 and Jiang et al., 2004
	MBNL1	Inclusion of exons 6 and 8 (brain) Inclusion of exons 5, 7 and 10 (skeletal muscle)	Both	Jiang et al., 2004; Dhaenens et al., 2008; Konieczny et al., 2014
	MBNL2	Inclusion of exons 7 and 8 (brain)	DM1	Nakamori et al., 2013 and Yamashita et al., 2012

Table 2. Key splicing events in Myotonic Dystrophy.

This table lists the 28 main genes that undergo aberrant splicing in DM1 and/or DM2. Genes marked with * are clinically significant because their mis-splicing directly contributes to hallmark disease symptoms. Genes are grouped by functional category. 'Both' = mis-splicing observed in DM1 and DM2; 'DM1' = only reported in DM1). (Information from Nakamori et al., 2012, Wang et al., 2019, and López-Martínez et al., 2020).

Collectively, the splicing alterations summarised in Table 2 affect diverse physiological systems, reflecting the multisystemic nature of DM1 and DM2. In skeletal muscle, aberrant inclusion of intron 1 and exon 7a in *CLCN1* reduces chloride conductance and causes myotonia (Charlet et al., 2002; Mankodi et al., 2002). Similarly, exclusion of *BIN1* exon 11 and *RYR1* exon 70 impairs excitation-contraction coupling (Kimura et al., 2005; Fugier et al., 2011). Metabolic complications, such as insulin resistance, arise from the exclusion of exon 11 in *INSR* (Savkur et al., 2001). Cardiac manifestations, including conduction defects and arrhythmias, are associated with mis-splicing of *TNNT2* (foetal exon 5 inclusion), *SCN5A* (foetal exon 6a inclusion), and *TTN* (Lin, 2006; Freyermuth et al., 2016; Bosè et al., 2019). In the CNS, exclusion of exons in MAPT and *NMDAR1* may contribute to cognitive dysfunction and neuropsychiatric features (Jiang, 2004; Goodwin et al., 2015). Thus, sequestration of MBNL proteins results in widespread reversions to foetal isoforms, providing a unifying molecular mechanism for the multisystemic phenotype of myotonic dystrophy and form the foundation for investigating additional layers of post-transcriptional regulation.

1.3 Epitranscriptomic modifications in DM1

Epitranscriptomic modifications are reversible and stable chemical changes to RNA that do not alter the nucleotide sequence but can influence RNA fate and gene expression (Chokkalla et al., 2021; Imbriano et al., 2023). To date, more than 170 chemically distinct epitranscriptomic modifications have been identified across all classes of RNAs, including transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), messenger RNAs (mRNAs), and noncoding RNAs such as long noncoding RNAs (lncRNAs) (Sordyl et al., 2026; MODOMICS database). These modifications occur on adenosine, cytidine, guanosine, uridine or ribose by adding functional groups (e.g. methyl, acyl, thioalkyl). The chemical diversity of these modifications confers distinct regulatory functions that are highly cell-type-specific and context-dependent.

Epitranscriptomic modifications are primarily regulated by three classes of proteins: ‘writers’, ‘readers’, and ‘erasers’, which add, interpret or remove RNA modifications, respectively (see Figure 5; Meyer & Jaffrey, 2014). Writers catalyse the installation of specific chemical marks on RNA, thereby generating recognition sites for the readers. Readers selectively recognise these modified nucleotides and translate the chemical marks into functional outcomes by modulating RNA stability, localisation, splicing, nuclear export, or translational efficiency. In contrast, erasers remove these modifications, enabling dynamic and reversible regulation of RNA in response to cellular stimuli. Through the coordinated activity of writers, readers, and erasers, epitranscriptomic regulation provides an additional layer of post-transcriptional control that is essential for development, cellular stress response, and disease (see Figure 5; Visconti et al., 2021). Given that DM1 is characterised by extensive disruption of RNA-protein interactions and post-transcriptional gene regulation, it is plausible that the function of epitranscriptomic regulators, including RNA modification readers, writers and erasers, may be altered in this disease context.

Furthermore, RNA modifications are subject to both spatial and temporal regulation (Akhtar et al.,2021). Modifications can be deposited co-transcriptionally, influencing RNA processing events such as splicing and export, or post-transcriptionally, where

they regulate RNA stability and translation in the cytoplasm. This regulation allows cells to rapidly fine-tune gene expression without altering the underlying DNA sequence. Disruption of these tightly regulated epitranscriptomic processes has been increasingly implicated in human disease, including neuromuscular and neurodegenerative disorders such as myotonic dystrophy type 1 (DM1).

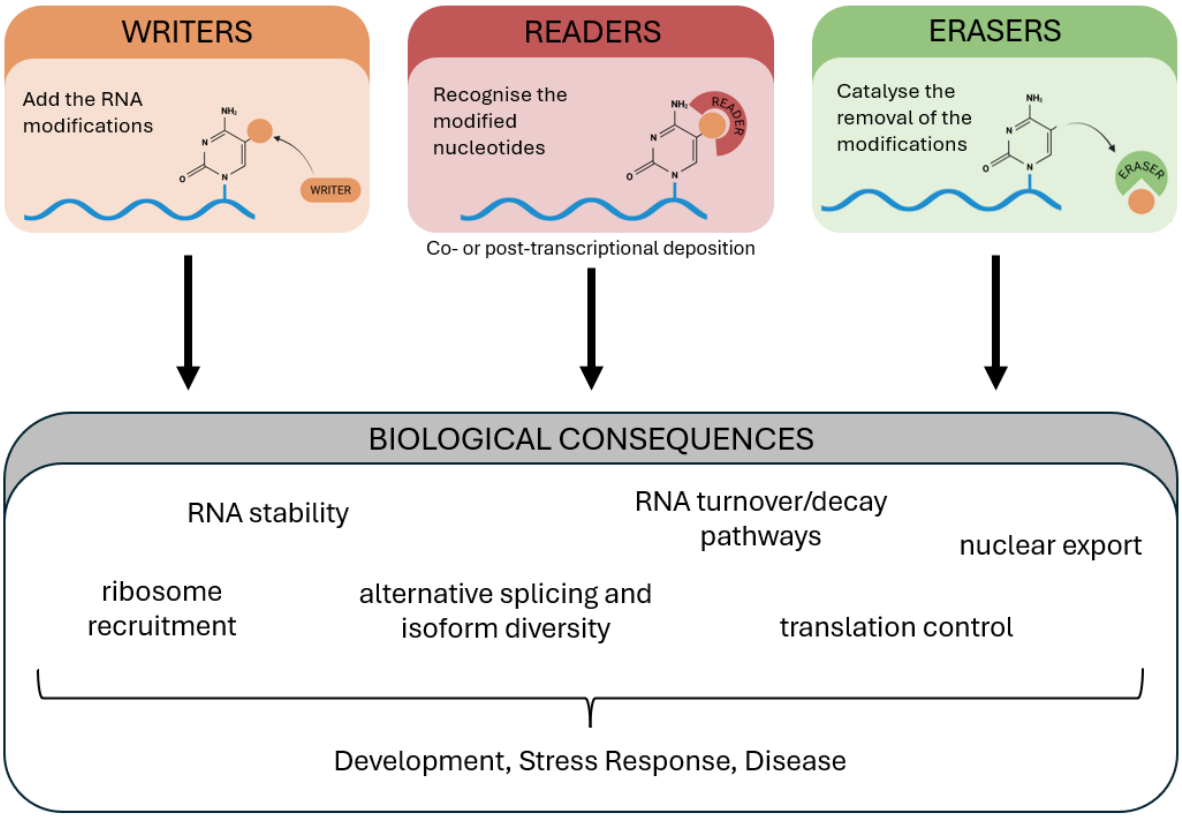

Figure 5. Overview of epitranscriptomic regulation and its biological consequences.

There are over 170 recognised epitranscriptomic modifications which can be found on the nucleosides or ribose in RNA molecules (Sordyl et al., 2026). These modifications are regulated by writers, readers and erasers. Through coordinated and spatially and temporally regulated activity, these factors influence multiple stages of the RNA lifecycle, including stability, splicing, nuclear export, translation, and decay, thereby contributing to development, cellular stress responses, and disease. Self-made diagram based upon the information in Section 1.3.

Epitranscriptomic modifications can be investigated by reducing the expression of the associated ‘writer’, ‘reader’, or ‘eraser’ proteins (Visconti et al., 2021). This approach enables the identification of these genes and allows assessment of

whether loss or reduction of these factors alters modification-dependent RNA processing events. Where no effect is observed, this may suggest that the modification is not dependent on the targeted factor under the tested experimental conditions, the modification is not present, or that compensatory mechanisms exist.

Genetic knockdown (KD) is an experimental technique that reduces or transiently inhibits gene expression without altering the underlying DNA sequence, most commonly via RNA interference mechanisms (Hannon, 2002). By using small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs), target mRNAs are degraded or translationally repressed, resulting in partial and often reversible depletion of protein levels. KD is therefore widely used to investigate gene function while preserving cellular viability and allowing recovery.

In contrast, genetic knockout (KO) permanently disrupts gene function through targeted modification of genomic DNA, typically using genome-editing technologies such as CRISPR-Cas9, zinc finger nucleases (ZFNs), and transcription activator-like effector nucleases (TALENs) (Gaj et al., 2013). CRISPR-Cas9 utilises a guide RNA to direct the Cas9 endonuclease to a complementary genomic sequence adjacent to a protospacer-adjacent motif (PAM), resulting in a double-strand break that is repaired by error-prone non-homologous end joining. ZFNs and TALENs achieve similar outcomes through protein-DNA recognition mechanisms. Therefore, these approaches are particularly suited to assessing the long-term consequences of gene ablation.

1.3.1 Focusing on the m^5C modification

Currently, the mechanisms supporting mRNA stability in DM1 remain unclear, particularly given the hundreds of RNA modifications that may be involved. Although direct evidence for epitranscriptomic modification of the expanded DMPK transcript is currently lacking, studies on other repeat-expanded RNAs suggest that pathogenic repeat transcripts can be subject to RNA modification, raising the possibility that similar mechanisms may operate in DM1. Among these, N6-methyladenosine (m^6A)

and 5-methylcytosine (m^5C) are of particular interest because they are two of the most studied modifications (Chokkalla et al., 2020).

m^5C is a novel and abundant modification found in DNA, mRNA, tRNA and lncRNA, particularly in non-coding RNAs at the 3' UTR, which is notable given that the CUG repeat expansion in DMPK occurs in the same region (Yang et al., 2017; Soltanzadeh, 2022). The m^5C modification involves the methylation of the C-5 on cytosine. It is frequently found in CG-rich regions, immediately downstream of translation initiation sites, exhibiting dynamic, tissue-specific and evolutionarily conserved distribution across mammalian transcriptomes (Yang et al., 2017). Key regulators of m^5C include NSUN and DNMT methyltransferase family members, RNA-binding proteins such as ALYREF, YBX1, and RAD52, and TET family demethylases (Huang et al., 2024). Previous studies suggest that m^5C influences mRNA localisation, stability, and translation in a range of biological processes, including stress responses, development, tumorigenesis and embryonic genome activation (Chen et al., 2019; Yang et al., 2019; Tang et al., 2020; Zou et al., 2020). It may also play tissue-specific roles in lipid metabolism and muscle development (Liu et al., 2022).

On the other hand, m^6A is the most prevalent internal cotranscriptional RNA modification in higher eukaryotic organisms (Jiang et al., 2021). This modification determines cell fate during the endothelial-to-hematopoietic transition, mRNA export and RNA metabolism, tumorigenesis, and spermatogenesis via the p53 signalling pathway (Zheng et al., 2013; Zhang et al., 2017). However, its direct role in DM1 seems unlikely because the methylation occurs at the N-6 of adenine, and there is no adenine in the DM1 CUG repeat expansion (Desrosiers et al., 1974; Zhong et al., 2020). Therefore, m^5C is hypothesised to be a more plausible epitranscriptomic modification to the development and severity of DM1 phenotypes.

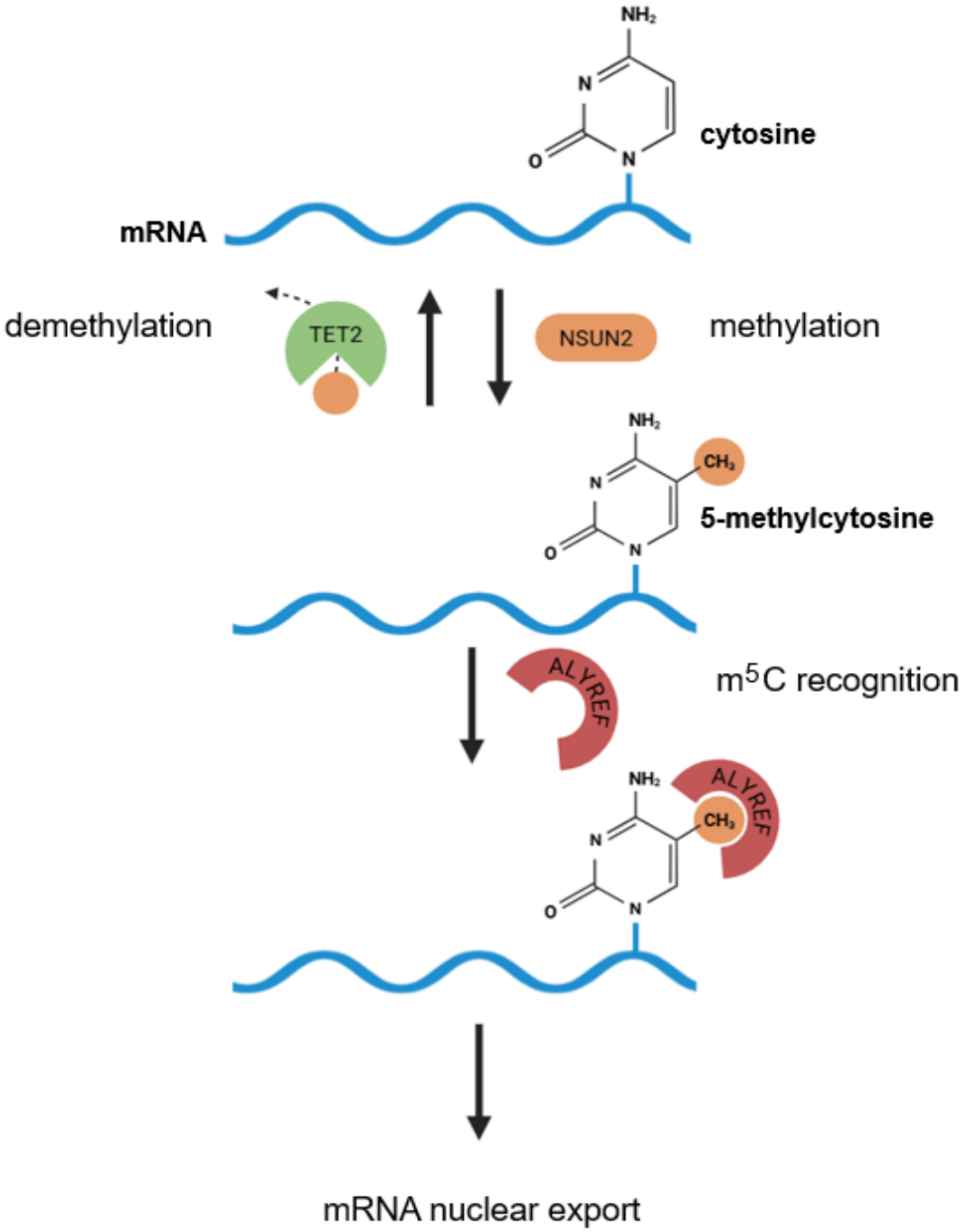

Figure 6. Proposed epitranscriptomic modification in DM1.

For this research project, the 'writer,' 'reader,' and 'eraser' are hypothetically attributed to the proteins encoded by *NSUN2*, *ALYREF*, and *TET2*, respectively. NSUN2 (orange) is proposed to catalyse the addition of the m^5C modification to cytosine residues on mRNA. The modified transcript can then undergo demethylation by TET2 or m^5C be recognised by ALYREF, a core adaptor of the transcription-export (TREX) complex. ALYREF binding is proposed to promote recruitment of the mRNA export machinery, facilitating TREX-mediated nuclear export of m^5C-modified transcripts through the nuclear pore complex. (Self-made diagram using Biorender.com, based on the information written in Section 1.3).

1.3.2 *NSUN2*

NOP2/Sun RNA methyltransferase 2 (NSUN2), also called *Misu,* is located on chromosome 5p.15.31-33 (Brzezicha et al., 2006; Frye et al., 2010). The encoded protein, NSUN2 (also called RNA C(5)-methyltransferase), consists of 767 amino acids and predominantly localises to the nucleolus, playing an essential role in cell proliferation, differentiation, and early development (Frye & Watt, 2006). By interacting with the nucleolar and spindle-associated protein, NuSAP, NSUN2 stabilises the mitotic spindle in rapidly dividing cells, whilst also contributing to protein synthesis through tRNA maintenance (Hussain et al., 2009; Van Haute et al., 2019). Genetic studies have also highlighted the physiological importance of NSUN2 in KO mice, which, although viable, displayed reduced body size, male sterility, cyclic alopecia, and neurological defects (Blanco et al., 2011; Blanco et al., 2014). In humans, homozygous NSUN2 mutations have been linked to intellectual disability and neurodevelopmental disorders (Abbasi-Moheb et al., 2012; Yang et al., 2024).

NSUN2 plays an important role in RNA modification. It catalyses the methylation of cytosine to 5-methylcytosine (m^5C) in a range of RNA substrates, including tRNAs, mRNAs, and non-coding RNAs – shown in Figure 6 (Yang et al., 2017; Van Haute et al., 2019). These modifications regulate RNA metabolism by promoting transcript stability, enhancing translation efficiency, and preventing RNA decay (Zhang et al., 2012; Chen et al., 2019). Furthermore, the m^5C modifications can label transcripts, which are recognised and exported from the nucleus by the export adaptor, ALYREF (Yang et al., 2017).

1.3.3 ALYREF

ALYREF, also called *THOC4*, encodes the ubiquitously expressed RNA-binding protein, ALY/REF export factor (ALYREF). ALYREF functions as a molecular chaperone and dual-region export adapter involved in the efficient nuclear export of spliced and, to a lesser extent, unspliced mRNA, ensuring coordinated export machinery recruitment at the 5′ and 3′ ends (Muravenko et al., 2000; Shi et al., 2017; Fan et al., 2019).

ALYREF has been implicated as a modifier of neurodegeneration in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), see Figure 3 (Berson et al., 2019; Kow et al., 2022). In Drosophila, knockdown of Ref1 (the ALYREF orthologue) suppressed pathogenic phenotypes caused by TDP-43 and expanded G4C2 repeats, primarily through reduced expression of the pathogenic transgenes, whilst ALYREF was upregulated in disease models and in neurons derived from ALS patients carrying C9orf72 repeat expansions (Berson et al., 2019). Although RNA foci formation was not directly assessed in this model, these findings demonstrate that ALYREF influences the cellular handling of repeat-containing transcripts. Complementary work in Caenorhabditis elegans showed that loss of *aly* homologues (*aly-2* and *aly-3*) similarly suppressed pathogenic phenotypes in Tau and TDP-43 models of neurodegeneration, through distinct mechanisms (Kow et al., 2022). Collectively, these studies establish ALYREF as a conserved disease modifier in neurodegeneration, while highlighting that its effects may arise from altered RNA processing or transcript regulation rather than direct binding to expanded repeat RNA.

Loss-of-function studies by Fan et al. (2019) revealed that *ALYREF* depletion results in the nuclear accumulation of non-polyadenylated mRNAs and a reduction in cytoplasmic transcript levels. In the context of DM1, *ALYREF* KD is predicted to exacerbate ribonuclear foci formation by disrupting transcription/export (TREX)-mediated mRNA export and promoting nuclear sequestration of mRNA substrates (Shi et al., 2017).

As an m⁵C ‘reader’, ALYREF recognises and binds directly to m⁵C-modified mRNAs through its RNA-binding domains, which preferentially interact with methylated cytosine residues installed by methyltransferases such as NSUN2 (Yang et al., 2017). The presence of the methyl group is proposed to alter local RNA chemical properties and/or secondary structure, thereby enhancing the stability of ALYREF-RNA interactions relative to unmodified transcripts (Shi et al., 2017). This enhanced binding affinity enables selective association with specific mRNA populations, particularly transcripts encoding proteins involved in translation, cell cycle

progression, and transcriptional regulation, thereby prioritising their nuclear export via the TREX pathway (Shi et al., 2017; Yang et al., 2017).

Given these roles in RNA export and neurodegeneration, ALYREF dysregulation may similarly influence pathogenic RNA foci in DM1, making it a candidate disease modifier in this context.

1.3.4 Functional Interplay Between NSUN2 and ALYREF in m⁵C-Dependent RNA Regulation.

NSUN2 and ALYREF are well-established as the 'writer' and 'reader' of m⁵C epitranscriptomic modifications, respectively. Their cooperative function has been implicated in various disease contexts, particularly cancer.

A recent study by Zuo et al. (2023) demonstrated that NSUN2-mediated m⁵C methylation promotes retinoblastoma progression by stabilising and enhancing the expression of phosphoribosylformylglycinamidine synthase (PFAS), a key enzyme in purine biosynthesis. ALYREF binds to the m⁵C-modified PFAS transcript, increasing its RNA stability and expression. Elevated m⁵C levels via NSUN2 were shown to enhance cancer cell proliferation and metastasis both *in vitro* and *in vivo*.

Similarly, Wang et al. (2023) reported that NSUN2 and ALYREF co-regulate urothelial bladder cancer malignancy through the activation of hypermethylated m⁵C-modified oncogenic RNAs. Both regulators were significantly upregulated in tumour tissues and were positively associated with several oncogenic pathways, suggesting a broader oncogenic role for this m⁵C-dependent regulatory axis.

Moreover, Yang et al. (2025) demonstrated that the NSUN2/ALYREF-driven m⁵C modification enhances PD-L1 expression in non-small cell lung cancer, thereby

facilitating tumour immune evasion. This further underscores the pathological relevance of NSUN2–ALYREF cooperation across diverse cancers.

Beyond cancer, NSUN2 and ALYREF have also been implicated in cellular stress responses. Huang et al. (2024) found that NSUN2 increases m⁵C methylation on Nrf2 mRNA, promoting liver cell susceptibility to Doxorubicin-induced damage by suppressing antioxidant stress responses. Notably, this regulatory effect was dependent on ALYREF, highlighting the functional interdependence of these two proteins.

Mechanistic insights into their interaction have been provided by Yang et al. (2017), who demonstrated that NSUN2 silencing leads to an increase in nuclear retention of ALYREF and a reduction in RNA-binding affinity. In contrast, ALYREF KD does not appear to affect NSUN2's RNA-binding capability, suggesting that ALYREF's RNA-binding and subcellular localisation are regulated in an m⁵C-dependent manner by NSUN2's catalytic activity.

These examples highlight the potential of NSUN2 and ALYREF to regulate RNA fate through the m⁵C pathway in diverse disease contexts. Although their involvement in DM1 has not yet been demonstrated, the evidence from other systems raises the possibility that analogous mechanisms may contribute to RNA dysregulation in DM1, thereby justifying further investigation of the epitranscriptomic m⁵C-modification.

1.3.5 TET2

Ten-eleven translocation methylcytosine deoxygenase *(TET2)*, first identified in 2009, is located on chromosome 4q24 (Delhommeau et al., 2009; Jankowska et al., 2009). It is one of three conserved TET genes (*TET1, TET2* and *TET3*) that regulate genome demethylation and affect histone modification (Tahiliani et al., 2009; Ito et al., 2011; Deplus et al., 2013; Kong et al., 2016; Joshi et al., 2022).

Current findings define TET2 as a multifunctional dioxygenase that regulates both DNA and RNA cytosine modifications. By linking transcriptional, post-transcriptional, and epigenetic control, TET2 may act as the ‘eraser’ of m5C modifications in DM1, as predicted in Figure 6.

TET proteins have a double-stranded β-helix fold (DSβH) domain, which contains three Fe^{2+}-binding sites and one α-ketoglutarate binding site, stabilised by zinc fingers that mediate CpG recognition and substrate binding (Hu et al., 2013). This structure enables TET2 to flip the target base into its catalytic pocket, orient the m^5C methyl group towards Fe^{2+} and drive iterative oxidation of m^5C to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Ito et al., 2011; Hu et al., 2013). Through this mechanism, TET2 drives active DNA demethylation and maintains epigenetic plasticity during development and differentiation (Handal et al., 2024).

Somatic TET2 mutations occur in ~15–25% of myeloid malignancies, particularly chronic myelomonocytic leukaemia and secondary acute myeloid leukaemia, driving proliferative, myelomonocytic phenotypes (Delhommeau et al., 2009; Iyama et al., 2025). TET2 loss promotes clonal haematopoiesis, leukemic blast migration, and leukaemia stem cell homing/self-renewal (Li et al., 2024). Mechanistically, TET2 deficiency causes aberrant m^5C accumulation in TSPAN13 mRNA, stabilised by YBX1, thereby enhancing leukemogenic potential (Li et al., 2024).

Beyond cancer, TET2 expression is linked to neuronal and stress responses. In chronic mild stress, hippocampal TET2 regulates mRNA stability via 5hmC–Upf1, supporting stress adaptation (Xia et al., 2023). In DM1, TET2 declines during differentiation, shifting from reversible to irreversible heterochromatin states. High TET2 in undifferentiated cells permits reversal of DNA methylation/H3K9me3 after repeat excision, whereas low TET2 in differentiated cells locks these marks, highlighting its role in epigenetic plasticity (Handal et al., 2024). In this context,

reduced TET2 activity may exacerbate NSUN2/ALYREF-mediated m5C accumulation, promoting persistent RNA dysregulation in DM1.

1.3.6 Current understanding and knowledge gaps of m^5C in DM1

Epitranscriptomic regulation has emerged as an important layer of post-transcriptional gene control, with modifications such as m^5C influencing RNA stability, localisation, translation, and nuclear export through the coordinated activity of writer, reader, and eraser proteins. In multiple biological and disease contexts, the addition of m^5C by NSUN family methyltransferases and its recognition by the export adaptor ALYREF have been shown to regulate RNA fate via the TREX-mediated export pathway. However, despite the central role of RNA dysregulation in DM1, direct evidence for epitranscriptomic modification of the expanded DMPK transcript is currently lacking, and the contribution of m^5C-dependent mechanisms to RNA foci formation and stability in DM1 remains unexplored. It is unknown whether altered activity of NSUN2 or ALYREF influences the pathogenic repeat-containing RNAs. Addressing these gaps is essential to determine whether epitranscriptomic pathways contribute to RNA toxicity in DM1 and to establish m^5C regulators as potential disease modifiers or therapeutic targets.

1.4 Aims of this study

This project aims to examine the role of *ALYREF* in the processing of the *DMPK* expansion repeat mRNA. The effects of transfecting KB-TeloMyoD fibroblast cells (a DM1 immortalised patient-derived cell line) with gene-specific siRNAs to knock down ALYREF (the reader) will be observed, due to its predicted involvement in the m^5C modification mechanism, as outlined in Sections 1.3.2. ALYREF protein KD will first be confirmed using western blotting with normalisation to the housekeeping protein, GAPDH. After KD is confirmed, changes in nuclear foci intensity, size, and number per nucleus will then be assessed using a FISH assay. Ribonuclear foci are a key indicator of the DM1 phenotype, as discussed in Section 1.2.3, so their analysis will help to evaluate the impact of the m^5C modification on our DM1 cell line.

2 Methods

2.1 Tissue Culture

2.1.1 Cell lines

SB-TeloMyoD fibroblast cells were derived from an unaffected male individual to be used as a control. The SB cell line has a maximum of 21 CTG repeats, which is within the normal, unaffected range (see Section 1.1.3).

Immortalised human KB-TeloMyoD fibroblast cells and XTB fibroblast cells were derived from two male patients with DM1 and were cultured to test the effect of potential therapeutic targets. The KB cell line has 300-400 CTG repeats; however, it has been reported to have extensive variation, showing readings of up to ~1500 CTG repeats (Xing et al., 2023). The XTB cell line has 800 CTG repeats. Both cell lines are classed as pathogenic (See Section 1.1.3).

2.1.2 Cell culture

Cells were cultured at 37°C with 5% CO_2. Dulbecco's Modified Eagle's Medium (DMEM; Gibco, 41965-039), supplemented with 15% foetal bovine serum (Sigma-Aldrich, F7524) and 2.4% penicillin-streptomycin (Sigma-Aldrich, P4333), was used across all cell types. Cells were split under a sterile laminar hood once the confluence exceeded 60%. T-75 flasks (ThermoFisher, 156800) were used for culturing. The old media was removed, and the cells were washed once with Dulbecco's phosphate-buffered saline (PBS; Gibco, 14190-086) and incubated with 0.5 mL of Trypsin-EDTA (Gibco, 25300054) for 5 minutes. Cells were resuspended in fresh media and split into new labelled T-75 flasks with a total volume of 20 mL.

2.1.3 Freezing cell lines

The media was removed; cells were washed once in PBS and incubated with 0.5 mL of Trypsin-EDTA for 5 minutes at 37 °C with 5% CO_2. Cells were resuspended in 10

mL of DMEM and transferred to a sterile 15 mL Falcon tube, then centrifuged for 5 minutes at 1000 g. The supernatant was removed, and the cell pellet was resuspended in 10% dimethyl sulfoxide (DMSO, Sigma-Aldrich, D2650) solution made up in complete media and transferred to a screw-capped vial. The vial was labelled and stored at -80 °C for 24 hours before being transferred to liquid nitrogen.

2.1.4 Defrosting cell lines

Vials were removed from liquid nitrogen, thawed in their screw-capped vials using a 37 °C water bath, and resuspended in 1 mL of pre-warmed DMEM. Cells were immediately transferred to a sterile 15 mL Falcon tube containing 10 mL of room-temperature DMEM (rapid transfer minimises cell death caused by DMSO toxicity) and centrifuged for 1 minute at 1000 g. The supernatant was carefully removed, and the pelleted cells were resuspended in 10 mL of DMEM and transferred to a T-75 flask with 10 mL of DMEM. The flask was labelled and placed in an incubator at 37 °C and 5% CO_2 overnight.

2.1.5 siRNA transfection of cell lines by electroporation

For siRNA-mediated ALYREF KD, *ALYREF*-specific siRNAs and a non-specific siRNA control were sourced from Dharmacon™ (see Table 3 for sequences). KB cells were grown to 60-70% confluency, and ~2 million cells per T75 flask were assumed (based upon previous work in the lab). Two flasks were used per condition to optimise cell survival following electroporation. Electroporation transiently permeabilises the plasma membrane with an electrical pulse, which, while enabling siRNA uptake, also causes acute membrane damage and induces metabolic and oxidative stress, resulting in increased cell death. Therefore, the use of two flasks ensured that a sufficient number of successfully transfected cells survived for subsequent experiments. The cells were transfected with *ALYREF*-specific siRNAs, the non-specific siRNA or siGLO Green Transfection Indicator (Dharmacon™, D-001630-01-05).

2.1.5.1 Preparation

Adherent cells were washed once in PBS and incubated with 0.5 mL Trypsin-EDTA for 5 minutes at 37 °C. Following incubation, cells were resuspended in 9.5 mL DMEM containing FBS to inactivate trypsin via serum protease inhibitors, rapidly stopping its proteolytic activity and preventing further proteolytic damage to cell-surface proteins. The cell suspension was transferred to a sterile 15 mL Falcon tube and centrifuged at 1000 g for 5 minutes. The resulting pellet was washed and resuspended in 10 mL PBS deficient in calcium and magnesium and centrifuged at 1000 g for 5 minutes to ensure all excess salts were removed before electroporation.

2.1.5.2 Transfection

For each sample, 100 µL of Nucleofector mix (82 µL Nucleofector solution + 18 µL Supplement 1 solution) from the Amaxa Basic Nucleofector Kit for primary mammalian fibroblasts (Lonza, VPI-1002) was added to a sterile Eppendorf tube, and 5 µL of a 20 µM stock solution of non-specific siRNA, target-specific siRNA or transfection control was added. One after the other, the supernatant was removed from the Falcon tube, the cell pellet was resuspended in the corresponding Nucleofector Mix, and the cell suspension was transferred to a cuvette for electroporation (Nucleofector Amaxa Electroporator set to the NHDF Human Adult High-Efficiency U-023 programme). Once electroporated, cell suspensions were transferred to the corresponding T75 flask containing 20 mL FGM®-2 Fibroblast Growth Medium-2 BulletKit® (Lonza, CC-3132). Cells were incubated at 37 °C with 5% CO_2 for ~48 hours.

siRNA	Sequence
ON-TARGETplus SMARTpool Human THOC4 Catalog: L-012078-00-0005 Lot: 240607	UCUCAGACGCCGAUAUUCA GUUAAACAGACCAGCAAAU GGAACUCUUUGCUGAAUUU CAAAACAACUUCCCGACAA
siGenome™ Control siRNA Non-Targeting siRNA #3 Catalog: D-001210-03-05 Lot: 3391572	AUGUAUUGGCCUGUAUUAG
Negative Control SpCas9 gRNA #1 SKU: 063-1010-000-000	GCACUACCAGAGCUAACUCA
Negative Control SpCas9 gRNA #2 SKU: 063-1011-000-000	sequence unknown

Table 3: ALYREF Dharmacon™ and NSUN2 Synthego reagents for siRNA transfection and CRISPR-Cas9 experiments.

This table shows the full details for the ALYREF KD siRNAs and the scrambled siRNA control (scramble 3) used in this project and used in the data presented in Section 3.4. The SMARTpool Human THOC4 (ALYREF) siRNA is composed of four different siRNAs. Scrambled gRNAs (scramble 1 and scramble 2) from the NSUN2 experiments (Section 3.5) are also shown.

2.2 SDS-PAGE and western blot

2.2.1 Sample preparation

For SDS-PAGE and western blot experiments, all cell lines mentioned in Section 2.1.1 were grown to 100% confluency. Under a sterile laminar hood, cells were washed once in PBS and incubated in 0.5 mL of Trypsin-EDTA for 5 minutes at 37 °C with 5% CO_2. Cells were neutralised with 10 mL DMEM, transferred to a 15 mL Falcon tube and spun at 1000 g for 5 minutes. The media was removed, and the cell pellet was resuspended in 1 mL PBS and transferred to an Eppendorf tube to spin down at 2300 g for 5 minutes at 4 °C. The PBS was removed, and the pellet was resuspended in 150 µL Radioimmunoprecipitation (RIPA) lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 1% (v/v) NP-40, 0.5% (w/v) sodium deoxycholate, 0.1% (w/v) SDS, 1 mM EDTA) and 1.5 µL 100X protease inhibitor (Cell Signalling Technology, 5872). The Eppendorf tube was placed at -80 °C for >2 hours.

Following this period of freezing, three cycles of freeze-thawing were performed before centrifuging at 13,400 g for 10 minutes at 4 °C. The supernatant was aliquoted into labelled Eppendorf tubes and stored at -80 °C. Using the Micro BCA™ Protein Assay Kit (ThermoFisher, 23235), the absorbance was measured at 562 nm on a spectrophotometer and protein concentration was calculated using the linear regression equation from the standard BSA vs absorbance curve.

Extracted protein (15/30 µg total protein) combined with 4X lithium dodecyl sulphate (LDS) sample buffer (NuPAGE, NP007) and 10X sample reducing agent (NuPAGE, NP0004). The Eppendorf tubes were placed on a heating block at 95-100 °C for 5 minutes and cooled at room temperature.

2.2.2 SDS-PAGE

Samples were loaded onto a 10% 1 mm x 10 well Bis-Tris pre-made polyacrylamide gel (Invitrogen, NP0301), along with 5 µl of SeeBlue marker (Invitrogen, LC5625) and 2 µL of Western MagicMark™ (Invitrogen, LC5602) in the XCell SureLock Mini-cell electrophoresis chamber with 800 mL 1X running buffer (5% (v/v) 20X MOPS, 95% (v/v) dH_2O). The gel was run on the 'NuPAGE GEL' setting at 200V for 50 minutes at room temperature.

2.2.3 Electrophoretic transfer to PVDF membrane

The polyvinylidene fluoride (PVDF) membrane (ThermoFisher, LC2005) was activated in 100% methanol for 1 minute, washed twice in dH_2O and submerged in transfer buffer (5% (v/v) 20X NuPAGE transfer buffer, 10% (v/v) methanol, 75% (v/v) dH_2O). The gel was trimmed and sandwiched between the PVDF membrane, two filter papers and four transfer buffer-soaked sponges (Figure 7) and returned to the electrophoresis chamber. Transfer buffer was added to the inner chamber, and dH_2O was added to the outer chamber. The transfer was run on the 'NuPAGE BLOT' setting at 30V for 1 hour at room temperature.

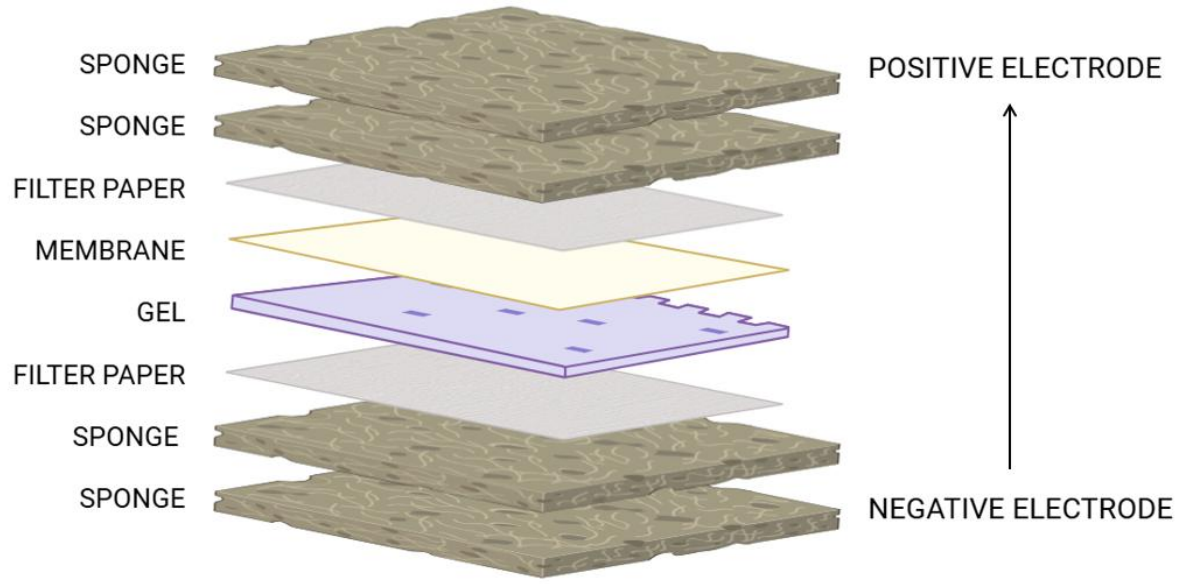

Figure 7: Western blot sandwich formation for transfer.

The schematic diagram represents the 'sandwich' built to achieve optimal transfer of the proteins in the gel to the membrane. The movement of the electrical current is represented by the arrow on the right, moving from the negative electrode to the positive electrode, making the correct placement of the gel and membrane crucial. (Self-made diagram using Biorender.com).

2.2.4 Immunodetection

The sandwich was removed from the chamber, and the membrane was placed into a 50 mL Falcon tube and washed three times in dH_2O before incubating in 10 mL of 5% milk powder (ChemCruz, sc-2324) in Tris-buffered saline containing 0.05% Tween-20 (TBST) for 2 hours at room temperature. Following incubation, the membrane was washed three times in TBST for 5 minutes and then incubated overnight at 4 °C in 10 mL of 5% milk in TBST containing a 1:1000 dilution of the primary antibody (Anti-ALY/Ref antibody, Abcam, ab202894). The milk/primary antibody mixture was removed, and the membrane was washed three times in TBST for 10 minutes. The membrane was incubated in 10 mL of 5% milk in TBST containing a 1:5000 dilution of goat anti-rabbit IgL (HRP) antibody (Abcam, ab205718) for 30 minutes at room temperature. Following incubation, the milk/antibody solution was removed, and the membrane was washed three times in TBST for 10 minutes. The membrane was then incubated in 3 mL of detection solution (1:1 Reagent A and Reagent B) (Clarity™ Western ECL Substrate, Bio-Rad

170-5061) for 5 minutes at room temperature. The detection solution was removed, and the membrane was sealed in plastic wrap before imaging with the Bio-Rad ChemiDoc MP Imaging System.

2.2.5 Re-probing the gel

Secondary staining with the loading control GAPDH was performed to ensure samples were consistently loaded onto the gel. The membrane was washed twice in dH_2O for 30 minutes and once in TBST for 10 minutes to ensure antibodies were removed. The membrane was incubated for 30 minutes at room temperature in 10 mL of 5% milk in TBST containing a 1:50,000 dilution of anti-GAPDH antibody (Abcam, ab8245). After incubation, the membrane was washed three times in TBST for 10 minutes and incubated in 10 mL of 5% milk in TBST containing a 1:5000 dilution of goat anti-mouse IgL (HRP) antibody (Abcam, ab205719) for 30 minutes at room temperature. The membrane was then washed three times in TBST for 10 minutes and incubated in 3 mL of detection solution (1:1 Reagent A and Reagent B) (Clarity™ Western ECL Substrate, Bio-Rad, 170-5061) at room temperature for 5 minutes. The excess detection solution was removed, and the membrane was imaged using the Bio-Rad ChemiDoc MP Imaging System. Band intensities were quantified using ImageJ software.

2.3 Fluorescence *In Situ* Hybridisation (FISH)

2.3.1 Plating of Cells for FISH

Using a haemocytometer, wild-type KB, *ALYREF* KD, non-specific siRNA control cells, and siGLO transfection control cells were plated in a Greiner black, clear-bottom 96-well plate (Greiner, 655090) at a density of ~10,000 cells per well and incubated at 37 °C for 48 hours in a humidified chamber.

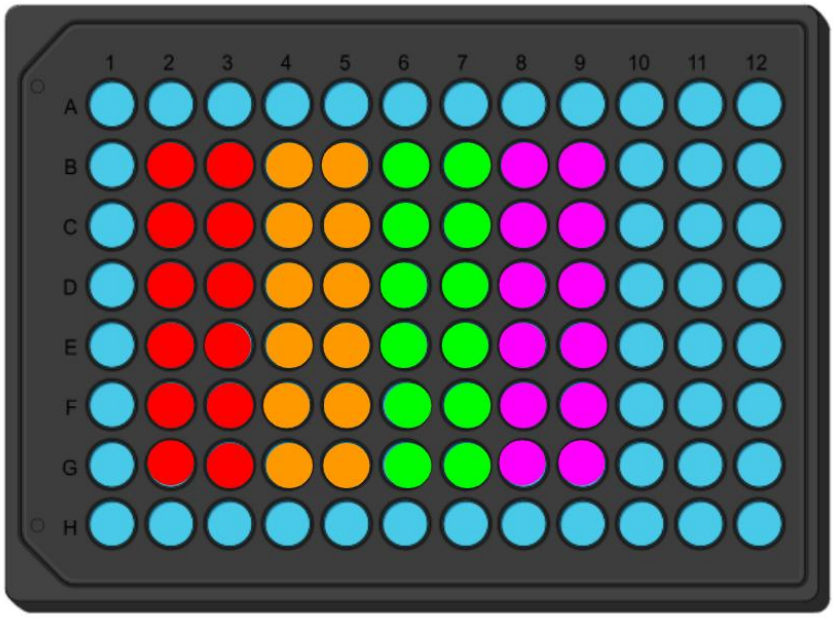

Figure 8: 96-well Plate Layout for ALYREF FISH.

In a Greiner black clear-bottomed 96-well plate, ~48 hours before Day 1 of FISH, cells were seeded at ~10,000 per well and made up to a total volume of 200 µL with complete DMEM. Four different conditions were seeded: KB cells (red), KB cells transfected with non-specific siRNA (orange), KB cells transfected with target siRNAs (green) and KB cells transfected with the siGLO Green Transfection Indicator. PBS was added to the surrounding wells.

2.3.2 Hybridisation

Following incubation, the culture medium was removed, and cells were washed three times with PBS. This was carried out by adding 200 µL PBS per well, incubating for 5 minutes per wash, and then removing the PBS before the next wash.

Cells were then fixed in 4% PFA (1X PBS in DEPC H_2O) and incubated for 30 minutes at room temperature. Following incubation, PFA was removed, and the cells were washed three times for 5 minutes in PBS. Cells were permeabilised in 80% ethanol for 2-3 hours at room temperature.

The ethanol was removed, and the cells were washed three times for 5 minutes in PBS. 50 µL of the pre-hybridisation solution (40% formamide, 10% SSC, 50% DEPC-treated H_2O) was added to each well, and cells were incubated at room temperature for 15 minutes. The pre-hybridisation solution was removed from columns 2-7 and replaced with 50 µL of hybridisation solution (4 mL pre-hybridisation solution, 40 µL of 10mg/mL SS-DNA (single-strand salmon testes DNA, Sigma-Aldrich, 68938-01-2), 8 µL of 20 mg/mL recombinant albumin (New England Biolabs, B9 200S), 4 µL of 200 mM VAC (ribonuclease vanadyl complexes, Sigma-Aldrich, R3380), 2 µL of 100 µg/mL Cy3-labelled $(CAG)_{10}$ probe (Eurofins Genomics, 36787403) and incubated overnight at 37 °C and 5% CO_2 in a humidified chamber. The plate was covered with foil to prevent probe degradation and remained covered where possible for method 3.3.3.

2.3.3 Plate imaging

The hybridisation solution was removed, and the cells were washed three times for 5 minutes in PBS supplemented with 10mM $MgCl_2$. Cells were incubated at room temperature in 100 µL Hoechst 33342 diluted 1:5000 in PBS supplemented with 10mM $MgCl_2$ per well for 15 minutes. The Hoechst stain was removed and replaced with 100 µL PBS supplemented with 10mM $MgCl_2$. Cells were visualised using a Zeiss Celldiscoverer 7 Microscope. Statistical analysis of extracted data was performed in GraphPad Prism. Plates were stored at 4°C, wrapped in foil.

3 Results

As discussed in the introduction (see Section 1.4), the main aim of this research project was to investigate the role of *ALYREF* in the processing of the *DMPK* expansion repeat mRNA. To address this, FISH was selected as the most appropriate assay to assess the effect of ALYREF KD on ribonuclear foci in DM1. FISH assays are readily adaptable to 96-well plates, allowing for multiple comparisons under different conditions within a single experiment. This approach enabled the assessment of three foci characteristics (area, fluorescent intensity and number per nucleus) across multiple cell lines (WT, KD and non-targeting siRNA). In this section, the non-targeting siRNA will be referred to as scramble 3. There are three different non-targeting siRNAs used: scramble 1, scramble 2 and scramble 3. The sequences can be found in Table 3.

Furthermore, unpublished data generated by two undergraduate project students (see Sections 3.4-3.5) were analysed and incorporated to support the findings presented here.

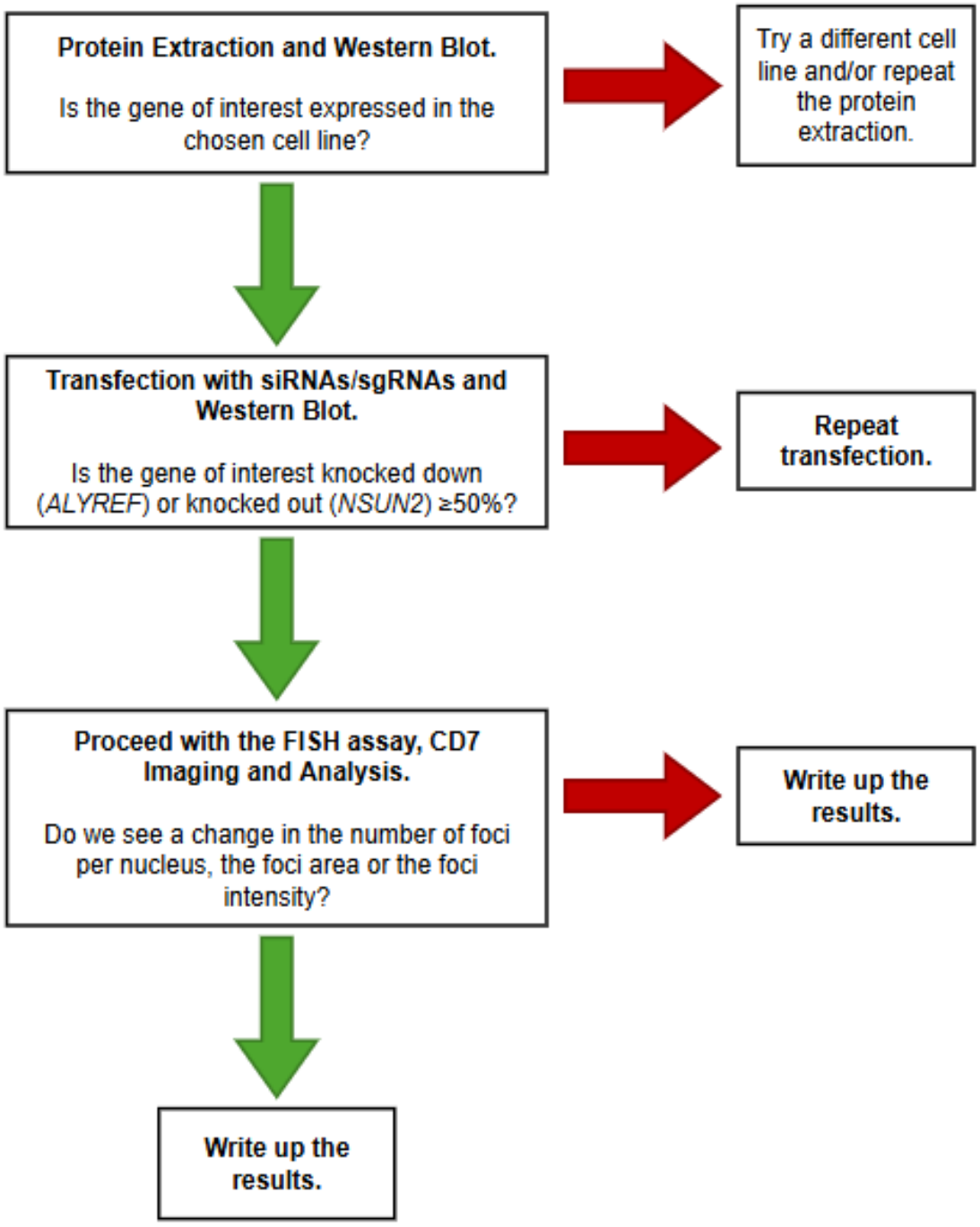

Figure 9. Overview of Experimental Design and Decision-Making Workflow.

This decision-making flowchart illustrates the process used to evaluate experimental quality and determine whether to proceed with certain experiments, and which data were included and which were excluded in this thesis. The arrows answer the questions in the box, green: YES, red: NO.

3.1 ALYREF is present in the DM1 cell lines

To assess the presence of ALYREF in the DM1 cell lines, protein lysates were prepared in duplicate from KB-TeloMyoD (KB) fibroblasts and XTB cells (both DM1 lines), as well as from SB-TeloMyoD (SB) fibroblasts (a non-DM1 control). The samples were extracted following the protocol set out in Section 2.2.1 and analysed by western blotting (see Sections 2.2.1-2.2.5).

Figure 10 shows ALYREF expression across all cell lines. Band identity was assigned based on migration relative to the molecular weight ladder, consistent with the predicted molecular weight of ALYREF (~27 kDa). Quantification of band intensities revealed that ALYREF/GAPDH ratios were consistent across most samples, with only a variation of ~0.13 between replicates, excluding XTB 1. The data from XTB 1, which appears abnormally high compared to the other samples, likely reflects a loading error.

The KB cell line was therefore chosen for further experimentation, given its reliable ALYREF expression and routine use in our laboratory.

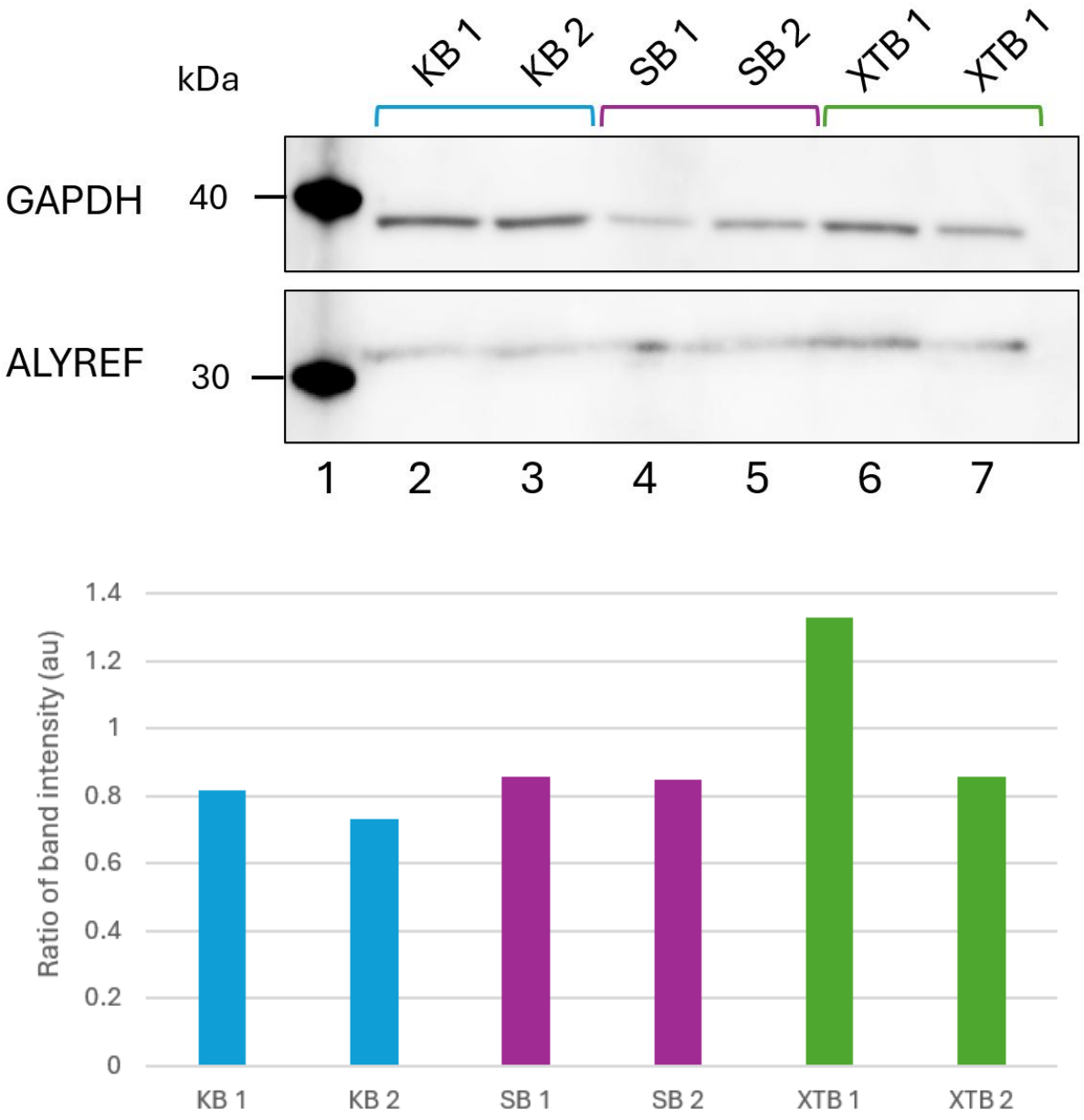

Figure 10. ALYREF is present in both DM1 cell lines.

Western blot showing the level of ALYREF expression in KB, XTB and SB cells. The expected molecular weight of ALYREF is ~27 kDa. Lane 1: MagicMark™ XP Western Protein Standard. For this assay: KB cells (lanes 2-3), SB cells (lanes 4-5), XTB cells (lanes 6-7). Samples were run in duplicate. All samples were loaded at 15 µg. GAPDH was used as a loading control (1:50,000). The expected molecular weight of GAPDH is ~37 kDa. The graph displays the normalised ALYREF/GAPDH band intensity ratios (arbitrary units), confirming comparable expression across all cell line protein samples.

3.2 ALYREF KD in transfected KB cell line

To ensure that the ALYREF-targeting siRNAs could effectively knock down ALYREF (>50%), KB cells were electroporated alone or transfected with ALYREF-specific siRNAs, scramble 3, or siGLO controls (see Section 2.1.5). WT KB cells without electroporation were also used to ensure that electroporation did not affect ALYREF expression.

Figure 11 shows that ALYREF expression was reduced by 89% and 87% in KB cells transfected with ALYREF-specific siRNAs. All western blot data (Figure 11, panel A) were analysed using ImageJ software and normalised against the control loading antibody, GAPDH, and then normalised against the scramble 3 (Figure 11, panel B).

KB cells transfected with scramble 3 (see Table 3 for siRNA sequence) were used as the primary control for ALYREF KD experiments, as this controls for non-specific effects of siRNA delivery and the electroporation/transfection procedure, allowing changes in ALYREF expression to be attributed specifically to ALYREF-targeting siRNAs.

Protein samples were loaded at both 15 µg and 30 µg, with 15 µg chosen for subsequent experiments due to clearer banding and slightly greater KD efficiency. GAPDH antibody dilution was also optimised from 1:5000 to 1:50,000 to improve band distinction.

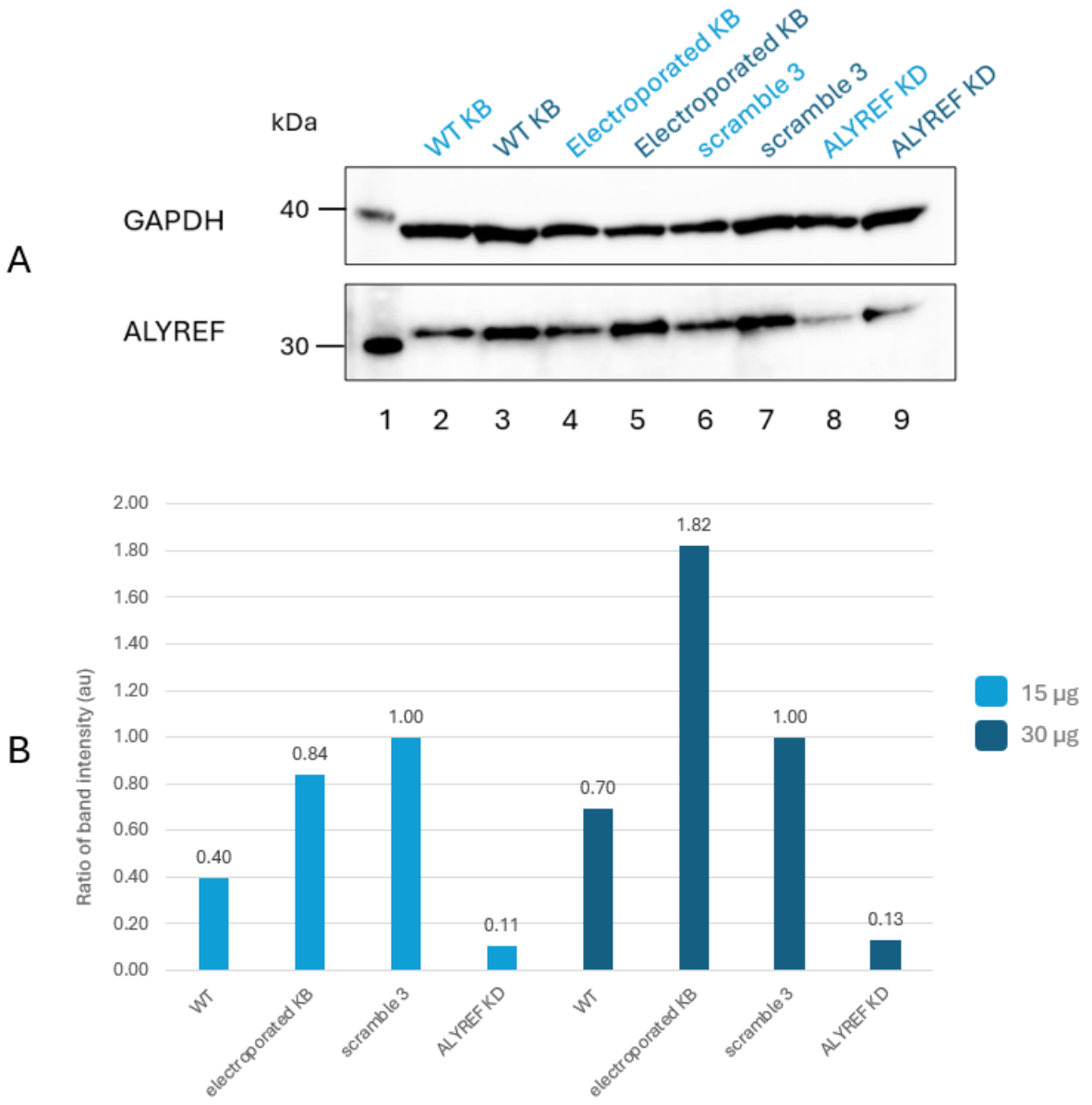

Figure 11. ALYREF KD efficiency relative to transfection with scramble 3.

Panel A: western blot showing ALYREF protein levels in KB cells. The expected molecular weight of ALYREF is ~27 kDa. Lane 1: MagicMark XP Western Protein Standard. Lanes 2, 4, 6, 8 contain 15 µg of protein (light blue); Lanes 3, 5, 7 and 9 contain 30 µg of protein (navy blue). For this assay: WT KB cells (lanes 2-3), electroporated KB control (lanes 4-5), scrambled control (lanes 6-7), ALYREF-specific siRNA (lanes 8-9). GAPDH was used as a loading control (1:50,000) to confirm effective ALYREF KD. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of ALYREF KD relative to scramble 3. Data are expressed as ALYREF/GAPDH ratios, normalised to scramble 3.

3.3 Effect of ALYREF KD on RNA

Although the FISH assay was attempted eight times during this project, limitations, including low protein yield and insufficient cell recovery, resulted in only three assays generating usable data. The three validated datasets (named FISH 1, 2 and 3) are presented below, all of which have followed the experimental workflow in Figure 9. FISH experiments were performed as stated in Section 2.3.

Section 3.2 demonstrated that ALYREF expression can be sufficiently reduced in the KB cell line. For FISH experiments 1, 2, and 3, ALYREF KD was confirmed each time using the same method. All western blot data for the three FISH experiments were analysed using ImageJ software, with ALYREF band densities normalised against GAPDH and then expressed relative to scramble 3 (Figures 11 and 13).

Foci morphology was assessed by measuring intensity (arbitrary units, au), number per nucleus, and area (μm²). Data were analysed in GraphPad Prism and presented as bar graphs (see Figures 12, 14, 16 and 17).

Overall, when pooled together, the three experiments show a significant reduction in foci area ($p = 0.0052$) and intensity ($p = 0.0013$) in the ALYREF KD cells, compared to scramble 3. The mean ($\bar{x}$) foci area decreased from 3.60 μm² in scramble 3 cells to 2.68 μm² in ALYREF KD cells, which is a 26% decrease in size. Foci intensity decreased from $\bar{x}$ = 2.15 million au in scramble 3 cells to $\bar{x}$ = 1.65 million au in ALYREF KD cells, a 23.26% decrease.

There is no significant change statistically in the number of foci; however, the mean number of foci decreased from 3.87 to 3.10 per nucleus.

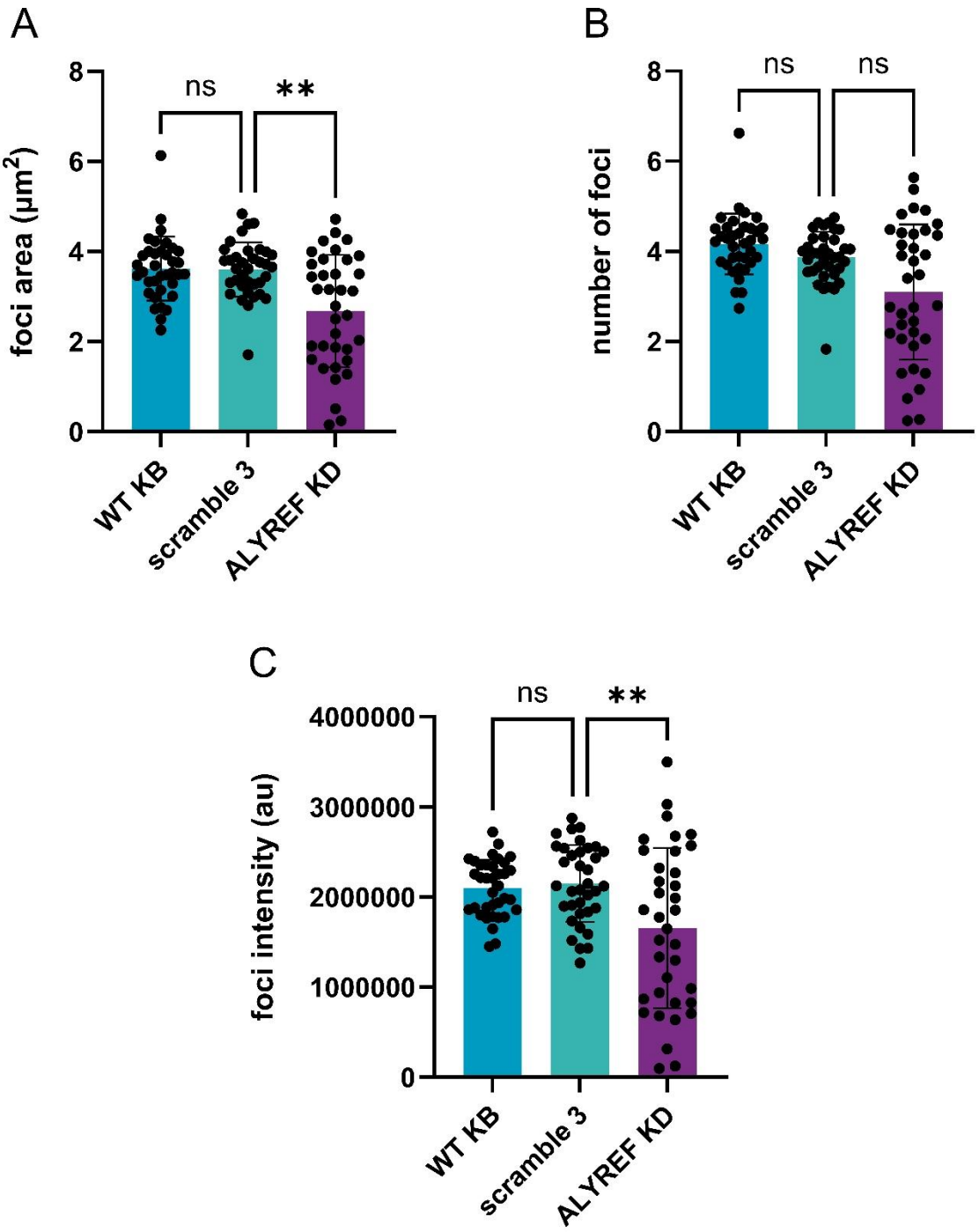

Figure 12. Combined analysis of three independent ALYREF KD experiments (FISH 1-3).

Bar graphs show the mean ± SEM for (A) foci area (µm²), (B) number of foci per nucleus, and (C) foci intensity (arbitrary units, au) in WT KB, scrambled control (scramble 3), and ALYREF KD cells. Each data point represents the mean value of one well (6 wells per condition, 5000-10,000 cells seeded per well). Data normality was assessed with the Shapiro–Wilk test, followed by one-way ANOVA or Kruskal–Wallis as appropriate. (** = $p < 0.01$, ns = not significant).

3.3.1 FISH 1: ALYREF KD and its impact on foci

In the first FISH assay, ALYREF protein expression was reduced by ~66% relative to scramble 3 (Figure 13). As this level of ALYREF KD exceeded the desired threshold set in Figure 9 (based on earlier work in the lab), transfected cells were analysed for ribonuclear foci.

ALYREF KD significantly decreased all foci characteristics in FISH 1. Foci area decreased by 46% from 4.05 μm² in scramble 3 cells to 2.19 μm² in ALYREF KD cells ($p = 0.0004$). The number of foci also decreased by ~25% from 3.97 in scramble 3 cells to 2.98 in ALYREF KD cells ($p = 0.015$). The greatest statistically significant difference was observed in the foci intensity between scramble 3 and ALYREF KD cells. The foci intensity decreased in the ALYREF KD cells by 40% from 1.97 million au to 1.18 million au ($p<0.0001$).

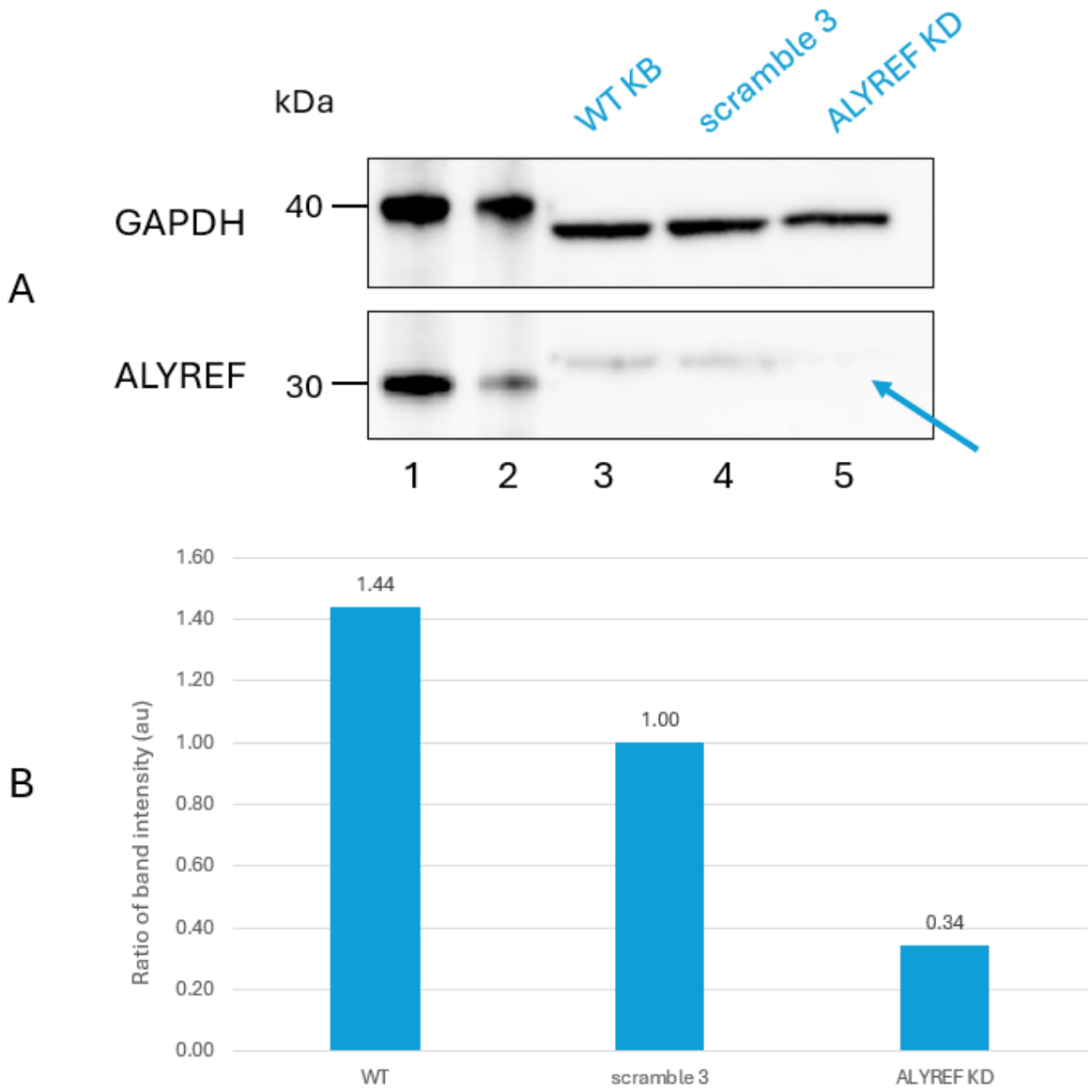

Figure 13. Confirmation of ALYREF KD in FISH 1.

Panel A: western blot showing ALYREF protein levels in KB cells from FISH 1. The expected molecular weight of ALYREF is ~27 kDa. Lane 1-2 MagicMark XP Western Protein Standard. For this assay: WT KB cells (lane 3), scrambled control (lanes 4), ALYREF-specific siRNA (lane 5). All lanes were loaded with 15 µg protein. GAPDH was used as a loading control (1:50,000) to confirm effective ALYREF KD. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of ALYREF KD relative to scramble 3. Data are expressed as ALYREF/GAPDH ratios, normalised to scramble 3.

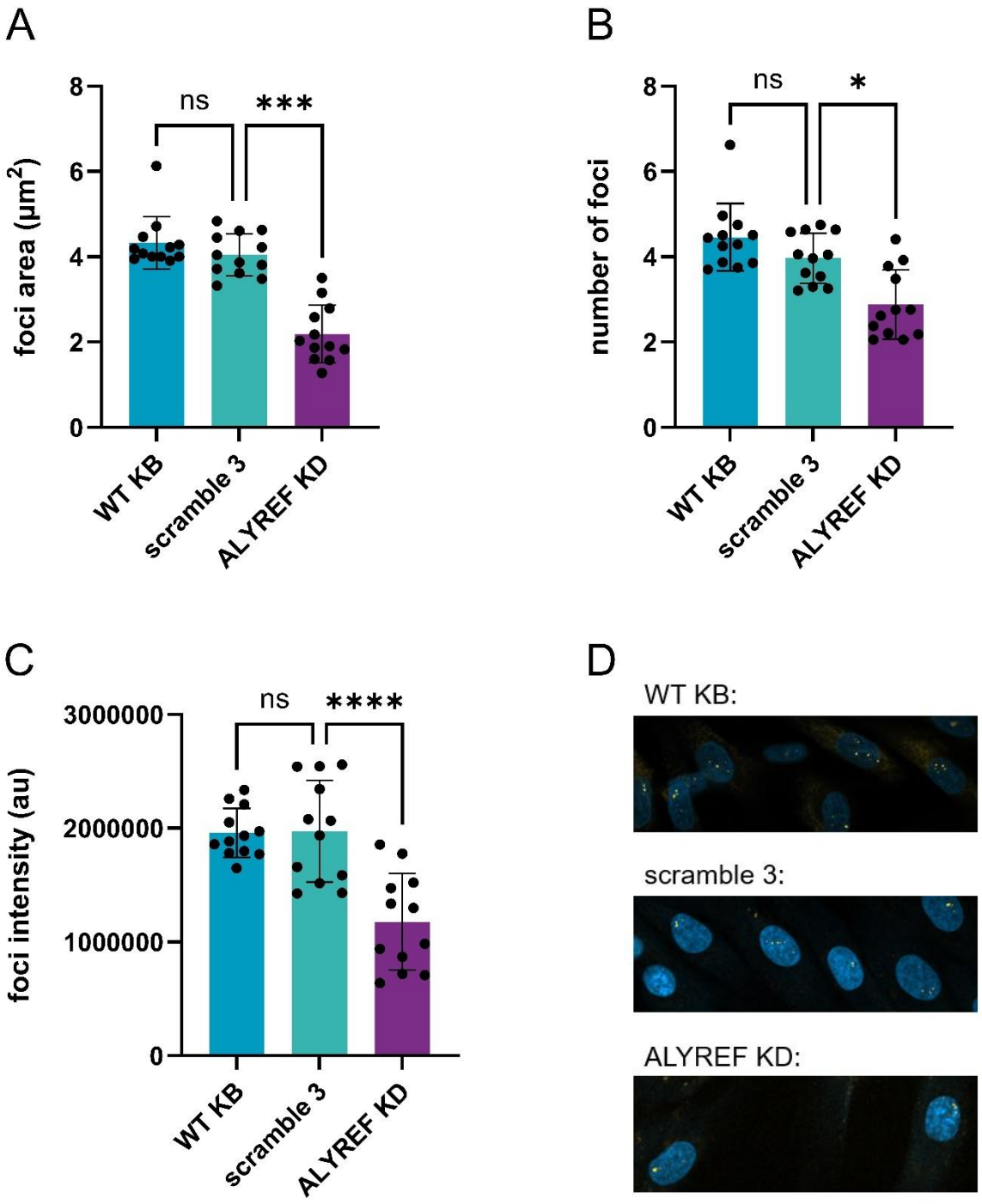

Figure 14. ALYREF KD decreased all foci characteristics FISH 1.

Bar graphs show the mean ± SEM for (A) foci area (µm²), (B) number of foci per nucleus, and (C) foci intensity (arbitrary units, au) in WT KB, scrambled control (scramble 3), and ALYREF KD cells. Each data point represents the mean value of one well (12 wells per condition, 5000-10,000 cells seeded per well). Data normality was assessed with the Shapiro–Wilk test, followed by one-way ANOVA or Kruskal–Wallis as appropriate. (**** = extremely significant, p<0.0001, *** = highly significant, p<0.001, * = statistically significant, p < 0.05, ns = not significant). Panel D shows foci images taken by the Zeiss Celldiscoverer 7 Microscope during plate imaging.

3.3.2 ALYREF KD in FISH experiments 2 and 3

ALYREF KD was validated for FISH assays 2 and 3 using western blotting (Figure 15). Quantification of band intensities (Figure 15, panel B) confirmed that ALYREF expression was reduced by 50% in FISH 2 and by 90% in FISH 3, relative to scramble 3. GAPDH was used as a loading control, and all values were normalised first to GAPDH and then to scramble 3.

The WT and siGLO controls showed comparable ALYREF expression to scramble 3 in both assays, especially in FISH 3, indicating that the observed reduction was specific to the ALYREF-specific siRNA. These results confirmed that both KDs met the minimum threshold (≥50%) defined in Figure 9, validating the datasets for subsequent foci analysis.

Due to experimental error, the protein was loaded at 30 µg, not 15 µg, but because of GAPDH normalisation, this does not affect the data.

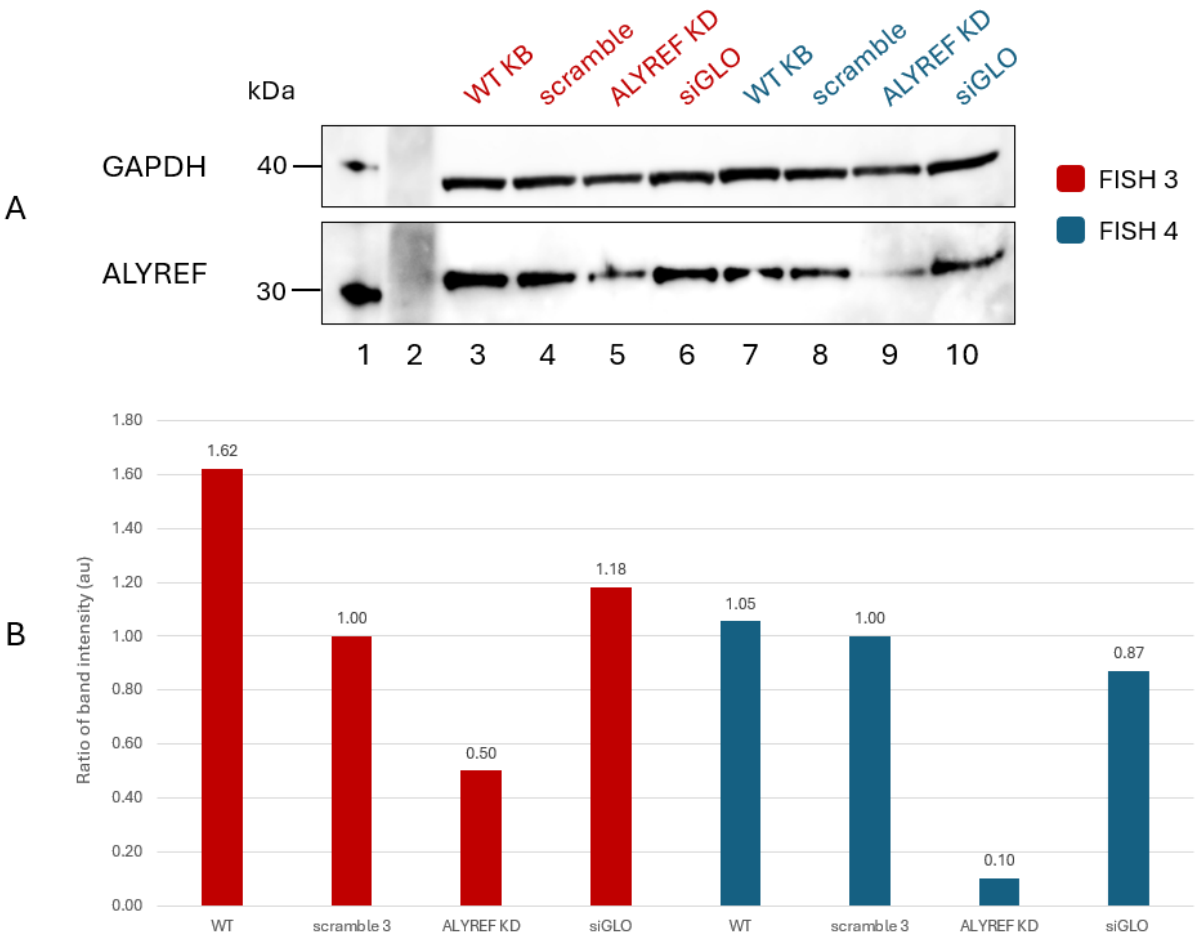

Figure 15. Confirmation of ALYREF KD in FISH 2 and 3.

Panel A: western blot showing ALYREF protein levels in KB cells from FISH 2 (red) and FISH 3 (navy blue). The expected molecular weight of ALYREF is ~27 kDa. Lane 1: MagicMark XP Western Protein Standard; Lane 2: SeeBlue Pre-stained Protein Standard. Lanes 3-6: FISH 2 samples; Lanes 7-10: FISH 3 samples. For each assay: WT KB cells (lanes 3, 7), scrambled control (lanes 4, 8), ALYREF-specific siRNA (lanes 5, 9), and siGLO control (lanes 6, 10). All lanes were loaded with 30 µg protein. GAPDH was used as a loading control (1:50,000) to confirm effective ALYREF KD. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of ALYREF KD relative to scramble 3. Data are expressed as ALYREF/GAPDH ratios, normalised to scramble 3.

3.3.3 FISH 2: Effect of ALYREF KD on the foci

In the second FISH experiment, ALYREF expression was reduced by 50% relative to scramble 3 (Figure 15). However, the data juxtaposes the data observed in FISH 1.

ALYREF KD significantly increased both the number of foci per nucleus ($p = 0.0029$) and the area of the foci ($p = 0.0458$), compared to scramble 3 (Figure 16). The number of foci per nucleus increased 1.25-fold from 4 to 5, whilst the foci area increased 1.15-fold from 3.25 μm^2 to 3.75 μm^2. In contrast, there was no statistical significance with foci intensity ($p = 0.0675$).

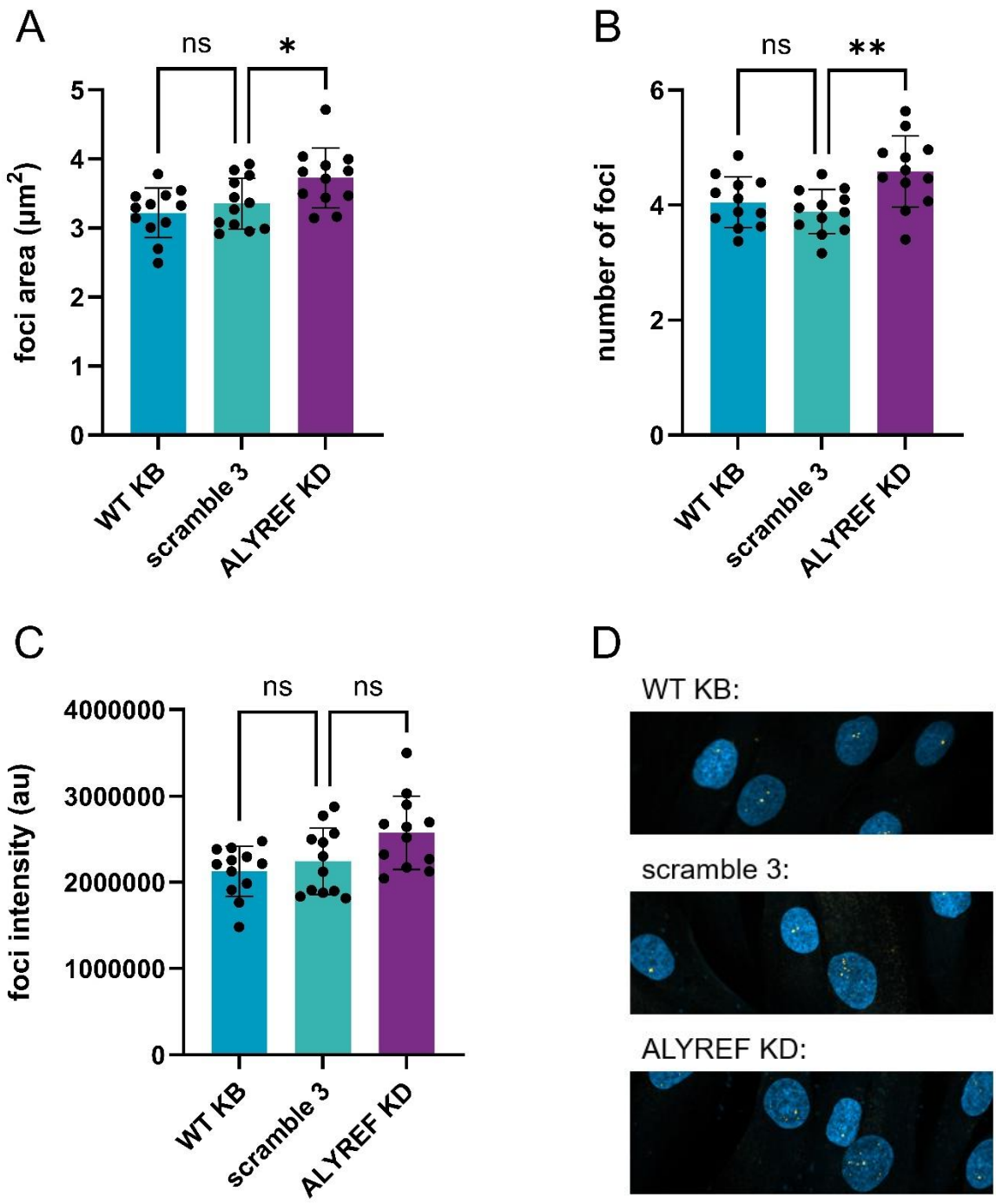

Figure 16. ALYREF KD increased foci number and area in FISH 2.

Bar graphs show the mean ± SEM for (A) foci area (μm²), (B) number of foci per nucleus, and (C) foci intensity (arbitrary units, au) in WT KB, scrambled control (scramble 3), and ALYREF KD cells. Each data point represents the mean value of one well (12 wells per condition, 10,000 cells seeded per well). Data normality was assessed in the same way as Figure 14. (** = very significant, $p < 0.01$, * = statistically significant, $p < 0.05$, ns = not significant). Panel D shows foci images taken by the Zeiss Celldiscoverer 7 Microscope during plate imaging.

3.3.4 FISH 3: Effect of ALYREF KD on the foci

The third FISH experiment, FISH 3, shows a reduction in ALYREF expression of 90% relative to scramble 3 (Figure 15). This is the highest protein KD observed in the three FISH assays presented in this thesis.

FISH 3 exhibited a trend in foci characteristics like that observed in FISH experiment 1 and strongly juxtaposed those of FISH 2. However, the data for FISH experiment 3 displayed considerable variability, resulting in a broad distribution and correspondingly wide error bars.

ALYREF KD significantly decreased both the number of foci per nucleus ($p = 0.0134$) and the foci intensity ($p = 0.0005$), compared to scramble 3 (Figure 17). The mean number of foci decreased by 51.5% from 3.75 to 1.82 foci per nucleus. The foci intensity decreased 45.6% from 2.2 million au to 1.2 million au. There was no statistical significance or notable change in the foci area because the range of data was too large ($p = 0.08$).

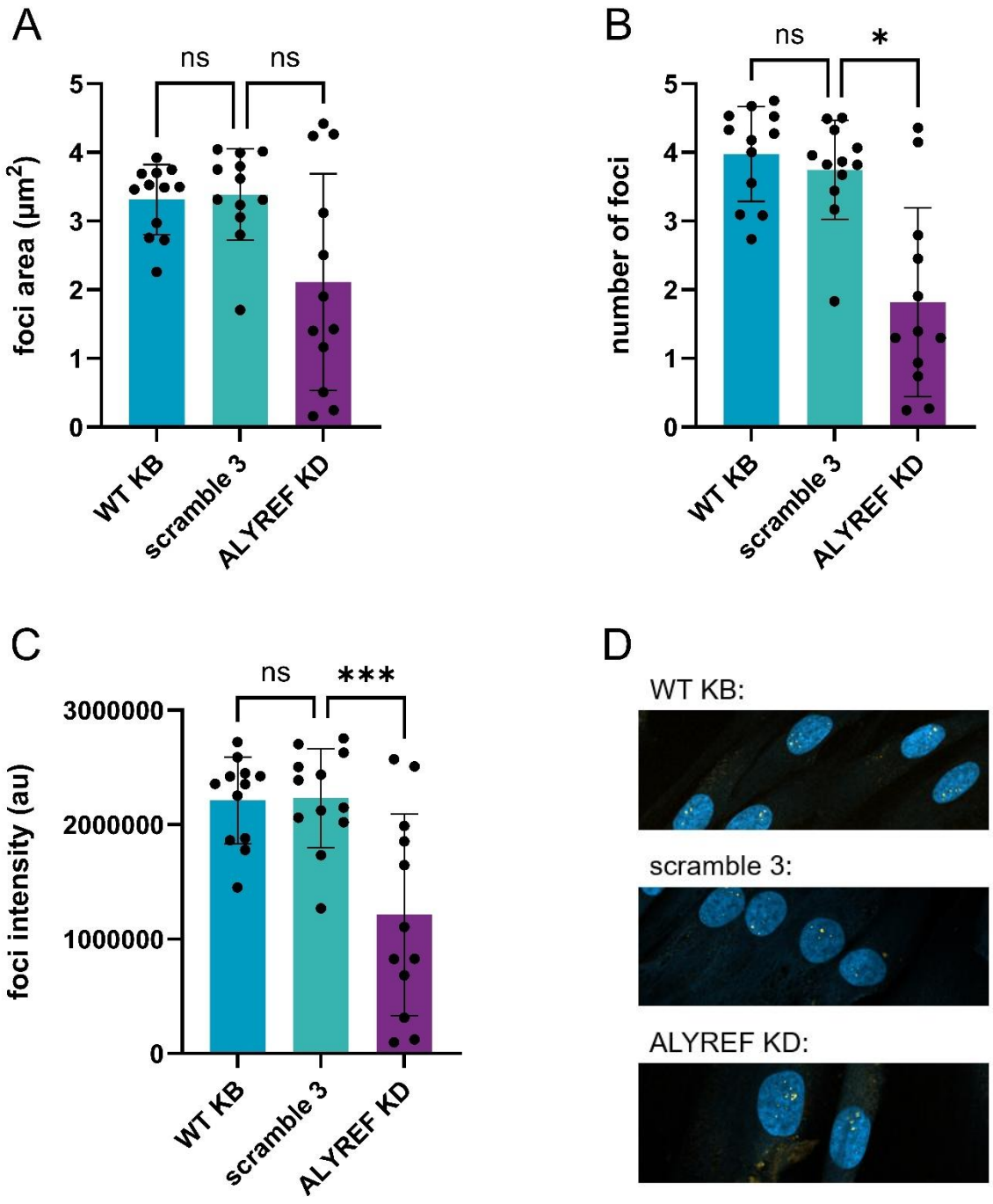

Figure 17. ALYREF KD decreased foci number and intensity in FISH 3.

Bar graphs show the mean ± SEM for (A) foci area (μm²), (B) number of foci per nucleus, and (C) foci intensity (arbitrary units, au) in WT KB, scrambled control (scramble 3), and ALYREF KD cells. Each data point represents the mean value of one well (12 wells per condition, 10,000 cells seeded per well). Data normality was assessed in the same way as Figure 14. (*** = highly significant, $p < 0.001$, * = statistically significant, $p < 0.05$, ns = not significant). Panel D shows foci images taken by the Zeiss Celldiscoverer 7 Microscope during plate imaging.

3.4 Additional ALYREF data

In addition to the data presented above, ALYREF KD analysis was also conducted in the lab by Grace Bamford (*G. Bamford*), an undergraduate project student. This unpublished data is presented below to provide a comparison to the data I generated.

3.4.1 Confirmation of ALYREF KD in KB cells

In *G. Bamford*'s transfection, the same ALYREF-specific siRNA used in the experiments above was chosen to transfect the KB cells. ALYREF expression was reduced by 91% compared to WT KB cells (Figure 18). As with previous data, the western blot was analysed with ImageJ and band intensities quantified in relation to the control (in this case, WT KB cells).

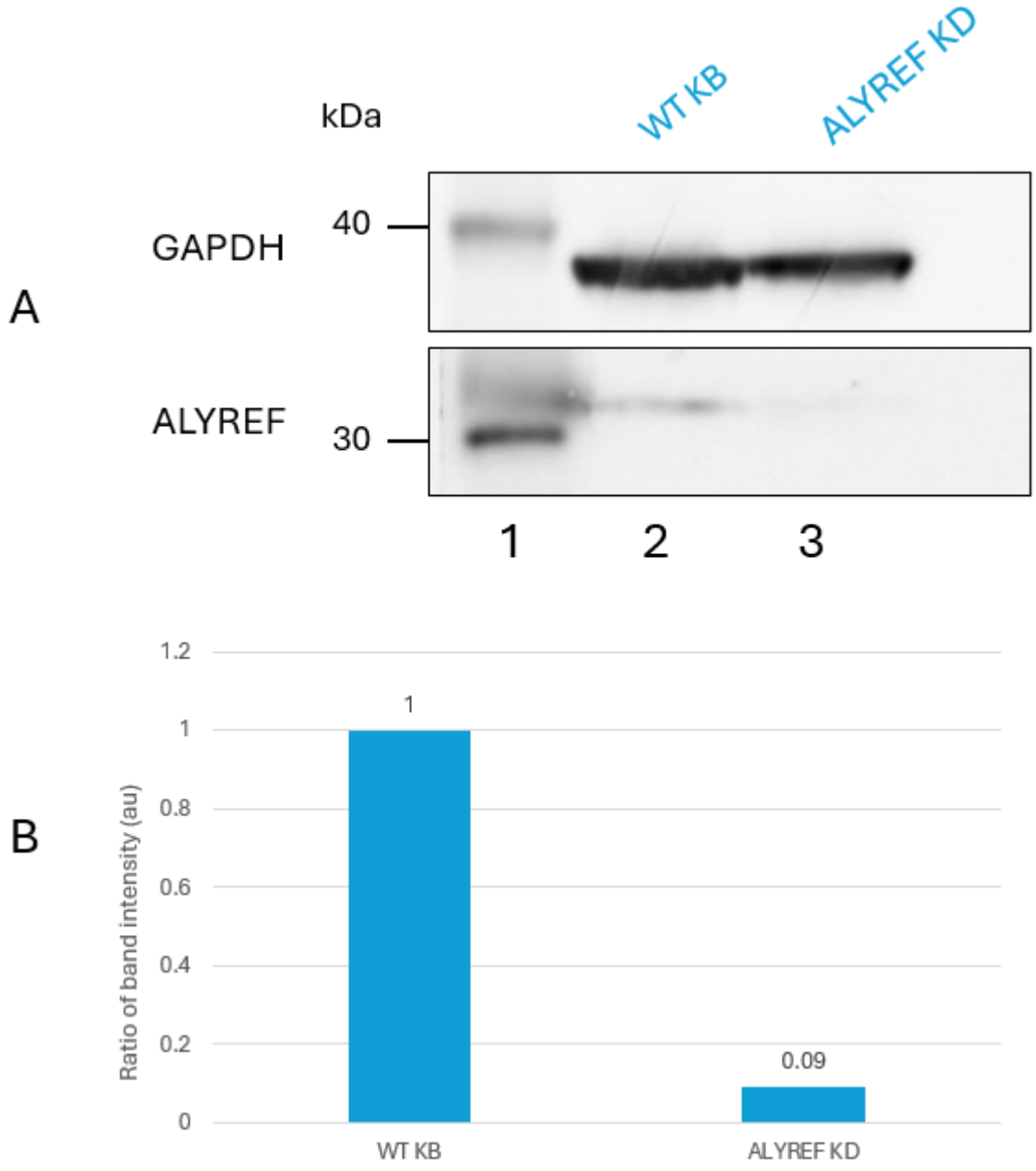

Figure 18. Confirmation of ALYREF KD.

Panel A: western blot showing ALYREF protein levels in KB cells. Lane 1: MagicMark XP Western Protein Standard; Lane 2: WT KB cells; Lane 3: KB cells transfected with ALYREF-specific siRNA. All lanes were loaded with 15 µg protein. GAPDH was used as a loading control (1:5000) to confirm effective KD of ALYREF. Panel B shows the quantification of ALYREF KD relative to WT KB cells. Data are expressed as ALYREF/GAPDH ratios, normalised to WT. Data produced by *G. Bamford*.

3.4.2 FISH 4 (*G. Bamford)*: ALYREF KD and its effect on the foci

The fourth FISH experiment shows a reduction in ALYREF expression of 88% relative to WT KB cells (Figure 19). WT KB cells were used as the control cell line.

FISH 4 showed a general increase in foci intensity, area and the number of foci per nucleus, similar to FISH 2 (FISH 2 data -Figure 16; FISH 4 data - Figure 20). This juxtaposes the findings in FISH 1 and FISH 3.

ALYREF KD significantly increased the number of foci per nucleus and the foci area (Figure 20, Panels A and B). The number of foci had a 2-fold increase from WT KB to ALYREF KD cells from ~5 to ~10 foci per nucleus ($p = 0.0001$). The average number of foci normally observed in WT cells is between 4 and 5; however, 10 is the greatest average observed in the data presented in this thesis. The foci area had a very significant increase with a 1.4-fold change from 5 to 7 μm^2 ($p = 0.0026$).

Interestingly, like FISH 2, there was no statistical significance or notable change in the foci intensity ($p = 0.4607$).

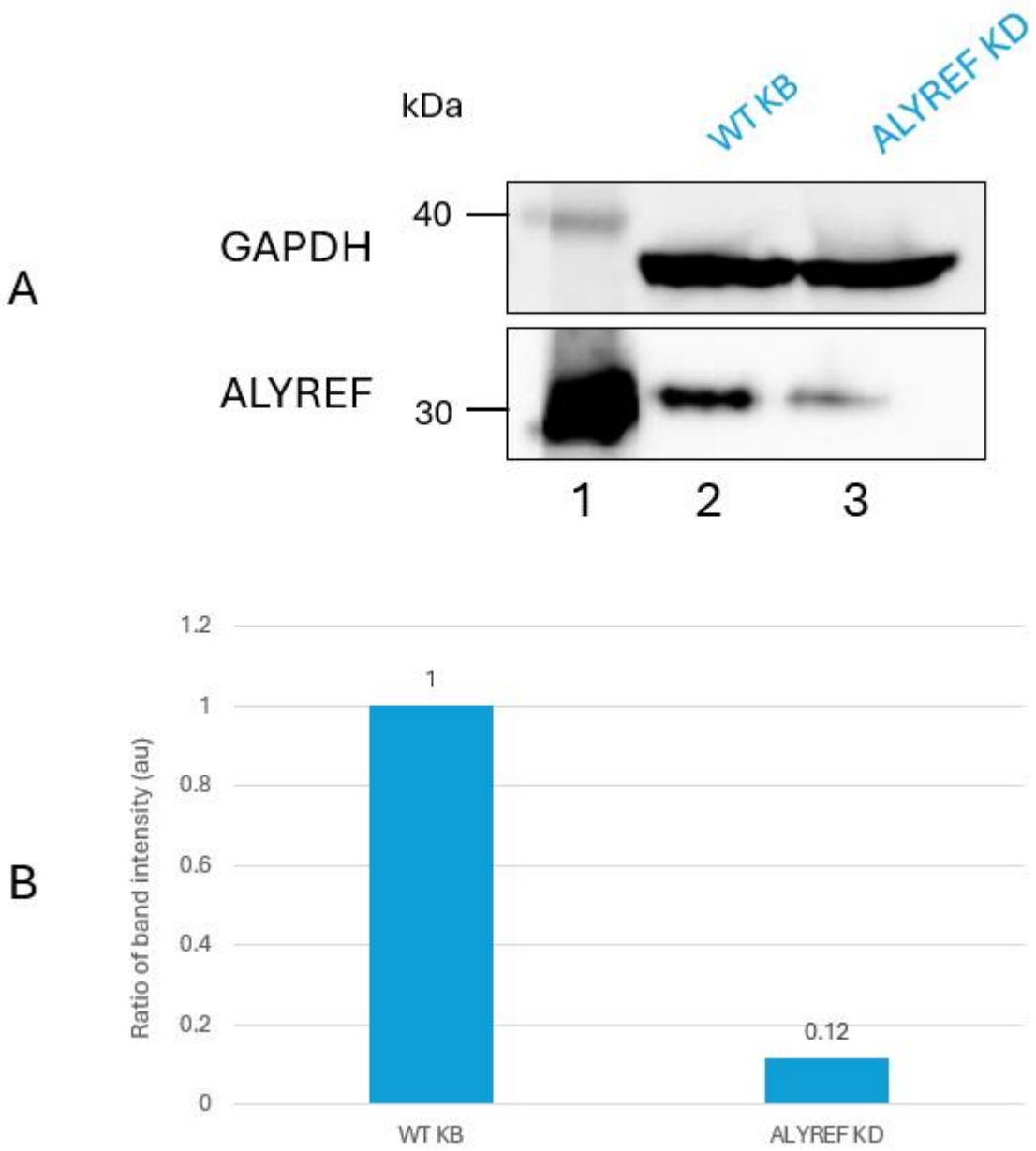

Figure 19. Confirmation of ALYREF KD in FISH 4.

Panel A: western blot showing ALYREF protein levels in KB cells from FISH 4. The expected molecular weight of ALYREF is ~27 kDa. Lane 1: MagicMark XP Western Protein Standard; Lane 2: WT KB cells; Lane 3: KB cells transfected with ALYREF-specific siRNA. All lanes were loaded with 15 µg protein. GAPDH was used as a loading control (1:5000) to confirm effective ALYREF KD. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of ALYREF KD relative to WT KB cells. Data are expressed as ALYREF/GAPDH ratios, normalised to WT. Data produced by *G. Bamford*.

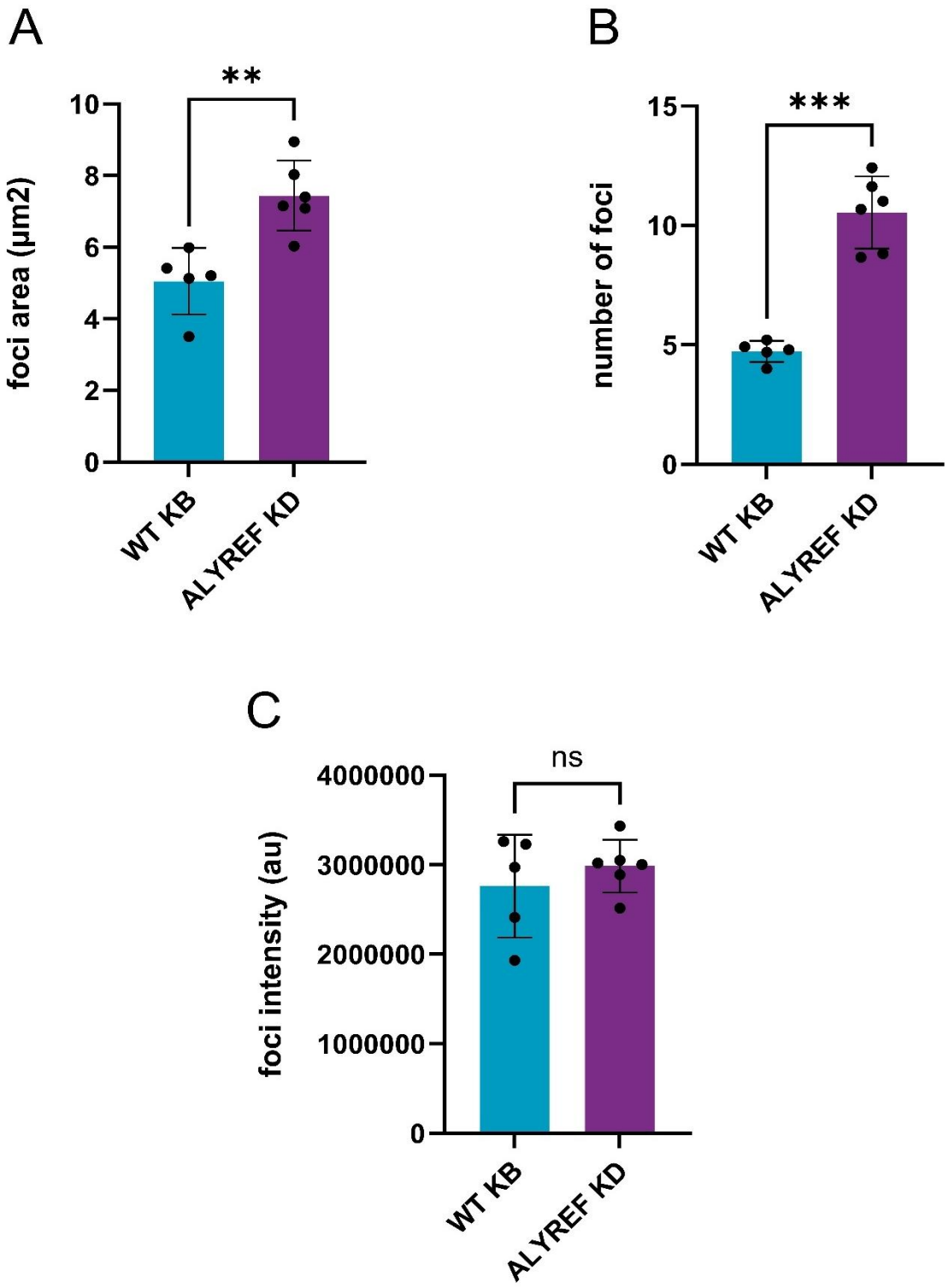

Figure 20. ALYREF KD increase foci number and area in FISH 4.

Bar graphs show the mean ± SEM for (A) foci area (µm²), (B) number of foci per nucleus, and (C) foci intensity (au) in WT KB and ALYREF KD cells. Each data point represents the average of individual wells (5 wells per WT, 6 wells per ALYREF KD, ~5,000 cells seeded per well). Data normality was assessed with the Shapiro–Wilk test, followed by Welch's t-test. (*** = highly significant, $p < 0.001$, ** = very significant, $p < 0.01$, ns = not significant). This data was produced by *G. Bamford*.

3.5 NSUN2 KO and its effect on foci

In addition to the data presented above, NSUN2 KO analysis was conducted in the lab by Owen Hannant (*O. Hannant*), an undergraduate project student. This unpublished data is presented below to provide a comparison to the data I generated and to add to the 'bigger picture' of epitranscriptomics in DM1.

As mentioned in Section 1.3, epitranscriptomic modification of RNA and actions arising from it are controlled by the coordinated activity of proteins known as the 'writer', 'reader' and 'eraser'. In this respect, ALYREF is a protein that reads the signals written by a writer protein such as NSUN2. Thus, it seemed reasonable to examine the effect of NSUN2 KO to determine how this related to or compared to ALYREF KD.

For NSUN2 KO, CRISPR Cas 9 genome engineering was employed by *O. Hannant.* He designed and tested PCR primers and single guide RNAs (sgRNAs) from Sigma-Aldrich and Synthego, respectively, before running a PCR (Taq protocol) and subsequent gel to be sequenced by Sanger sequencing. As well as his NSUN2 sgRNA, scramble 1 and scramble 2 were used. For transfection, the protocol was similar to that in Section 2.1.5; however, the nucleofector solution was added to the sgRNA/Cas9 complex (10 µl of 30 µM sgRNA, 3 µl Cas9 protein) before resuspending the cell pellet and electroporating the samples.

3.5.1 NSUN2 KO with CRISPR Cas9 (O. Hannant)

In the first NSUN2 KO experiment, NSUN2 protein expression was reduced by 99.9% relative to scramble 1 (Figure 21). In the second NSUN2 KO experiment, NSUN2 protein expression was reduced by 99.4% relative to scramble 2 (Figure 22). These levels of knock out significantly exceed the threshold set in Figure 9, so the data have been presented.

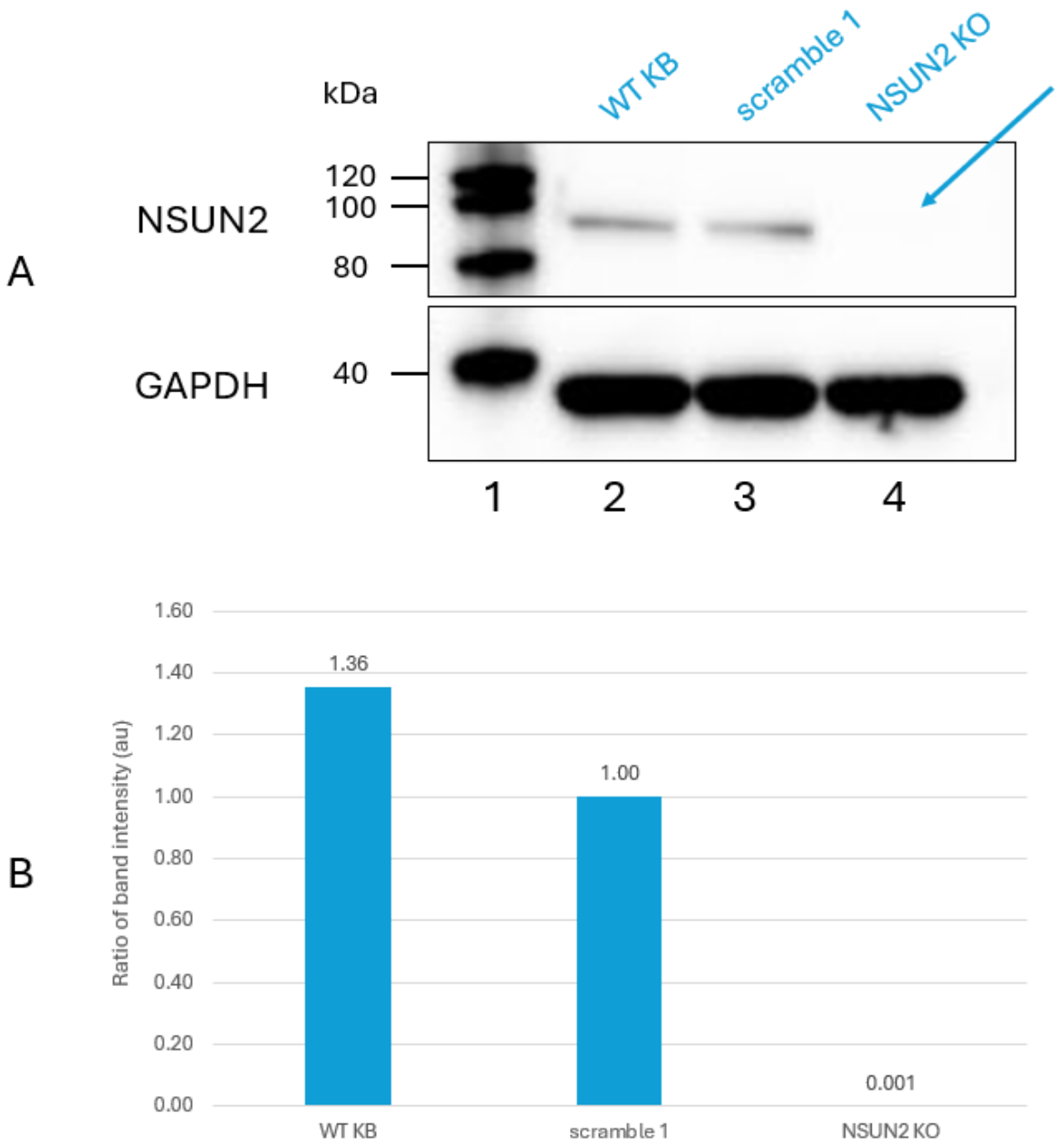

Figure 21. Confirmation of NSUN2 KO to scramble.

Panel A: western blot showing NSUN2 protein levels in KB cells from FISH 5. The expected molecular weight of NSUN2 is ~85-100 kDa. Lane 1: MagicMark XP Western Protein Standard; Lane 2: WT KB cells; Lane 3: KB cells transfected with scramble 1; Lane 4: KB cells transfected with NSUN2-specific sgRNAs. All lanes were loaded with 15 µg protein. GAPDH was used as a loading control (1:5000) to confirm effective NSUN2 KO. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of NSUN2 KO relative to WT KB cells. Data are expressed as NSUN2/GAPDH ratios, normalised to scramble 1. Data produced by *O. Hannant*.

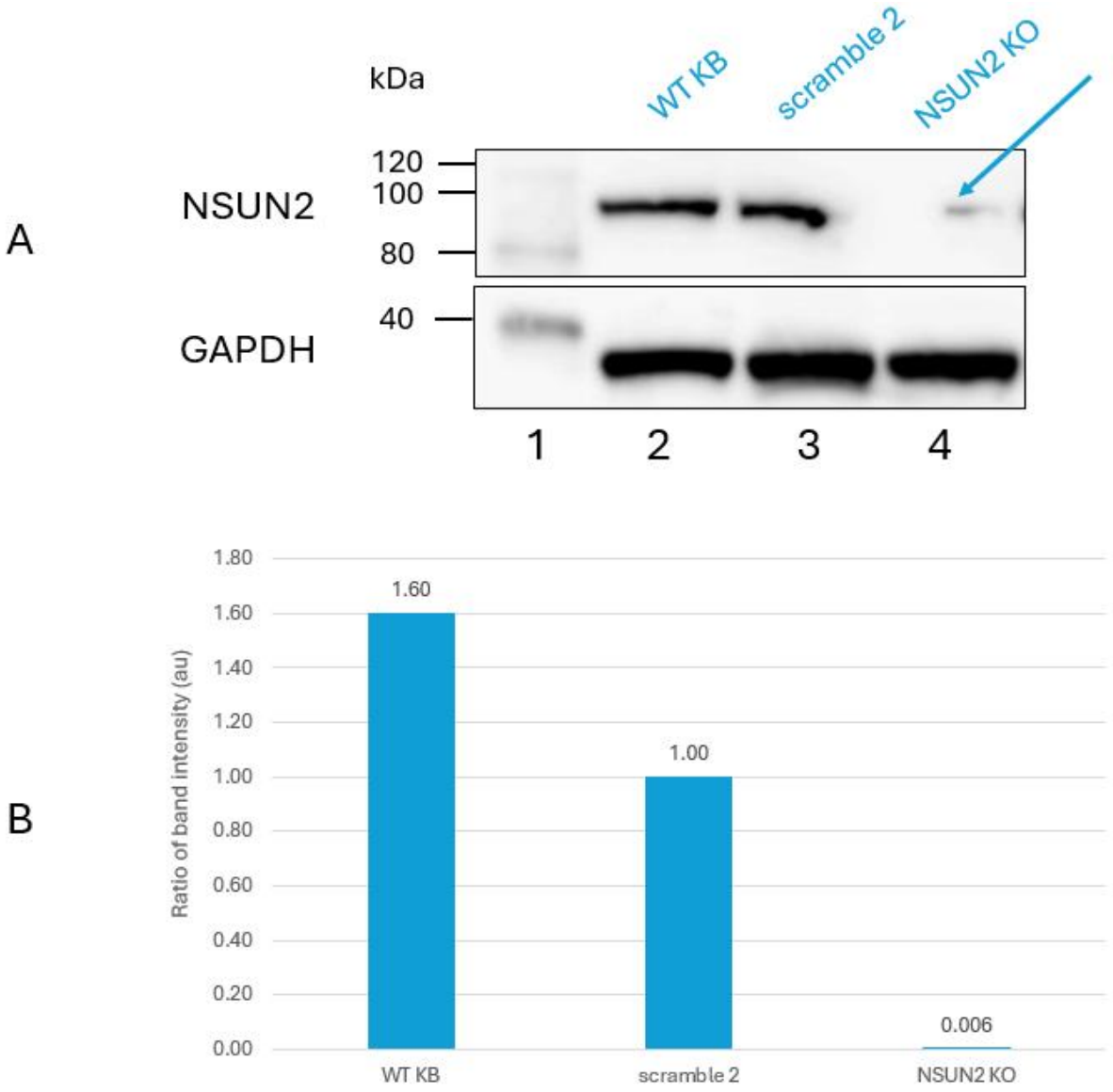

Figure 22. Confirmation of NSUN2 KO to scramble 2.

Panel A: western blot showing NSUN2 protein levels in KB cells from FISH 6. The expected molecular weight of NSUN2 is ~85-100 kDa. Lane 1: MagicMark XP Western Protein Standard; Lane 2: WT KB cells; Lane 3: KB cells transfected with scramble 2; Lane 4: KB cells transfected with NSUN2-specific sgRNAs. All lanes were loaded with 15 µg protein. GAPDH was used as a loading control (1:5000) to confirm effective NSUN2 KO. The expected molecular weight of GAPDH is ~37 kDa. Panel B shows the quantification of NSUN2 KO relative to WT KB cells. Data are expressed as NSUN2/GAPDH ratios, normalised to scramble 2. Data produced by *O. Hannant*.

3.5.2 NSUN2 KO and its effect on the foci (*O. Hannant)*

Figure 23 presents a combined analysis of two independent FISH experiments examining the effect of NSUN2 KO on ribonuclear foci characteristics in KB cells. In Figure 23, the hollow circle symbolises experiment 1, and the solid circle symbolises experiment 2. Each experiment employed a distinct scramble control, which is shown separately to account for experimental variability. WT KB cells are included as a baseline reference.

3.5.2.1 NSUN2 KO foci area

The reduction in NSUN2 expression resulted in a consistent and significant decrease in ribonuclear foci area relative to scramble 2 (Figure 23A). When data from both experiments were combined, the mean foci area decreased from 6.45 μm² for scramble 1 and 12.30 μm² in scramble 2 to 5.01 μm² in NSUN2 KO cells.

When analysed separately, both experiments demonstrated a significant reduction in foci area following NSUN2 KO. In the first experiment, the mean foci area decreased 47.13% ($p = 0.0004$) from 6.45 μm² to 3.41 μm² in scramble 1 and NSUN2 KO cells. In the second experiment, the foci area decreased by 46.26% ($p = 0.0066$) from $\bar{x}$ = 12.30 μm² in scramble 2 cells to $\bar{x}$ = 6.61 μm² in NSUN2 KO cells.

3.5.2.2 NSUN2 KO number of foci per nucleus

A significant decrease in the number of foci per nucleus was observed following NSUN2 KO (Figure 23B). Compared to scramble 1 and scramble 2 controls ($\bar{x}$ = 4.32 and $\bar{x}$ = 5.66 foci per nucleus, respectively), NSUN2 KO cells showed a 27.5% reduction and 44.7% reduction with a mean of 3.13 foci per nucleus ($p = 0.0146$ and $p \leq 0.0001$), indicating a robust reduction in foci abundance across independent experiments.

When analysed separately, the decrease in the number of foci per nucleus in the first experiment was highly significant ($p = 0.0006$) between scramble 1 and NSUN2 KO

cells, from $\bar{x}$ = 4.32 foci per nucleus to $\bar{x}$ = 2.89. This shows a 33.10% reduction. The second experiment also showed a highly significant ($p < 0.0002$) decrease in the number of foci per nucleus between scramble 2 and NSUN2 KO cells, from $\bar{x}$ = 5.66 foci per nucleus to $\bar{x}$ = 3.14 ($p = 0.0002$).

3.5.2.3 NSUN2 KO foci intensity

Foci intensity was reduced in NSUN2 KO cells compared with scramble controls (Figure 23C). While absolute intensity values varied between experiments, NSUN2 KO consistently resulted in a significant reduction in signal intensity compared to scramble 1 (67.85% decrease; $p = 0.0019$). The variability between scramble 1 and 2 likely reflects technical differences between independent FISH experiments rather than biological inconsistency, as the direction of effect was preserved across replicates.

In the first experiment, foci intensity decreased from 863356 au in the scramble 1 control cells to 459685 au in the ALYREF KD cells. That is a 46.76% decrease and extremely significant ($p < 0.0001$). In the second experiment, foci intensity also decreased in the NSUN2 KO cells compared to scramble 2 ($p = 0.0128$), which is not represented in Figure 23. There was a 44.75% reduction from $\bar{x}$ = 172663 for scramble 2 to $\bar{x}$ = 95400 for NSUN2 KO. However, the average foci intensity observed across all conditions in previous FISH assays was greater than 1 million au, significantly higher than the values obtained in both experiments.

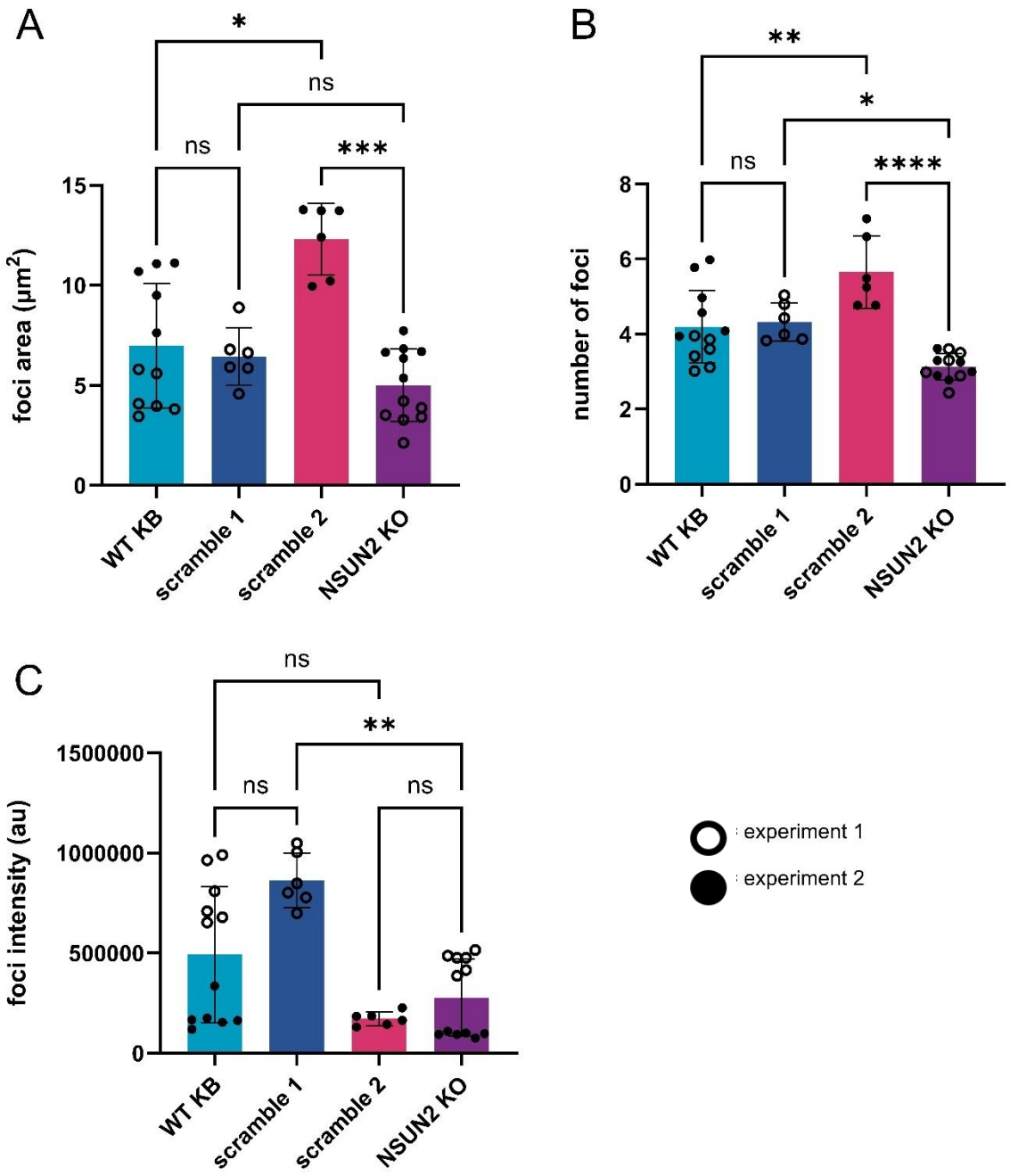

Figure 23. NSUN2 KO decreases all observed characteristics.

Bar graphs show the mean ± SEM for (A) foci area, (B) number of foci per nucleus, and (C) foci intensity in WT KB cells, KB cells transfected with either scramble 1 or scramble 2 and NSUN2 KO cells. Scramble 1 and scramble 2 represent independent experimental controls from separate FISH experiments. Each data point represents the average of individual wells (6 wells per condition, ~5,000 cells/well). Data normality was assessed with the Shapiro–Wilk test, followed by one-way ANOVA or Kruskal-Wallis, as appropriate. (**** = extremely significant, $p \leq 0.0001$; *** = highly significant, $p \leq 0.001$; ** = highly significant, $p \leq 0.01$; * = statistically significant, $p \leq 0.05$; ns = not significant). Data produced by *O. Hannant*.

4 Discussion

The results obtained in this project remain preliminary until further experimental repeats are conducted; however, the data provide valuable insight into the potential roles of ALYREF and NSUN2 in DM1. A consistent trend was observed regarding foci in NSUN2 KO cells. In contrast, the effects of ALYREF KD were less conclusive, with two pairs of FISH experiments (FISH 1 and 3, FISH 2 and 4) demonstrating opposing outcomes.

These findings, therefore, require careful interpretation alongside existing literature and underline the need to account for both technical and biological factors when evaluating the role of ALYREF and NSUN2 in DM1. Importantly, both proteins are functionally interdependent within the m^{5}C pathway, so their effects may converge even when experimental outcomes appear divergent.

4.1 ALYREF

At the outset, it was necessary to confirm ALYREF protein expression in one of the DM1 cell lines to ensure that knockdown would yield a measurable effect. Three well-established cell lines were assessed: KB, a DM1 cell line routinely used in the laboratory, XTB, a DM1 cell line used to a lesser extent in the laboratory, and SB, a non-DM1 control. Because ALYREF is a ubiquitous protein, it was expected to be expressed equally in both diseased and non-diseased cells. As shown in Figure 10, ALYREF expression was confirmed in all cell lines, and KB cells were selected for subsequent experiments. Since the same cell line was used throughout, the number of (CTG)n repeats can be assumed to remain consistent. This is important because previous studies (Botta et al., 2008; Ballester-Lopez et al., 2020; Núñez-Manchón et al., 2024) have demonstrated that, in DM1, the number of foci per nucleus correlates with (CTG) repeat length. Using a single cell line across all experiments therefore minimised variability arising from repeat-length differences.

The next stage was to determine whether ALYREF could be effectively knocked down in KB cells by transfection with ALYREF-specific siRNAs (see Table 3, Section 2.1.5). Previous data from *G. Bamford* suggested that this was achievable (see Figures 18 and 19), and the same protocol and siRNAs were employed to replicate these findings. In my experiments, two controls were included to confirm that any observed effects were specific to ALYREF KD: (i) KB cells electroporated without additional reagents, to assess whether electroporation itself affected the cells, and (ii) KB cells transfected with scramble 3 siRNA, to evaluate whether siRNA transfection had a non-specific impact on ALYREF protein expression.

Notably, as shown in Figure 13, both the electroporation and scramble 3 exhibited elevated ALYREF protein expression compared with WT KB cells. This may reflect a cellular stress response to electroporation and/or siRNA exposure, or experimental variability in protein detection. However, only the ALYREF-specific siRNAs produced a marked reduction in protein expression relative to scramble 3, confirming a successful and specific knockdown. This highlights the importance of including multiple controls when interpreting KD data.

Next, to investigate the functional effects of ALYREF KD, ribonuclear foci were assessed by FISH. All four experiments examined changes in foci intensity, area, and number per nucleus, since ribonuclear foci are a defining feature of DM1 (Figure 4) and can be readily visualised microscopically using this approach (Swanson et al., 1993; Taneja et al., 1995; Michalowski et al., 1999). This data provides important insight into the extent of RNA aggregation, which was hypothesised to increase upon knockdown of the mRNA export factor ALYREF.

FISH	% KD	Change in foci characteristics		
		Intensity	Area	Number
1	66	↓↓↓↓	↓↓↓	↓
2	50	ns	↑	↑↑
3	90	↓↓↓	ns	↓
combined	-	↓↓	↓↓	ns
4	88	ns	↑↑	↑↑↑

Table 4: Summary of ALYREF FISH experiments.

Trends observed in foci intensity, area, and number per nucleus across the four ALYREF FISH assays. A downward arrow (↓) indicates a decrease, and an upward arrow (↑) indicates an increase. Multiple arrows denote greater statistical significance, while “ns” indicates no significant change. % KD represents the percentage protein reduction.

Overall, FISH 1-3 show a significant reduction in foci intensity and area compared to the scrambled control (Figure 12, panels A and C). There was no significance in overall foci number (Figure 12, panel B) which may indicate that ALYREF depletion may alter foci properties without reducing their abundance. It must also be noted that there is a large range of data for ALYREF KD across all three foci characteristics, and thus a high standard deviation. This is most likely due to the variation and thus the effect of knockdown achieved across experiment. As the effect of ALYREF KD on the foci in DM1 has not previously been looked into, it is difficult to spot outliers.

In FISH 1 (~66% KD), foci intensity and foci area were reduced by ~40% and 46%, respectively (Figure 14), suggesting ALYREF contributes to the stability and/or maturation of foci rather than the initial formation. This is important because whilst foci number reflects how many initial aggregation sites form, foci intensity and area reflect how much RNA and protein accumulate at each site. Therefore, this data implies that ALYREF KD may destabilise existing foci, preventing them from fully maturing or maintaining their size/intensity, making them smaller and less intense than in WT KB cells. Furthermore, although the average number of foci per nucleus

decreased by ~25%, indicating that ALYREF depletion may alter foci properties whilst reducing their abundance.

A similar pattern was observed in FISH 3 (~90% KD), where both intensity and number of foci decreased (Figure 17). Together, these experiments suggest that substantial ALYREF depletion destabilises ribonuclear foci, reducing their accumulation but not their abundance (to a certain degree). This suggests a threshold effect, where a greater loss of ALYREF has more pronounced consequences for foci stability. However, FISH 3 also displayed wide error bars, indicating substantial variability between replicates and underlining the need for additional repeats to confirm these trends.

On the other hand, FISH 2 (~50% KD) and FISH 4 (~88% KD, performed independently by *G. Bamford*) produced the opposite outcome, showing significant increases in both foci number and area, while foci intensity remained unchanged (FISH 2 - Figure 16, FISH 4 - Figure 20). Notably, these two experiments yielded the highest absolute foci counts observed in the dataset, with up to ten foci per nucleus in FISH 4. Importantly, the observed increase in foci accumulation aligns with the previous literature, which shows that normally transcripts are exported from the nucleus by ALYREF, and so ALYREF KD disrupts TREX-mediated mRNA export, leading to nuclear accumulation of the mutant mRNA and thus foci (see Section 1.3.3; Shi et al., 2017; Yang et al., 2017; Fan et al., 2019).

In the case of FISH 4, however, the absence of a scrambled siRNA or transfection efficiency controls (see Section 3.4) prevents direct comparison with other assays, and it cannot be excluded that transfection itself contributed to the observed effects. Nevertheless, the similarity between FISH 2 and FISH 4 suggests that under certain conditions, ALYREF depletion may enhance, rather than suppress, RNA aggregation.

Overall, the variability between experiments suggests that the effect of ALYREF KD is not linear but may depend on the degree of depletion. This echoes observations in other RNA-binding proteins, such as TDP-43, where partial versus complete loss produced distinct phenotypes (Yang et al., 2014). It also resonates with stress-response literature, where Huang et al. (2024) showed that NSUN2–ALYREF cooperation influenced Nrf2 mRNA regulation in a dose- and context-dependent manner. Moreover, Hojka-Osinska. (2021) showed that if an RNA-binding protein is only partly depleted, other factors may compensate but if it is almost entirely knocked out, compensation may fail, producing a different phenotype which may explain the differences between moderate (50-80%) and severe (>80%) ALYREF KD.

Alternatively, variability could reflect technical differences, such as transfection efficiency or protein loading, or biological heterogeneity within the cell population. Thus, while all experiments confirm that ALYREF influences foci dynamics, the precise direction of this effect appears sensitive to the degree of KD, cellular stress responses, or technical variability.

4.2 NSUN2

Preliminary CRISPR-Cas9 KO experiments (n=2) performed in this lab suggest a direct role for *NSUN2* in modulating the DM1 disease pathway. In NSUN2 KO cells, mutant *DMPK* RNA transcript levels were reduced, indicating that m^5C may contribute to nuclear retention of the RNA and transcriptional dysregulation.

CRISPR-mediated KO was chosen by *O.Hannant* because it produces stable, near-complete depletion of NSUN2, allowing clearer observation of foci dynamics. In contrast, siRNA KD is transient and often incomplete, which may not sufficiently reduce m^5C to reveal strong phenotypic effects on foci formation and maturation. *G.Bamford* did attempt CRISPR-mediated KO with ALYREF but could not get it to work.

NSUN2 data	% KO	Change in foci characteristics		
		Intensity	Area	Number
Experiment 1	99.9	↓↓↓↓	↓↓↓	↓↓↓
Experiment 2	99.4	↓↓	↓↓	↓↓↓↓
combined	-	↓↓ (scramble 1) ns (scramble 2)	ns (scramble 1) ↓↓↓ (scramble 2)	↓ (scramble 1) ↓↓↓↓ (scramble 2)

Table 5: Summary of NSUN2 FISH experiments.

Trends observed in foci intensity, area, and number per nucleus across the two NSUN2 FISH assays performed by *O.Hannant*. A downward arrow (↓) indicates a decrease, and an upward arrow (↑) indicates an increase. Multiple arrows denote greater statistical significance, while “ns” indicates no significant change. % KO represents the percentage protein reduction.

Both NSUN2 experiments represent severe NSUN2 KOs (>99%) and showed statistically significant decreases in all three foci characteristics, strongly contrasting the variable ALYREF results (Figure 23). Although baseline foci area differed between experiments, the direction and magnitude of the NSUN2-dependent reduction were consistent, indicating a reproducible effect of NSUN2 loss on foci size.

NSUN2 is a well-characterised methyltransferase ‘writer’ in the m^5C pathway, which plays an important role in mRNA export (Yang et al., 2017). The methylcytosine (m^5C) is recognised by the mRNA export adaptor ALYREF, facilitating the translocation of mRNA from the nucleus to the cytoplasm. Thus, the hypothesis suggests that the reduction of NSUN2 (KO) might produce a similar effect to the reduction (KD) of ALYREF. In this case, the data suggest that NSUN2 KO produces a significant reduction in nuclear foci.

4.3 NSUN2 and ALYREF

Previous findings (Yang, 2017) demonstrated that NSUN2 silencing increases nuclear retention of ALYREF while reducing RNA-binding affinity. In the context of DM1, this suggests that although mutant RNA remains in the nucleus, it cannot efficiently assemble into mature foci in the absence of NSUN2-mediated m^5C modification. An alternative explanation is that in the absence of methylation, the mutant RNA is targeted for degradation.

These results also highlight the tight functional coupling of NSUN2 and ALYREF. As mentioned in the introduction, Huang et al. (2024) showed that NSUN2-mediated m^5C methylation of Nrf2 mRNA influenced antioxidant stress responses in an ALYREF-dependent manner. Although studied in hepatocytes under Doxorubicin exposure, this illustrates a general principle: disruption of NSUN2 impairs ALYREF-mediated RNA handling, altering stress pathways and RNA metabolism. Applying this framework to DM1, the inability of mutant RNA to form mature foci in NSUN2 KO cells may reflect a similar failure of the NSUN2–ALYREF axis to stabilise and correctly process transcripts under stress conditions.

4.4 Limitations and Further Research

The main limitation of this project was the lack of successful biological repeats for ALYREF KD. As outlined previously, this was largely due to an insufficient reduction in ALYREF protein expression, which rendered the results inconclusive and unusable. In addition, a low cell count in the FISH assay was occasionally observed, most likely due to poor cell adhesion and subsequent cell loss during PBS washing steps.

Further studies should therefore aim to replicate the ALYREF KD and FISH experiments at least three additional times to generate a larger dataset and allow more robust comparisons. To ensure validity and comparability, these experiments should be conducted using the same protocols, cell line, and reagents described in

this project (see Section 2). If difficulties in achieving effective ALYREF KD persist (see Figure 9), the transfection protocol (see Section 2.1.5) should be troubleshooted and systematically optimised. Future work should also investigate complementary approaches, such as RNA immunoprecipitation or transcriptome-wide mapping, to clarify whether ALYREF directly regulates foci formation or stability and to determine whether m^5C readers represent viable therapeutic targets for mitigating RNA toxicity in DM1.

In addition to the technical limitations, an alternative explanation for the observed changes in RNA foci dynamics following ALYREF KD is the potential alteration of wider gene transcription. As ALYREF plays a central role in mRNA export and transcript processing, its KD may indirectly influence transcriptional output or transcript abundance, rather than directly modulating foci formation or stability. Changes in foci number, area, or intensity may therefore reflect altered levels of mutant DMPK transcripts or widespread dysregulation of RNA metabolism, rather than direct modification-specific effects. While this possibility was not directly assessed in the present study, future work incorporating transcriptional profiling approaches, such as RT-qPCR or RNA sequencing, would be necessary to distinguish between direct epitranscriptomic regulation of RNA foci and secondary effects arising from altered gene transcription.

With respect to *NSUN2*, the data obtained in this project provides a stronger foundation. Ongoing work in the laboratory is focusing on the effect of NSUN2 KO on R-loop formation. R-loops are three-stranded nucleic acid structures that form co-transcriptionally and occupy approximately 5% of the mammalian genome (Sanz et al., 2016). They have been implicated in both physiological and pathological processes, as they can stall DNA replication and transcription, disrupt 3′ end processing, promote genome instability, and drive repeat expansions (García-Muse & Aguilera, 2019). Consequently, R-loops are associated with a wide range of human diseases, including neurodegenerative disorders, cancer, and autoimmune conditions (García-Muse & Aguilera, 2019).

In the context of DM1, R-loops are predicted to form at the CTG repeat expansion within DMPK. Preliminary findings from this laboratory suggest that NSUN2 KO enhances R-loop formation at the *DMPK* locus while simultaneously reducing m^5C methylation, indicating an inverse correlation between m^5C modification and R-loop accumulation. These observations highlight the potential importance of NSUN2-mediated RNA methylation in modulating R-loop dynamics in DM1. It would also be informative to investigate the effect of ALYREF KD on R-loop formation.

In parallel, it would be valuable to investigate the effect of TET2 KD/KO in DM1 cells. As outlined in Section 1.3.5, TET2 is considered the third component of the m^5C epitranscriptomic pathway in DM1. Functioning as an 'eraser,' TET2 removes the modified mRNA through demethylation. Such experiments would provide a more comprehensive understanding of the role of the m^5C modification and contribute to a broader perspective on epitranscriptomic regulation in DM1 pathogenesis. I predict that TET2 KD or KO would lead to increases in all foci characteristics. In the absence of TET2, the m^5C epitranscriptomic modification would persist on the mutant RNA, promoting its nuclear retention, and ALYREF-mediated mRNA export would likely be insufficient to counteract the accumulation of mutant transcripts.

4.5 Conclusion

In summary, this study suggests that ALYREF and NSUN2 act as interdependent regulators of RNA metabolism in DM1, with effects that are both dose- and context-dependent. While ALYREF KD produced variable outcomes, NSUN2 KO consistently reduced foci, consistent with its role in enabling ALYREF function via m^5C modification. These findings support a model in which the NSUN2-ALYREF interaction, together with TET2, orchestrates the addition, interpretation, and removal of m^5C marks, thereby influencing RNA aggregation, export, and genome stability in DM1.

5 Bibliography

Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, *et al.* (2012). Mutations in NSUN2 cause autosomal-recessive intellectual disability. American Journal of Human Genetics. 90(5): 847-855.

Akhtar J, Lugoboni M, Junion G. (2021). m6A RNA modification in transcription regulation. *Transcription*. 12(5): 266-276.

Alsaggaf R, Pfeiffer RM, Pearce EE, Greene MH, Lochmuller H, Gadalla SM. (2024). Mortality Trends and Causes of Death in Myotonic Dystrophy Type 1 Patients From the UK Clinical Practice Research Datalink. *Muscle & Nerve.* 71: 229-236.

Antonucci L, D'Amico D, Di Magno L, Coni S, Di Marcotullio L, Cardinali B, *et al.* (2013). CNBP regulates wing development in Drosophila melanogaster by promoting IRES-dependent translation of dMyc. *Cell Cycle*. 13(3): 434-439.

Armas P, Agüero TH, Borgognone M, Aybar MJ, Calcaterra NB. (2008). Dissecting CNBP, a zinc-finger protein required for neural crest development, in its structural and functional domains. *Journal of Molecular Biology*. 382(4):1043-1056.

Balasubramanyam A, Iyer D, Stringer JL, Beaulieu C, Potvin A, Neumeyer AM, *et al.* (1998). Developmental changes in expression of myotonic dystrophy protein kinase in the rat central nervous system. *Journal of Comparative Neurology*. 394(3): 309-25.

Ballester-Lopez A, Núñez-Manchón J, Koehorst E, Linares-Pardo I, Almendrote M, Lucente G, *et al.* (2020). Three-dimensional imaging in myotonic dystrophy type 1: Linking molecular alterations with disease phenotype. *Neurology Genetics*. 6(4): 484.

Benhalevy D, Gupta SK, Danan CH, Ghosal S, Sun HW, Kazemier HG, *et al.* (2017). The Human CCHC-type Zinc Finger Nucleic Acid-Binding Protein Binds G-Rich Elements in Target mRNA Coding Sequences and Promotes Translation. *Cell Reports*. 18(12): 2979-2990.

Berger DS, Ladd AN. (2012). Repression of nuclear CELF activity can rescue CELF-regulated alternative splicing defects in skeletal muscle models of myotonic dystrophy. *PLoS Currents*. 4: RRN1305.

Berson A, Goodman LD, Sartoris AN, Otte CG, Aykit JA, Lee VM, *et al.* (2019). Drosophila Ref1/ALYREF regulates transcription and toxicity associated with ALS/FTD disease etiologies. *Acta Neuropathologica Communications*. 7(1): 65.

Blanco S, Kurowski A, Nichols J, Watt FM, Benitah SA, Frye M. (2011). The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. *PLoS Genetics*. 7(12): e1002403.

Blanco S, Dietmann S, Flores JV, Hussain S, Kutter C, Humphreys P, *et al.* (2014). Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. *EMBO Journal*. 33(18): 2020-2039.

Bosè F, Renna LV, Fossati B, Arpa G, Labate V, Milani V, *et al.* (2019). TNNT2 Missplicing in Skeletal Muscle as a Cardiac Biomarker in Myotonic Dystrophy Type 1 but Not in Myotonic Dystrophy Type 2. *Frontiers in Neurology.* 10: 992.

Botta A, Rinaldi F, Catalli C, Vergani L, Bonifazi E, Romeo V, *et al.* (2008). The CTG repeat expansion size correlates with the splicing defects observed in muscles from myotonic dystrophy type 1 patients. *Journal of Medical Genetics*. 45(10): 639.

Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, *et al.* (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3 ′ end of a transcript encoding a protein kinase family member. *Cell*. 68(4): 799-808.

Brzezicha B, Schmidt M, Makalowska I, Jarmolowski A, Pienkowska J, Szweykowska-Kulinska Z. (2006). Identification of human tRNA:m5C methyltransferase catalysing intron-dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic Acids Research*. 34(20): 6034-6043.

Buj-Bello A, Furling D, Tronchère H, Laporte J, Lerouge T, Butler-Browne GS, Mandel JL. (2002). Muscle-specific alternative splicing of myotubularin-related 1 gene is impaired in DM1 muscle cells. *Human Molecular Genetics*. 11(19): 2297-2307.

Burian HM & Burns CA. (1967). Ocular changes in myotonic dystrophy. *American Journal of Ophthalmology*. 63(1): 22-34.

Bush EW, Helmke SM, Birnbaum RA, Perryman MB. (2000). Myotonic Dystrophy Protein Kinase Domains Mediate Localization, Oligomerization, Novel Catalytic Activity, and Autoinhibition. *Biochemistry*. 39(29): 8480-8490.

Cardani R, Mancinelli E, Rotondo G, Sansone V, Meola G. (2006). Muscleblind-like protein 1 nuclear sequestration is a molecular pathology marker of DM1 and DM2. *European Journal of Histochemistry*. 50(3): 177-82.

Cardani R, Giagnacovo M, Rossi G, Renna LV, Bugiardini E, Pizzamiglio C, *et al.* (2014). Progression of muscle histopathology but not of spliceopathy in myotonic dystrophy type 2. *Neuromuscular disorders*. 24(12): 1042-1053.

Carrascosa-Sàez M, Colom-Rodrigo A, González-Martínez I, Pérez-Gómez R, García-Rey A, Piqueras-Losilla D, *et al.* (2025). Use of HSALR female mice as a model for the study of myotonic dystrophy type 1. *Lab Animals*. 54: 92-102.

Charlet BN, Savkur RS, Singh G, Philips AV, Grice EA, Cooper TA. (2002). Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. *Molecular Cell*. 10(1): 45-53.

Chen W, Liang Y, Deng W, Shimizu K, Ashique AM, Li E, Li YP. (2003). The zinc-finger protein CNBP is required for forebrain formation in the mouse. *Development*. 130(7): 1367-1379.

Chen W, Wang Y, Abe Y, Cheney L, Udd B, Li YP. (2007). Haploinsuffciency for Znf9 in Znf9+/− Mice Is Associated with Multiorgan Abnormalities Resembling Myotonic Dystrophy. *Journal of Molecular Biology.* 368(1): 8-17.

Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, *et al.* (2019) 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. *Nature Cell Biology*. 8(21): 978-990.

Chen Z, Morris HR, Polke J, Wood NW, Gandhi S, Ryten M, *et al.* (2025). Repeat Expansion Disorders. *Practical Neurology.* 25: 204-216.

Chokkalla AK, Mehta SL, Vemuganti R. (2020). Epitranscriptomic regulation by m(6)A RNA methylation in brain development and diseases. *Journal of Cerebral Blood Flow & Metabolism.* 40(12): 2331-2349.

Chokkalla AK, Mehta SL, Vemuganti R. (2021). Epitranscriptomic Modifications Modulate Normal and Pathological Functions in CNS. *Translational Stroke Research*. 13(1): 1-11.

Crisafulli S, Sultana J, Fontana A, Salvo F, Messina S, Trifiro G. (2020). Global epidemiology of Duchenne muscular dystrophy: an updated systematic review and meta-analysis. *Orphanet Journal of Rare Diseases*. 15: 141.

Cumming SA, Hamilton MJ, Robb Y, Gregory H, McWilliam C, Cooper A, *et al.* (2018). De novo repeat interruptions are associated with reduced somatic instability and mild or absent clinical features in myotonic dystrophy type 1. *European Journal of Human Genetics*. 26: 1635-1647.

Cumming SA, Jimenez-Moreno C, Okkersen K, Wenninger S, Daidj F, Hogarth F, *et al.* (2019). Genetic determinants of disease severity in the myotonic dystrophy type 1 OPTIMISTIC cohort. *Neurology Journals*. 93(10): 995-1009.

David AP, Pipier A, Pascutti F, Binolfi A, Weiner AMJ, Challier E, *et al.* (2019). CNBP controls transcription by unfolding DNA G-quadruplex structures. *Nucleic Acids Research*. 47(15): 7901-7913.

Day JW, Ricker K, Jacobsen JF, Rasmussen LJ, Dick KA, Kress W, *et al.* (2003). Myotonic dystrophy type 2: molecular, diagnostic and clinical spectrum. *Neurology Journals*. 60(4): 657-664.

de Die-Smulders CE, Höweler CJ, Thijs C, Mirandolle JF, Anten HB, Smeets HJM, *et al.* (1998). Age and Causes of Death in Adult-Onset Myotonic Dystrophy. *Brain*. 121: 1557-1563.

Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Massé A, *et al.* (2009). Mutation in TET2 in myeloid cancers. New England Journal of Medicine. 360(22): 2289-2301.

Depienne C & Mandel JL. (2021). 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *AJHG*. 108(5): 764-785.

Deplus R, Delatte B, Schwinn MK, Defrance M, Méndez J, Murphy N, *et al.* (2013). TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. EMBO Journal. 32(5): 645-655.

Desrosiers R, Friderici K, Rottman F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences of the USA*. 71(10): 3971-3975.

Dhaenens CM, Schraen-Maschke S, Tran H, Vingtdeux V, Ghanem D, Leroy O, *et al.* (2008). Overexpression of MBNL1 fetal isoforms and modified splicing of Tau in the DM1 brain: two individual consequences of CUG trinucleotide repeats. *Experimental Neurology*. 210(2): 467-478.

Dodge PR, Gamstorp I, Byers RK, Russell P. (1966). Myotonic dystrophy in infancy and childhood. *Paediatrics*. 35: 3-19. (Cited from Harper, 1989)

Dogan C, De Antonio M, Hamroun D, Varet H, Fabbro M, Rougier F, *et al.* (2016). Gender as a modifying factor influencing myotonic dystrophy type 1 phenotype severity and mortality: a nationwide multiple databases cross-sectional observational study. *PLoS ONE*. 11(2): 148264.

Dunn LJ, Dierker LI. (1973). Recurrent hydramnios in association with myotonia dystrophica. *Obstetrics & Gynecology*. 42(1): 104-106.

El Boujnouni N, Ripken L, Willemse M, van der Sanden B, Neveling K, Hoischen A, *et al.* (2025). Repeat length as a key determinant for disease severity and antisense oligonucleotide activity in myotonic dystrophy type 1. *Molecular Therapy.* 33(3): 101502.

Ellerby LM. (2020). Repeat Expansion Disorders: Mechanisms and Therapeutics. *Neurotherapeutics.* 16: 924-927.

Emery AE. (2002). The muscular dystrophies. *Lancet*. 359(9307): 687-695.

Falcetta D, Quirim S, Cocchiararo I, Chabry F, Théodore M, Stiefvater A, *et al.* (2024). CaMKIIβ deregulation contributes to neuromuscular junction destabilization in Myotonic Dystrophy type I. *Skeletal Muscle*. 14(1): 11.

Fan J, Wang K, Du X, Wang J, Chen S, Wang Y, *et al.* (2019). ALYREF links 3'-end processing to nuclear export of non-polyadenylated mRNAs. *EMBO Journal*. 38(9): 99910.

Fortune MT, Vassilopoulos C, Coolbaugh MI, Siciliano MJ, Monckton DG. (2000). Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. *Human Molecular Genetics.* 9(3): 439-445.

Freyermuth F, Rau F, Kokunai Y, Linke T, Sellier C, Nakamori M, *et al.* (2016). Splicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy. *Nature Communications.* 7: 11067.

Frye M, Watt FM. (2006). The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Current Biology*. 16(10): 971-981.

Frye M, Dragoni I, Chin SF, Spiteri I, Kurowski A, Provenzano E, *et al.* (2010). Genomic gain of 5p15 leads to over-expression of Misu (NSUN2) in breast cancer. *Cancer Letters*. 289(1): 71-80.

Fugier C, Klein AF, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, *et al.* (2011). Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nature Medicine*. 17: 720-725.

Gao Z, Cooper TA. (2013). Reexpression of pyruvate kinase M2 in type 1 myofibers correlates with altered glucose metabolism in myotonic dystrophy. *Proceedings of the National Academy of Sciences of the USA*. 110(33): 13570-13575.

Gaj T, Gersbach CA, Barbas CF 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology*. 31(7): 397-405.

García-Muse T, Aguilera A. (2019). R Loops: From Physiological to Pathological Roles. *Cell*. 179(3): 604-618.

Goedert M, Spillantini MG, Jakes R, Rutherford D, Crowther RA. (1989). Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron*. 3(4): 519-526.

Goers ES, Purcell J, Voelker RB, Gates DP, Berglund JA. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. *Nucleic Acids Research*. 38(7): 2467-2484.

Goodwin M, Mohan A, Batra R, Lee KY, Charizanis K, Fernández Gómez FJ, *et al.* (2015). MBNL Sequestration by Toxic RNAs and RNA Misprocessing in the Myotonic Dystrophy Brain. *Cell Rep*orts. 12(7): 1159-1568.

Greenfield J. (1911). Notes on a family with myotonia atrophica and early cataract with report of an additional case of myotonia atrophica. *Revue Neurologiue Psychiatric*. 9: 169-181. (Cited from Harper, 1989)

Groh WJ, Groh MR, Saha C, Kincaid JC, Simmons Z, Ciafaloni E, *et al.* (2008). Electrocardiographic Abnormalities and Sudden Death in Myotonic Dystrophy Type 1. *The New England Journal of Medicine*. 358(25): 2688-2697.

Handal T, Juster S, Abu Diab M, Yanovsky-Dagan S, Zahdeh F, Aviel U, *et al.* (2024). Differentiation shifts from a reversible to an irreversible heterochromatin state at the DM1 locus. *Nature Communications*. 15(1): 3270.

Hannon GJ. (2002). RNA interference. *Nature*. 418(6894): 244-251.

Hanoun S, Sun Y, Ebrahimi F, Ghasemi M. (2022). Speech and language abnormalities in myotonic dystrophy: An Overview. *The Journal of Clinical Neuroscience*. 96: 212-220.

Harley HG, Rundle SA, MacMillan JC, Myring J, Brook JD, Crow S, *et al.* (1993). Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *The American Journal of Human Genetics.* 52(6): 1164-1174.

Harper PS. (1975). Congenital myotonic dystrophy in Britain 2.Genetic Basis. *Archives of Disease in Childhood*. 50: 514-521. (Cited from Harper 1989).

Harper PS. (1989). Myotonic dystrophy: 2nd edition. W.B. Saunders. 1-28.

Hartshorne DJ, Ito M, Erdodi F. (1998). Myosin light chain phosphatase: subunit composition, interactions and regulation. *The Journal of Muscle Research and Cell Motility*. 19(4): 325-341.

Hawkins AM, Hawkins CL, Abdul Razak K, Khoo TK, Tran K, Jackson RV. (2019). Respiratory dysfunction in myotonic dystrophy type 1: A systematic review. *Neuromuscular Disorders*. 29(3): 198-212.

Heller SA, Shih R, Kalra R, Kang PB. (2020). Emery-Dreifuss muscular dystrophy. *Muscle Nerve*. 61(4): 436-448.

Ho TH, Bundman D, Armstrong DL, Cooper TA. (2005). Transgenic mice expressing CUG-BP1 reproduce splicing mis-regulation observed in myotonic dystrophy. *Human Molecular Genetics*. 14(11): 1539-1547.

Ho G, Carey KA, Cardamone M, Farrar MA. (2019). Myotonic dystrophy type 1: clinical manifestations in children and adolescents. *Archives of Disease in Childhood*. 104(1): 48-52.

Hoffman EP, Fischbeck HK, Brown MD, Johnson M, Medori R, Loike JD, *et al.* (1988). Characterisation of dystrophin in muscle biopsy specimens from patients with Duchenne's or Becker muscular dystrophy. *New England Journal of Medicine*. 318: 1363-1368. (Cited from Harper, 1989)

Hojka-Osinska A, Chlebowski A, Grochowska J, Owczarek EP, Affek K, Kłosowska-Kosicka K, *et al.* (2021). Landscape of functional interactions of human processive ribonucleases revealed by high-throughput siRNA screenings. iScience. 24(9): 103036.

Höweler CJ, Busch HFM, Geraedts JPM, Niermeijer MF, Staal A. (1989). Anticipation in Myotonic Dystrophy: Fact or Fiction?. *Brain*. 112(3): 779-797.

Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, *et al.* (2013). Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell*. 155(7): 1545-1555.

Huang Y, Li X, Wei L, Ma S, Ma L, Zan Y, He X, Tang, Ding Y. (2024). NSUN2 relies on ALYREF to regulate Nrf2-mediated oxidative stress and alleviate Dox-induced liver injury. *Biology Direct*. 19(1): 32.

Huml RA. (2015). Muscular Dystrophy: A Concise Guide. Springer. 1: 1-4/ 2: 5-8/ 5: 37-40.

Hunter A, Tsilfidis C, Mettler G, Jacob P, Mahadevan M, Surh L, Korneluk R. (1992). The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *Journal of Medical Genetics*. 29(11): 774-779.

Imbriano C, Moresi V, Belluti S, Renzini A, Cavioli G, Maretti E, Molinari S. (2023). Epitranscriptomics as a New Layer of Regulation of Gene Expression in Skeletal Muscle: Known Functions and Future Perspectives. International Journal of Molecular Science. 24(20): 15161.

Ito S, Li S, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. (2011). Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science*. 333(6047): 1300-1303.

Iyama S, Chi S, Idogawa M, Ikezoe T, Fukushima K, Utsu Y, *et al.* (2025). Prognostic impact of TET2 mutations in patients with acute myeloid leukemia: HM-SCREEN-Japan 01 and 02 study. Annals of Hematology. 104(1): 275-284.

Jacobs D, Willekens D, de Die-Smulders C, Frijns J, Steyaert J. (2017). Delusional and psychotic disorders in juvenile myotonic dystrophy type-1. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. 174(4): 359-366.

Jankowska AM, Szpurka H, Tiu RV, Makishima H, Afable M, Huh J, *et al.* (2009). Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood*. 113(25): 6403-6410.

Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. (2004). Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Human Molecular Genetics*. 13(24): 3079-3088.

Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Yang C, Chen Y. (2021). The role of m6A modification in the biological functions and diseases. *Signal Transduction Target Therapy*. 6(1): 74.

Johnson NE, Butterfield RJ, Mayne K, Newcomb T, Imburgia C, Dunn D, *et al.* (2021). Population-Based Prevalence of Myotonic Dystrophy Type 1 Using Genetic Analysis of Statewide Blood Screening Program. *Neurology*. 96(7): e1045-e1053.

Joosten IBT, Hellebrekers DMEI, de Greef BTA, Smeets HJM, de Die-Smulders CEM, Faber CG, Gerrits MM. (2020). Parental repeat length instability in myotonic

dystrophy type 1 pre- and protomutations. *European Journal of Human Genetics*. 28(7): 956-962.

Joshi K, Liu S, Breslin SJP, Zhang J. (2022). Mechanisms that regulate the activities of TET proteins. *Cellular and Molecular Life Sciences*. 79(7): 363.

Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, Cooper TA. (2008). A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proceedings of the National Academy of Sciences of the United States of America*. 105(51): 20333-20338.

Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, *et al.* (2003) A muscleblind knockout model for myotonic dystrophy. Science, 302, 1978-1980.

Kimura K, Ito M, Amano M, Chihara K, Fukata Y, Nakafuku M, *et al.* (1996). Regulation of myosin phosphatase by Rho and Rho-associated kinase (Rho-kinase). *Science*. 273(5272): 245-248.

Kimura T, Nakamori M, Lueck JD, Pouliquin P, Aoike F, Fujimura H, *et al.* (2005). Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+-ATPase in myotonic dystrophy type 1. *Human Molecular Genetics*. 14(15): 2189-2200.

Kino Y, Mori D, Oma Y, Takeshita Y, Sasagawa N, Ishiura S. (2004). Muscleblind protein, MBNL1/EXP, binds specifically to CHHG repeats. *Human Molecular Genetics*. 13(5): 495-507.

Kiyan E, Okumus G, Cuhadaroglu C, Deymeer F. (2010). Sleep apnea in adult myotonic dystrophy patients who have no excessive daytime sleepiness. *Sleep Breath*. 14(1): 19-24.

Koebis M, Ohsawa N, Kino Y, Sasagawa N, Nishino I, Ishiura S. (2011). Alternative splicing of myomesin 1 gene is aberrantly regulated in myotonic dystrophy type 1. *Genes Cells*. 16: 961-972.

Klug WS, Cummings MR, Spencer CA, Palladino MA, Killian DJ. (2020). Concepts of Genetics. *Pearson Education Limited*. 4: 121, 456.

Koch MC, Grimm T, Harley HG, Harper PS. (1991). Genetic risks for children of women with myotonic dystrophy. *American Journal of Human Genetics*. 48(6): 1084-1091.

Konieczny P, Stepniak-Konieczna E, Sobczak K. (2014). MBNL proteins and their target RNAs, interaction and splicing regulation. *Nucleic Acids Research*. 42(17): 10873-10887.

Kong L, Tan L, Lv R, Shi Z, Xiong L, Wu F, *et al.* (2016). A primary role of TET proteins in establishment and maintenance of De Novo bivalency at CpG islands. Nucleic Acids Research. 44(18): 8682-8692.

Kow RL, Black AH, Saxton AD, Liachko NF, Kraemer BC. (2022). Loss of aly/ALYREF suppresses toxicity in both tau and TDP-43 models of neurodegeneration. *GeroScience*. 44(2): 747-761.

Lam LT, Pham YCN, Nguyen thi Man, Morris GE. (2000). Characterization of a monoclonal antibody panel shows that the myotonic dystrophy protein kinase, DMPK, is expressed almost exclusively in muscle and heart. *Human Molecular Genetics*. 9(14) 2167-2173.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature*. 409(6822): 860-921.

Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, *et al.* (2013). Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Molecular Medicine*. 5(12): 1887-1900.

Li Y, Xue M, Deng X, Dong L, Nguyen LXT, Ren L, *et al.* (2024). TET2-mediated mRNA demethylation regulates leukaemia stem cell homing and self-renewal. *Cell Stem Cell*. 30: 1072–1090.

Liao Q, Zhang Y, He J, Huang K. (2022). Global Prevalence of Myotonic Dystrophy: An Updated Systematic Review and Meta-Analysis. *Neuroepidemiology*. 56(3): 163-173.

Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, *et al.* (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Human Molecular Genetics*. 15(13): 2087-2097.

Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL. *et al.* (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science*. 293(5531): 816-817.

Liu Y, Yang Y, Wu R, Gao CC, Liao X, Han X, *et al.* (2022). mRNA m 5 C inhibits adipogenesis and promotes myogenesis by respectively facilitating YBX2 and SMO mRNA export in ALYREF-m 5 C manner. *Cellular and Molecular Life Sciences*. 79(9): 481.

Lo Mauro A & Aliverti A. (2016). Physiology of respiratory disturbances in muscular dystrophies. *Breathe*. 12(4): 318-327.

López-Martínez A, Soblechero-Martin P, de-la-Puente-Ovejero L, Nogales-Gadea G, Archavala-Gomeza V. (2020). An Overview of Alternative Splicing Defects Implicated in Myotonic Dystrophy Type 1. *Genes*. 11(9): 1109.

Lueck JD, Mankodi A, Swanson MS, Thornton CA, Dirksen RT. (2007). Muscle chloride channel dysfunction in two mouse models of myotonic dystrophy. *Journal of General Physiology*. 129(1): 79-94.

Mah JK, Korngut L, Fiest KM, Dykeman J, Day LJ, Pringsheim T, Jette N. (2016). A Systematic Review and Meta-analysis on the Epidemiology of the Muscular Dystrophies. *Canadian Journal of Neurological Sciences*. 43(1): 163-177.

Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, *et al.* (1992). Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3' Untranslated Region of the Gene. *Science*. 255(5049): 1253-1255.

Mahadevan MS, Yadava RS, Mandal M. (2021). Cardiac Pathology in Myotonic Dystrophy Type 1. *International Journal of Molecular Science*. 22(21): 11874.

Malerba A, Klein P, Bachtarzi H, Jarmin SA, Cordova G, Ferry A, *et al.* (2017). PABPN1 gene therapy for oculopharyngeal muscular dystrophy. *Nature Communications*. 8: 14848.

Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, *et al.* (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Human Molecular Genetics*. 10(19): 2165-2170.

Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, *et al.* (2002). Expanded CUG Repeats Trigger Aberrant Splicing of ClC-1 Chloride Channel Pre-mRNA and Hyperexcitability of Skeletal Muscle in Myotonic Dystrophy. *Molecular Cell*. 10(1): 35-44.

Mankodi A, Teng-Umnuay P, Krym M, Henderson D, Swanson M, Thornton CA. (2003). Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2. *Annals of Neurology*. 54(6): 760-768.

Martorell L, Martinez JM, Carey N, Johnson K, Baiget M. (1995). Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five-year period. *Journal of Medical Genetics*. 32(8): 593-596.

Martorell L, Monckton DG, Gamez J, Baiget M. (2000). Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. *European Journal of Human Genetics*. 8(6): 423-430.

Mathieu J, Allard P, Potvin L, Prévost C, Bégin P. (1999). A 10-year study of mortality in a cohort of patients with myotonic dystrophy. *Neurology*. 52(8): 1658-1662.

Meola G, Cardani R. (2015). Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochimica et Biophysica Acta*. 1852(4): 594-606.

Meyer KD, Jaffrey SR. (2014). The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nature Review Molecular Cell Biology*. 15(5): 313-26.

Michalowski S, Miller JW, Urbinati CR, Paliouras M, Swanson MS, Griffith J. (1999). Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Research*. 27(17): 3534-3542.

Miller JW, Urbinati CR, Teng-umnuay P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS. (2000). Recruitment of human muscleblind proteins to CUG(n) expansions associated with myotonic dystrophy. *EMBO Journal*. 19(17): 4439-4448.

Milone M & Liewluck T. (2019). The unfolding spectrum of inherited distal myopathies. *Muscle Nerve*. 59(3): 283-294.

Misra C, Bangru S, Lin F, Lam K, Koenig SN, Lubbers ER, *et al.* (2020). Aberrant Expression of a Non-muscle RBFOX2 Isoform Triggers Cardiac Conduction Defects in Myotonic Dystrophy. *Developmental Cell*. 52(6): 748-763.

Monckton DG, Wong LJ, Ashizawa T, Caskey CT. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. Human Molecular Genetics. 4(1): 1-8.

Mounsey JP, John 3rd JE, Helmke SM, Bush EW, Gilbert J, Roses AD, *et al.* (2000). Phospholemman Is a Substrate for Myotonic Dystrophy Protein Kinase. *Journal of Biological Chemistry*. 275(30): 23362-23367.

Murányi A, Zhang R, Liu F, Hirano K, Ito M, Epstein HF, Hartshorne DJ. (2001). Myotonic dystrophy protein kinase phosphorylates the myosin phosphatase targeting subunit and inhibits myosin phosphatase activity. *FEBS Letters*. 493(2): 80-84.

Muravenko OV, Gizatullin RZ, Al-Amin AN, Protopopov AI, Kashuba VI, Zelenin AV, Zabarovsky ER. (2000). Human ALY/BEF gene Map position 17q25.3. *Chromosome Research*. 8(6): 562.

Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, *et al.* (2013). Splicing biomarkers of disease severity in myotonic dystrophy. *Annals of Neurology*. 74(6): 862-872.

Norman A, McNeil S, Lunt P. (1989). The genetic basis of myotonic dystrophy: Evidence for a trinucleotide repeat expansion. *Nature*. 337(6205): 346-348.

Norwood FL, Harling C, Chinnery PF, Eagle M, Bushby K, Straub V. (2009). Prevalence of genetic muscle disease in Northern England: in-depth analysis of a muscle clinic population. *Brain*. 132(11): 3175-3186.

Novelli G, Gennarelli M, Fattorini C, Abbruzzese C, Dallapiccola B. (1993). The Dynamic Genomics of Myotonic Dystrophy and its Clinical Relevance: An Overview. *Biomedicine & Pharmacotherapy*. 47(8): 321-330.

Núñez-Manchón J, Capó J, Martínez-Piñeiro A, Juanola E, Pesovic J, Mosqueira-Martín L, González-Imaz K, *et al.* (2024). Immortalized human myotonic dystrophy type 1 muscle cell lines to address patient heterogeneity. *iScience*. 27(6): 109930.

Ohsawa N, Koebis M, Suo S, Nishino I, Ishiura S. (2011). Alternative splicing of PDLIM3/ALP, for α-actinin-associated LIM protein 3, is aberrant in persons with myotonic dystrophy. Biochemical and Biophysical Research Communications. 409(1): 64-69.

Pavićević D.S, Miladinović J, Brkušanin M, Šviković S, Djurica S, Brajušković G, Romac S. (2013). Molecular genetics and genetic testing in myotonic dystrophy type 1. *BioMed Research International.* 2013: 391821.

Ranum LPW, Rasmussen P, Benzow K, Koob M, Day JW. (1998). Genetic mapping of a second myotonic dystrophy locus. *Nature Genetics*. 19(2): 196-198.

Richard P, Broustet A, Hamroun D, Touraine R, Kuntzer T. (1998). The poly(A)-binding protein (PABPN1) gene is responsible for oculopharyngeal muscular dystrophy. *Nature Genetics*. 18: 332-335.

Rosas MG, Centola C, Torres M, Mougular VS, David AP, Piga EJ, *et al.* (2024). The transcription of the main gene associated with Treacher–Collins syndrome (TCOF1) is regulated by G-quadruplexes and cellular nucleic acid binding protein (CNBP). *Scientific Reports*. 14: 7472.

Rossi S, Della Marca G, Ricci M, Perna A, Nicoletti TF, Brunetti V, *et al.* (2019). Prevalence and predictor factors of respiratory impairment in a large cohort of patients with Myotonic Dystrophy type 1 (DM1): A retrospective, cross-sectional study. *Journal of the Neurological Sciences*. 399: 118-124.

Russo V, Sperlongano S, Gallinoro E, Rago A, Papa AA, Golino P, *et al.* (2020). Prevalence of left ventricular systolic dysfunction in myotonic dystrophy type 1: a systematic review. *Journal of Cardiac Failure*. 26(10): 849-856.

Sacconi S, Lemmers RJ, Balog J, van der Vliet PJ, Lahaut P, van Nieuwenhuizen MP, *et al.* (2013). The FSHD2 gene SMCHD1 is a modifier of disease severity in families affected by FSHD1. *American Journal of Human Genetics*. 93(4): 744-751.

Salari N, Fatahi B, Valipour E, Kazeminia M, Fatahian R, Kiaei A, *et al.* (2022). Global prevalence of Duchenne and Becker muscular dystrophy: a systematic review and meta-analysis. *Journal of Orthopaedic Surgery and Research*. 17(1): 96.

Salisbury E, Schoser B, Schneider-Gold C, Wang GL, Huichalaf C, Jin B, *et al.* (2009). Expression of RNA CCUG repeats dysregulates translation and degradation of proteins in myotonic dystrophy 2 patients. American Journal of Pathology. 175(2): 748-762. doi:

Sanz LA, Hartono SR, Lim YW, Steyaert S, Rajpurkar A, Ginno PA, *et al.* (2016). Prevalent, Dynamic, and Conserved R-Loop Structures Associate with Specific Epigenomic Signatures in Mammals. *Molecular Cell*. 63(1): 167-178.

Savkur RS, Philips AV, Cooper TA. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Natural Genetics*. 29(1): 40-47.

Savkur RS, Philips AV, Cooper TA, Dalton JC, Moseley ML, Ranum LP, Day JW. (2004). Insulin receptor splicing alteration in myotonic dystrophy type 2. *American Journal of Human Genetics*. 74(6): 1309-1313.

Shaw DJ, Harper PS. (1989). Myotonic dystrophy: developments in molecular genetics. *British Medical Bulletin*. 45(3): 745-759.

Shi M, Zhang H, Wu X, He Z, Wang L, Yin S, *et al.* (2017). ALYREF mainly binds to the 5′ and the 3′ regions of the mRNA in vivo. *Nucleic Acids Research*. 45(16): 9640-9653.

Shi Q, Dai M, Ma Y, Liu J, Liu X, Wang XJ. (2024). DRED: A Comprehensive Database of Genes Related to Repeat Expansion Diseases. *Geomics Proteomics Bioinformatics.* 22(5): qzae068.

Shi Y, Niu Y, Zhang P, Luo H, Liu S, Zhang S, *et al.* (2023). Characterization of genome-wide STR variation in 6487 human genomes. *Nature Communications*. 14(1): 2092.

Sordyl D, Boileau E, Bernat A, Maiti S, Mukherjee S, Moafinejad SN, *et al.* (2026). MODOMICS: a database of RNA modifications and related information. 2025 update and 20th anniversary. *Nucleic Acids Research*. 54(D1): D219-D225.

Steinert H. (1909). Myopathologische Beiträge 1. Über das klinische und anatomische Bild des Muskelschwunds der Myotoniker. *Deutsche Zeitschrift für Nervenheilkunde*. 37: 58-104. (taken from Harper)

Stokes M, Varughese N, Iannaccone S, Castro D. (2019). Clinical and genetic characteristics of childhood-onset myotonic dystrophy. *Muscle & Nerve*. 60(6): 732-738.

Strachan T, Lucassen A. (2023). Genetics and Genomics in Medicine. *CRC Press*. 7: 194, 196-198.

Swanson MS, Cortese R, Dreyfuss G. (1993). RNA-binding proteins in human hereditary disease. *FASEB Journal*. 7(8): 672–678.

Sznajder ŁJ, Khan M, Ciesiołka A, Tadross M, Nutter CA, Taylor K, *et al.* (2025). Autism-related traits in myotonic dystrophy type 1 model mice are due to MBNL sequestration and RNA mis-splicing of autism-risk genes. *Nature Neuroscience*. 28(6): 1199-1212.

Taghizadeh E, Rezaee M, Barreto GE, Sahebkar A. (2019). Prevalence, pathological mechanisms, and genetic basis of limb-girdle muscular dystrophies: A review. *Journal of Cellular Physiology*. 234(6): 7874–7884.

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 324(5929): 930-935.

Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *Journal of Cell Biology*. 128(6): 995–1002.

Tang ZZ, Yarotskyy V, Wei L, Sobczak K, Nakamori M, Eichinger K, Moxley RT, Dirksen RT, Thornton CA. (2012). Muscle weakness in myotonic dystrophy associated with misregulated splicing and altered gating of Ca(V)1.1 calcium channel. *Human Molecular Genetics*. 21(6): 1312-1324.

Tang Y, Gao CC, Gao Y, Yang Y, Shi B, Yu JL, *et al.* (2020) OsNSUN2-mediated 5-methylcytosine mRNA modification enhances rice adaptation to high temperature. *Developmental Cell*. 3(53): 272-286.

Theadom A, Rodrigues M, Roxburgh R, Balalla S, Higgins C, Bhattacharjee R, *et al.* (2015). Prevalence of Muscular Dystrophies: A Systematic Literature Review. *Neuroepidemiology 1*. 43(3): 259-268.

Thornton CA, Griggs RC, Moxley RT 3rd. (1994). Myotonic dystrophy with no trinucleotide repeat expansion. *Annals of Neurology*. 35(3): 269-272.

Thorpe J, Osei-Owusu IA, Avigdor BE, Tupler R, Pevsner J. (2020). Mosaicism in Human Health and Disease. *Annual Reviews of Genetics*. 54: 487-510.

Tomé S, Gourdon G. (2020). DM1 Phenotype Variability and Triplet Repeat Instability: Challenges in the Development of New Therapies. International Journal of Molecular Science. 21(2): 457.

Tsien JZ, Huerta PT, Tonegawa S. (1996). The essential role of hippocampal CA1 NMDA receptor-dependent synaptic plasticity in spatial memory. *Cell*. 87: 1327-1338.

Udd B, Krahe R. (2012). The myotonic dystrophies: molecular, clinical, and therapeutic challenges. *The Lancet Neurology*. 11(10): 891-905.

Van der Maarel SM, Tawil R, Tapscott SJ. (2012). Facioscapulohumeral muscular dystrophy and DUX4: breaking the silence. *Trends in Molecular Medicine*. 17(5): 252-258.

Van Haute L, Lee SY, McCann BJ, Powell CA, Bansal D, Vasiliauskaitė L, *et al.* (2019). NSUN2 introduces 5-methylcytosines in mammalian mitochondrial tRNAs. *Nucleic Acids Research*. 47(16): 8720-8733.

Visconti VV, Centofanti F, Fittipaldi S, Macri E, Novelli G, Botta A. (2021). Epigenetics of Myotonic Dystrophies: A Minireview. *International Journal of Molecular Sciences*. 22(22): 12594.

Wagner SD, Struck AJ, Gupta R, Farnsworth DR, Mahady AE, Eichinger K, *et al.* (2016). Dose-Dependent Regulation of Alternative Splicing by MBNL Proteins Reveals Biomarkers for Myotonic Dystrophy. *PLoS Genetics*. 12(9): e1006316.

Wang ET, Cody NAL, Jog S, Biancolella M, Wang TT, Treacy DJ, *et al.* (2012). Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell*. 150(4): 710-724.

Wang ET, Treacy D, Eichinger K, Struck A, Estabrook J, Olafson H, *et al.* (2019). Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Human Molecular Genetics*. 28(8): 1312-1321.

Wang N, Chen R, Deng M, Wei W, Zhou Z, Ning K, *et al.* (2023). m5C-dependent cross-regulation between nuclear reader ALYREF and writer NSUN2 promotes urothelial bladder cancer malignancy through facilitating RABL6/TK1 mRNAs splicing and stabilization. *Cell Death and Disease*. 14(2): 139.

Wansink DG, van Herpen RE, Coerwinkel-Driessen MM, Groenen PJ, Hemmings BA, Wieringa B. (2003). Alternative splicing controls myotonic dystrophy protein kinase structure, enzymatic activity, and subcellular localization. *Molecular Cell Biology.* 23(16): 5489-5501.

Warf MB & Berglund JA. (2007). MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA*. 13(12): 2238-2251.

Winchester CL, Ferrier RK, Sermoni A, Clark BJ, Johnson KJ. (1999). Characterization of the Expression of DMPK and SIX5 in the Human Eye and Implications for Pathogenesis in Myotonic Dystrophy. *Human Molecular Genetics*. 8(3): 481-492.

Wojciechowska M and Krzyzosiak WJ. (2011). Cellular toxicity of expanded RNA repeats: focus on RNA foci. *Human Molecular Genetics*. 20(19): 3811-3821.

Wojciechowska M, Sobczak K, Kozlowski P, Sedehizadeh S, Wojtkowiak-Szlachcic A, Czubak K, *et al.* (2018). Quantitative Methods to Monitor RNA Biomarkers in Myotonic Dystrophy. *Scientific Reports*. 8: 5885

Wood L, Bassez G, Bleyenheuft C, Campbell C, Cossette L, Jimenez-Moreno AC, *et al.* (2018). Eight years after an international workshop on myotonic dystrophy patient

registries: Case study of a global collaboration for a rare disease. *Orphanet Journal of Rare Diseases*. 13: 155.

Xia M, Yan R, Wang W, Kong A, Zhang M, Miao Z, Ge W, Wan B, Xu X. (2023). The Tet2-Upf1 complex modulates mRNA stability under stress conditions. *Frontier Genetics*. 14: 1158954.

Xing X, Markus R, Ghosh T, Buxton S, Nieves DJ, Wojciechowska M, Brook JD. (2023). Regulation of Toxic RNA Foci and Mutant DMPK Transcripts: Role of MBNL Proteins and RNA Decay Pathways. *bioRxiv*. 559487.

Yadava RS, Frenzel-McCardell CD, Yu Q, Srinivasan V, Tucker AL, Puymirat J, *et al.* (2008). RNA toxicity in myotonic muscular dystrophy induces NKX2-5 expression. *Nature Genetics*. 40(1): 61-68.

Yamashita Y, Matsuura T, Shinmi J, Amakusa Y, Masuda A, Ito M, *et al.* (2012). Four parameters increase the sensitivity and specificity of the exon array analysis and disclose 25 novel aberrantly spliced exons in myotonic dystrophy. *Journal of Human Genetics*. 57(6): 368-374.

Yang C, Wang H, Qiao T, Yang B, Aliaga L, Qiu L, *et al.* (2014). Partial loss of TDP-43 function causes phenotypes of amyotrophic lateral sclerosis. Proceedings of the National Academy of Sciences of the United States of America. 111(12): E1121-1129.

Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, *et al.* (2017). 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell research*. 27(5): 606-625.

Yang Y, Wang L, Han X, Yang WL, Zhang M, Ma HL, *et al.* (2019). RNA 5-methylcytosine facilitates the maternal-to-zygotic transition by preventing maternal mRNA decay. *Molecular Cell*. 6(75): 1188-1202.

Yang Q, Zhang Q, Qin Z, Yi S, Luo J. (2024). A novel variant in NSUN2 causes intellectual disability in a Chinese family. *BMC Medical Genomics*. 17(1): 95.

Yang Y, Cao L, Xu X, Li D, Deng Y, Li L, *et al.* (2025). NSUN2/ALYREF axis-driven m5C methylation enhances PD-L1 expression and facilitates immune evasion in non-small-cell lung cancer. *Cancer Immunology and Immunotherapy*. 74(4): 132.

Yin Q, Wang H, Li N, Ding Y, Xie Z, Jin L, *et al.* (2020). Dosage effect of multiple genes accounts for multisystem disorder of myotonic dystrophy type 1. *Cell Research*. 30(2): 133-145.

Yu Z, Goodman LD, Shieh SY, Min M, Teng X, Zhu Y, Bonini NM. (2015). A fly model for the CCUG-repeat expansion of myotonic dystrophy type 2 reveals a novel interaction with MBNL1. *Human Molecular Genetics*. (24)4: 954-962.

Yuan Y, Compton SA, Sobczak K, Stenberg MG, Thornton CA, Griffith JD, Swanson MS. (2007). Muscleblindlike 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. *Nucleic Acids Research*. 35(16): 5474-5486.

Zambon AA, Muntoni F. (2021). Congenital muscular dystrophies: What is new?. *Neuromuscular Disorders*. 31(10): 931-942.

Zhang X, Liu Z, Yi J, Tang H, Xing J, Yu M, *et al.* (2012). The tRNA methyltransferase NSun2 stabilizes p16INK⁴ mRNA by methylating the 3'-untranslated region of p16. *Nature Communications*. 3: 712.

Zhang C, Chen Y, Sun B, Wang L, Yang Y, Ma D, *et al.* (2017). m6A modulates haematopoietic stem and progenitor cell specification. *Nature*. 549: 273-276.

Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, *et al.* (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*. 49(1): 18-29.

Zhong H, Tang HF, Kai Y. (2020). N6-methyladenine RNA Modification (m6A): An Emerging Regulator of Metabolic Diseases. *Current Drug Targets*. 21(11): 1056-1067.

Zou F, Tu R, Duan B, Yang Z, Ping Z, Song X, *et al.* (2020) Drosophila YBX1 homolog YPS promotes ovarian germ line stem cell development by preferentially recognizing 5-methylcytosine RNAs. *Proceedings of the National Academy of Sciences of USA*. 7(117): 3603-3609.

Zuo S, Li L, Wen X, Gu X, Zhuang A, Li R, *et al.* (2023). NSUN2-mediated m5C RNA methylation dictates retinoblastoma progression through promoting PFAS mRNA stability and expression. *Clinical and Translational Medicine*. 13(5): 1273.