MUSCLEBLIND IN DEVELOPMENT AND DISEASE

By

MICHAEL GUSTAVE POULOS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

To my family

ACKNOWLEDGMENTS

I would like to thank my wife, Ashley Poulos, for her unconditional love and support during my graduate school journey. I am also indebted to my parents, David and Eileen Poulos, and my brother, Nathan Poulos, for providing me with constant support and encouragement throughout my life. I am grateful to the mentors in my undergraduate and graduate studies that helped to foster my scientific development. At Grand Valley State University, Dr. Patrick Thorpe and Dr. Roderick Morgan introduced me to molecular biology and allowed me to explore my scientific curiosity in their laboratories, while teaching me the basic fundamentals of research in an encouraging and fun setting. I would especially like to thank my graduate mentor, Dr. Maury Swanson, who taught me how to critically evaluate and approach science, and has provided me with the necessary tools to be successful at the next level. I would also like to thank my committee members, Drs. Brian Harfe, William Hauswirth, and Jake Streit for providing guidance during my graduate career. Additionally, Molecular Genetics and Microbiology faculty Drs. Henry Baker, Rich Condit, Jim Resnick, and Al Lewin have always found time to talk science or to offer advice, and have greatly enriched my graduate school experience. Fellow Swanson lab graduate students, Rahul Kanadia, Yuan Yuan, Jason O'Rourke, Jihae Shin, and Ranjan Batra have always made laboratory an intellectually stimulating and fun experience. I would also like to thank Joyce Conners for her continuous support, help with the administrative aspects of my graduate career, and generally making my life as easy as possible.

TABLE OF CONTENTS

page

LIST OF FIGURES

LIST OF ABBREVIATIONS

DM	myotonic dystrophy
CDM	congenital myotonic dystrophy
HD	Huntington Disease
SMA	spinal muscular atrophy
SCA	spinal cerebellar ataxia
UTR	untranslated region
Kb	kilobase
HP1γ	heterochromatin protein 1 γ
HSA^{LR}	human skeletal actin long repeats
PKR	protein kinase R
dsRNA	double stranded RNA
ssRNA	single stranded RNA
miRNA	microRNA
MBNL	muscleblind-like
AAV	adeno-associated virus
CNS	central nervous system
NPC	nuclear pore complex
Dscam	Down syndrome cell adhesion molecule
snRNP	small nuclear ribonucleoprotein particle
SF1	splicing factor 1
U2AF65	U2 auxiliary factor 65
SR	serine/arginine rich protein
hnRNP	heterogenous nuclear ribonucleoprotein
RRM	RNA recognition motif

ESE/ESS	exonic splicing enhancer/silencer
ISE/ISS	intronic splicing enhancer/silencer
SMN	survival of motor neuron protein
IR	insulin receptor
cTNT/TNNT2	cardiac troponin T
TNNT3	fast skeletal muscle troponin T
CLCN1	muscle specific chloride channel
SCN4A	voltage gated sodium channel subunit A
ClaLC	clathrin light chain
Mbl	muscleblind
PTB	polypyrimidine tract binding protein
E15	embryonic day 15
P1	postnatal day 1
CUGBP1	CUG binding protein 1
LDHA	lactate dehydrogenase A
Mhc	myosin heavy chain
GluI	glutamine synthetase
TA	tibialis anterior
H&E	hematoxylin and eosin
ESC	embryonic stem cell
MEF	mouse embryonic fibroblast
Neo^R	neomycin resistance gene
TK	Herpes Simplex Virus thymidine kinase gene

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MUSCLEBLIND IN DEVELOPMENT AND DISEASE

By

Michael Poulos

May 2010

Chair: Name Maurice Swanson
Major: Medical Sciences – Genetics

Myotonic dystrophy (DM) may be the most variable human disorder described, affecting every age group from newborns to the adult-onset form which can range from adolescent manifestations to late-onset disease. Frequent disease characteristics include skeletal muscle weakness, myotonia, cataracts, and distinct changes in alternative splicing patterns. Congenital patients present with additional symptoms at birth, consisting of immature muscle, pulmonary insufficiencies, and central nervous system (CNS) involvement. DM arises from the expansion of two similar non-coding microsatellites in the *DMPK* and *CNBP* genes which have been proposed to cause disease through a common mechanism, a toxic RNA gain-of-function which can either inhibit or activate specific proteins. One of these candidates, the muscleblind-like (MBNL) family of proteins encoded by three genes, *MBNL1*, *MBNL2*, and *MBNL3*, are sequestered into discrete nuclear foci by RNA repeat expansions, preventing interactions with endogenous RNA targets and compromising their activity. However, it is unclear how the inhibition of MBNL function leads to disease and the extent of MBNL involvement in the diverse presentation of symptoms in DM. The goal of this study was to investigate the normal function of MBNL proteins which are compromised in DM and

their role in congenital disease. Our working hypothesis is that MBNL genes show distinct temporal and spatial expression patterns that influence age-of-onset and disease-associated pathological changes.

For MBNL1, we first demonstrate that this protein is an alternative splicing factor that interacts with pre-mRNAs misspliced in DM and that this interaction is necessary for splicing responsiveness. Additionally, we show that MBNL1 interacts with DM1 pathogenic and non-pathogenic repeat RNAs, but inherent differences in these interactions contribute to the ability to promote disease associated changes in alternative splicing. This and additional evidence supports the hypothesis that loss of MBNL1 function in DM causes defects in the alternative splicing of specific genes during postnatal development which leads to distinct pathological features in adult-onset disease, including myotonia and insulin insensitivity.

Since MBNL3 shows a restricted expression pattern during development and regeneration, we employed an *in vivo* approach to test the hypothesis that *Mbnl3* is necessary for normal embryonic development and loss of this protein results in congenital myotonic dystrophy (CDM). *Mbnl3* expression patterns, particularly in tissues affected in CDM including skeletal muscle and lung, as well as the cellular localization of Mbnl3 protein isoforms, suggest a function that is distinct from other Mbnl family members. Indeed, siRNA-mediated knockdown analysis of Mbnl3 in the C2C12 cell culture model of myogenesis demonstrates that Mbnl3 is essential for normal myogenic differentiation. To test the hypothesis that Mbnl3 was required for normal myogenesis *in vivo*, we generated *Mbnl3* isoform knockout mice that fail to express the major Mbnl3 isoform. While these $Mbnl3^{\Delta E2/Y}$ knockout mice do not recapitulate the

CDM phenotype, upregulation of other *Mbnl3* isoforms occurs in this line suggesting the possibility of functional complementation. A possible role for *Mbnl3* expression in adult muscle regeneration following injury was also discovered. We conclude that *Mbnl3* null lines must be created to determine if expression of this *Mbnl* gene is essential for normal embryonic muscle development. Overall, these results demonstrate that MBNL1 is an alternative splicing factor that regulates gene expression during postnatal life while *MBNL3* expression is essential for normal myogenic differentiation *in vitro* and possibly *in vivo*.

CHAPTER 1
OVERVIEW OF MYOTONIC DYSTROPHY

Myotonic Dystrophy Type 1

Myotonic dystrophy type 1 (DM1) is a dominantly inherited, multisystemic, neuromuscular disorder. It is the most common adult-onset form of muscular dystrophy, affecting 1 in 8000 individuals worldwide (Harper, 2001). Common DM1 symptoms include subcapsular cataracts, progressive distal skeletal muscle weakness and wasting, insulin resistance, cardiac arrhythmia, cognitive impairment, myotonia (a failure to relax skeletal muscle following a voluntary contration) (Avaria and Patterson, 1994; Groh et al., 2002; Lacomis et al., 1994; Milner-Brown and Miller, 1990; Stuart et al., 1983), and a molecular defect in the alternative splicing of specific pre-mRNAs (Ranum and Cooper, 2006). An interesting hallmark of this disease is the remarkable variability seen in the age of onset and penetrance of these clinical phenotypes. DM1 symptoms have been reported from adolescence to adulthood with early onset cataracts, muscle weakness, and myotonia being the most frequent disease manifestation (Harper, 2001). DM1 also displays genetic anticipation, a phenomenon in which each successive generation afflicted with disease presents with an earlier age of onset accompanied by more severe symptoms. Additionally, at its earliest time of onset, infants are affected with a distinct and severe form of disease (see congenital myotonic dystrophy). The dynamic nature of the characteristics of this disease have historically presented problems for defining the disorder in terms of symptoms. However, the variability demonstrated in DM1 became more clear in 1992, when the molecular defect was uncovered (Brook et al., 1992; Buxton et al., 1992; Fu et al., 1992; Harley et al., 1992; Mahadevan et al., 1992).

DM1 is caused by a $(CTG)_n$ microsatellite repeat expansion in the 3' untranslated region (UTR) of the *DMPK* gene. In the normal population, CTG repeats can range from 5 to 37. In disease, repeats expand to $(CTG)_{50\text{-}1000}$ for adult onset patients and $(CTG)_{1000\text{-}4000}$ for congenital patients (Anvret et al., 1993; Botta et al., 2008; Hedberg et al., 1999) (Fig 1-1A). Individuals with the initial expansion, called the protomutation, generally have smaller repeats and present with more mild symptoms and a later age of onset. Further transmission of the expanded allele results in a longer repeat region and more severe disease in the following generations (Harper, 2001). Interestingly, genetic anticipation, is also seen in other dominantly inherited trinucleotide repeat disorders which are subject to expansions, including Huntington disease (HD) and several spinal cerebellar ataxias (SCA) (Koshy and Zoghbi, 1997; La Spada et al., 1994). $(CTG)_{>50}$ repeats also display somatic instability that result in heterogenous expansions in tissues, likely contributing to the high degree of phenotypic variability seen in DM1. Although $(CTG)_n$ size can differ within the same individual due to somatic mosaicism, disease severity correlates positively with the length of the CTG expansion itself (Ashizawa et al., 1993; Martorell et al., 1997; Wong et al., 1995).

While mechanism of expansion of CTG/CAG repeats are not well understood, multiple models have been proposed to explain this process and have focused on errors during DNA replication and repair. During leading strand replication, CTG/CAG repeats can adopt stable DNA secondary structures that disassociate the 3' end of the synthesizing strand from the parental strand and allow it to reanneal out of register within the repeats of the parental strand (Pearson et al., 2005) (Fig. 1-2). Alternatively, Okazaki fragments located entirely within the expanded repeat itself during lagging

strand synthesis, which are not anchored to non-repetitive sequence, can slip and produce unpaired repeats (Richards and Sutherland, 1994). Repair and further rounds of replication promote expansion. It is, however, unclear what causes the protomutation from $(CTG)_{5-37}$ to $(CTG)_{>37}$, making the repeats susceptible to further expansion. Transgenic mice expressing *DMPK*-$(CTG)_{55}$ demonstrate intergenerational expansions in repeat size, while *DMPK*-$(CTG)_{20}$ remain static (Seznec et al., 2000). This result substantiates the observation in humans that CTG repeats become unstable between $(CTG)_{37-50}$ and are subject to further expansion in the germline, but it fails to explain the underlying cause of the protomutation.

Congenital Myotonic Dystrophy

Congenital myotonic dystrophy (CDM) is caused by $(CTG)_{1000-4000}$ repeat expansions (Tsilfidis et al., 1992). This is the same CTG mutation in DM1, however, the larger expansion affects newborns with a distinct set of clinical features. Individuals born with CDM present with immature skeletal muscle, poor suckling, talipes (club foot), mental retardation, pulmonary involvement, and hyotonia (a lack of baseline muscle tone resulting in movement impairment) (Harper, 1975; Harper, 2001). The main cause of mortality in CDM patients is due to respiratory deficiencies, believed to result from hypotonia of the diaphragm and intercostal muscles, as well as lung involvement (Campbell et al., 2004; Rutherford et al., 1989). CDM children that overcome the early-onset symptoms generally present with delays in developmental milestones until 3-4 years of age (Echenne et al., 2008). Remarkably, delays in skeletal muscle development and pulmonary involvement is resolved in early childhood. However, most CDM patients develop adult-onset DM1 symptoms by adolescence.

Haploinsufficiency Model of Disease

Dominantly inherited microsatellite repeat expansion diseases, including HD and SCA1, have CAG repeats located within their coding regions. These repeats code for a polyglutamine expansion, resulting in a toxic gain-of-function that affects multiple cellular processes (Imarisio et al., 2008; Kang and Hong, 2009). However, DM1-associated $(CTG)_n$ expansions in the *DMPK* 3'UTR are also inherited in a dominant manner, but do not affected the predicted protein product. This raises an intriguing question; how does a non-coding mutation cause a dominantly inherited disease? One potential explanation for this observation is the haploinsufficiency model, which proposes that the $(CTG)_n$ repeat expansion causes down-regulation of *DMPK* and the flanking genes, *DMWD* and *SIX5*. *DMWD*, *DMPK*, and *SIX5* are located within a 20 kilobase (kb) region of chromosome 19q13. Previous studies have reported evidence of decreased mRNA expression from these genes in DM1 and CDM tissues when compared to unaffected individuals (Eriksson et al., 2001; Fu et al., 1993; Hamshere et al., 1997; Inukai et al., 2000; Klesert et al., 1997; Novelli et al., 1993; Sabouri et al., 1993). $(CTG)_n$ repeats are flanked by CTCF insulator elements, which bind CTCF proteins and establish boundaries preventing the unwanted spread of heterochromatin at the *DMPK* locus (Phillips and Corces, 2009). In DM1, CTCF binding sites are methylated, preventing CTCF from binding (Filippova et al., 2001). In the absence of CTCF, antisense transcription from the downstream gene *SIX5* extends thorugh the CTG repeats and promotes the spread of heterchormatin formation by recruiting heterochromatin protein 1 γ (HP1γ), a protein involved in localization of silenced genes to the nuclear lamin (Cho et al., 2005). Additionally, loss of DNaseI hypersensitivity

sites at the *DMPK* locus supports the idea that heterochromatin formation could lead to loss of expression at the *DMPK* locus (Otten and Tapscott, 1995).

To test the idea that haploinsufficieny of the *DMPK* locus, caused by the expansion of $(CTG)_n$ repeats, promoted DM1 and CDM, *Dmpk* and *Six5* knockout mice were generated. *Dmpk*$^{-/-}$ mice display a mild late-onset skeletal muscle myopathy, characterized by a reduction in force generation, increased fibrosis, and a decrease in organization of the Z-line (Reddy et al., 1996). *Dmpk*$^{-/-}$ mice also demonstrate cardiac conduction defects, including first, second and third atrioventricular blocks (Berul et al., 1999). However, only a first degree atrioventricular block and no skeletal muscle defects are observed in *Dmpk*$^{+/-}$ mice. *Six5*$^{+/-}$ mice develop ocular cataracts, while *Six5*$^{-/-}$ males are sterile and display hypogonadism (Klesert et al., 2000; Sarkar et al., 2000; Sarkar et al., 2004). However, the age-related nuclear cataracts in *Six5*$^{+/-}$ mice are different than the subcapsular cataracts seen in DM1. *Six5*$^{+/-}$ mice also display mild cardiac conduction defects (Berul et al., 2000; Wakimoto et al., 2002). Neither of these models present with cardinal DM1 phenotypes, including skeletal muscle wasting, myotonia, subcapsular cataracts or changes in alternative splicing (Personius et al., 2005). In addition, skeletal muscle phenotypes and hypogonadism are only observed in *Dmpk*$^{-/-}$ and *Six5*$^{-/-}$ mice, not in *Dmpk*$^{+/-}$ and *Six5*$^{+/-}$ mice as the haploinsufficieny model would predict. While loss of DM1 associated gene expression may contribute to the disease, it is clear that haploinsufficiency of these genes is not solely responsible for disease onset. To date, no one has generated a *Dmwd* knockout mouse to assay for its contribution to the haploinsufficieny model.

Myotonic Dystrophy Type 2

The symptoms of DM1 have been well documented in the literature. However, a small percentage of families with myotonic dystrophy symptoms fail to develop genetic anticipation or a congenital form of the disease. Additionally, these individuals display proximal skeletal muscle weakness (as opposed to distal weakness in DM1) and normal $(CTG)_{5\text{-}37}$ repeats in the *DMPK* 3'UTR (Ricker et al., 1995; Thornton et al., 1994). In 2001, a $(CCTG)_n$ mutation linked to these families was identified in the first intron of the *CNBP* gene (formerly known as *ZNF9*), so the disease was renamed myotonic dystrophy type 2 (DM2) (Liquori et al., 2001). In the normal population, unaffected individuals posses $(CCTG)_{>27}$, while the DM2 pathogenic expansions contain $(CCTG)_{75\text{-}11,000}$ (Fig. 1-1B). These patients comprise approximately 5% of total myotonic dystrophy cases. Interestingly, DM1 and DM2 share a significant overlap in adult-onset disease manifestation (Fig. 1-3) despite being caused by two independent mutations in seemingly unrelated genes. *Cnbp*$^{+/-}$ mice have a reduced expression of *Cnbp* mRNA and show skeletal muscle degeneration and enlarged hearts (Chen et al., 2007). However, unlike DM1, there is no evidence that DM2 expansions alter expression from the locus harboring the $(CCTG)_n$ expansion (Botta et al., 2006). Therefore, haploinsufficiency of *CNBP* is unlikely to be a cause of DM2. Taken together, these observations predict a common primary mechanism of pathogenesis in DM1 and DM2, that is disconnected from the specific disease associated loci.

RNA Dominance and Sequestration Models of Disease

How do two different non-coding $C(CTG)_n$ repeat expansions in unrelated genes result in a dominantly inherited disease with remarkably similar phenotypes? The first evidence of a common mechanism emerged from the observation that both $(CUG)_{>50}$

and $(CCUG)_{>75}$ RNAs localize to discrete nuclear foci in DM1/DM2 skeletal muscle sections and cell culture but not in normal controls (Davis et al., 1997; Margolis et al., 2006). Relocalization into nuclear foci inhibits *DMPK* mRNA from nuclear export and translation (Davis et al., 1997). Additionally, a transgenic mouse expressing $(CTG)_{250}$ in the 3'UTR of a human skeletal actin transgene (HSA^{LR}) developed nuclear RNA foci similar to those seen in DM1 and DM2 and phenocopied characteristic DM skeletal muscle symptoms, including myotonia, centralized nuclei (indicating damaged and regenerating muscle fibers), and alternative splicing defects (Mankodi et al., 2000). HSA^{SR} control mice expressing $(CTG)_5$ were unaffected. Interestingly, the severity of the DM phenotype in mice correlated with expression of the transgene. Highly expressing HSA^{LR} mice were more affected than moderately expressing mice, while transgenes inserted into heterochromatic regions that did not express detectable levels of $(CUG)_{250}$ RNA did not develop DM symptoms. The presence of three unrelated and non-coding $C(CUG)_n$ repeat expansions that aberrantly localize in nuclear foci and promote expression dependant DM phenotypes suggest a toxic gain-of-function at the RNA level.

One model proposes that toxic $C(CUG)_n$ RNAs activate cellular proteins which, in turn, can contribute to the onset of disease through unwanted downstream events. First, protein kinase R (PKR) is a double-stranded RNA (dsRNA) binding protein that functions in innate immunity by interacting with viral dsRNAs in infected cells (Sadler and Williams, 2007). Activation of PKR through binding of dsRNA promotes cell death by inhibiting cellular translation, via phosphorylation of eIF-2, and activation of other cellular stress response mechanisms, including RNase L. PKR binds toxic $(CUG)_n$ RNA

in a length-dependent manner with increasing affinity *in vitro* and activates PKR in cell culture (Tian et al., 2000). However, PKR is not activated in DM1 patient tissues and *Pkr*$^{-/-}$; *HSA*^{LR} mice develop DM phenotypes, demonstrating that Pkr is not necessary for onset of disease in the presence of expanded $(CUG)_{250}$ (Mankodi et al., 2003). Second, expanded $(CUG)_n$ repeats can also act as substrates for DICER, a dsRNA binding protein with endonuclease activity that cleaves pre-microRNAs into 22 nucleotide fragments in the microRNA (miRNA) pathway (Krol et al., 2007). Processing of $(CUG)_n$ repeats by DICER can promote the downregulation of CAG containing transcripts through the siRNA pathway *in vitro*. Nevertheless, CUG siRNAs have not been detected in DM1 tissues or in a mouse model expressing $(CUG)_{250}$ repeats in skeletal muscle (Osborne et al., 2009). No additional evidence has substantiated CUG siRNA involvement in DM1 pathogenesis. Finally, CUGBP1, a single-stranded RNA (ssRNA) binding protein involved in developmentally regulated alternative splicing and RNA turnover, is upregulated ~ 2-fold in DM1 and DM2 skeletal muscle and heart when compared to the normal tissues (Timchenko et al., 2001). CUGBP1 upregulation is caused by increased stability of the protein through PKC-mediated hyperphosphorylation (Kuyumcu-Martinez et al., 2007). Mice expressing a tamoxifen-inducible transgene with $(CUG)_{960}$ repeats in the *DMPK* 3'UTR produce increased levels of CUGBP1, promote skeletal muscle wasting and heart conduction abnormalities, and die within 3 weeks of induction (Orengo et al., 2008). While these mice exhibit skeletal muscle and heart phenotypes, they do not recapitulate an adult-onset progressive myopathy seen in DM patients. It is unclear how toxic $(CUG)_n$ repeats activate PKC and promote CUGBP1 stability. One possibility is that $C(CUG)_n$ repeats play a direct role in

activating PKC through an unknown mechanism. Alternatively, PKC activation could be a byproduct of acute damage to the heart and skeletal muscle by expression of a toxic molecule *in vivo*.

The sequestration model proposes that toxic $C(CUG)_n$ RNAs inhibit cellular proteins by presenting novel binding sites which can inhibit their normal function through through dominant-negative interactions. Interestingly, $(CUG)_n$ and $(CCUG)_n$ repeat expansions are predicted to form similar thermodynamically stable RNA hairpin structures which have been visualized by electron microscopy (Michalowski et al., 1999). As the repeats expand, $(CUG)_n$ and $(CCUG)_n$ dsRNA would potentially become more toxic by sequestering an increasing amount of cellular protein and leading to more severe disease. To support this hypothesis, an *in vitro* binding assay using increasing lengths of $(CUG)_n$ RNA demonstrated a length-dependant interaction with a dsRNA binding protein, muscleblind-like 1 (MBNL1) (Miller et al., 2000). Human *MBNL1* is homologous to the *muscleblind* gene that is responsible for terminal muscle and photoreceptor development in *Drosophila* (Pascual et al., 2006). Moreover, MBNL1 protein colocalizes with pathogenic $(CUG)_n$ and $(CCUG)_n$ in discrete nuclear foci in DM1 and DM2 tissues. If inhibition of MBNL1 function by a dominant-negative interaction with toxic $(CUG)_n$ and $(CCUG)_n$ repeat RNAs is responsible for disease (Fig. 1-4), then loss of Mbnl1 in a mouse model should faithfully recapitulate DM phenotypes. Remarkably, *Mbnl1*$^{\Delta E3/\Delta E3}$ mice, which do not express Mbnl1 protein, phenocopy subcapsular cataracts, electrical myotonia, characteristic skeletal muscle pathology, and alternative splicing defects seen in DM (Kanadia et al., 2003a). Therefore, loss of Mbnl1 alone in a mouse model, is sufficient to cause multiple DM adult-onset

phenotypes. In an *in vivo* complementation assay, adeno-associated virus (AAV) mediated overexpression of myc-tagged Mbnl1 in the skeletal muscle of the HSA^{LR} mouse model, expressing $(CUG)_{250}$, rescued the myotonia and alternative splicing defects (Kanadia et al., 2006). This result provides further evidence that MBNL1 loss of function is the primary mechanism of adult-onset disease. While this data supports the idea that loss of MBNL1 function via sequestration by toxic $C(CUG)_n$ repeats causes multiple cardinal disease phenotypes in DM1 and DM2, there are many questions that have yet to be addressed.

Multiple Genetic Contributions and the Complex Etiology of DM

Myotonic dystrophy may be the most variable human genetic disorder in terms of its clinical presentation. Variably penetrant symptoms can occur from embryogenesis to late adult-onset with a wide variety of tissues affected. Germline expansion and somatic mosaicism of toxic $C(CUG)_n$ repeats likely contribute to the unique nature of this disease. Current data provides considerable evidence that MBNL1 function is compromised in DM and that this loss-of-function is the primary cause of many of the most characteristic symptoms of the adult-onset disease. However, this and other models fail to recapitulate key features of disease, including CDM, progressive skeletal muscle wasting, and central nervous system (CNS) phenotypes. $C(CUG)_n$ expansions potentially affect many genes by altering local chromatin structure at the DNA level and acting in a dominant-negative manner at the RNA level, activating and inhibiting multiple pathways. Like the MBNL1 loss-of-function model, disruption of a single gene by repeat expansions may be sufficient to recapitulated unmodeled DM phenotypes, including CDM and progressive skeletal muscle wasting. Alternatively, combinatorial genetic models may be necessary to fully reconstruct the onset of disease. The potential

contribution of multiple molecular pathways disrupted in DM makes determining the specific genetic contributions to disease a daunting task.

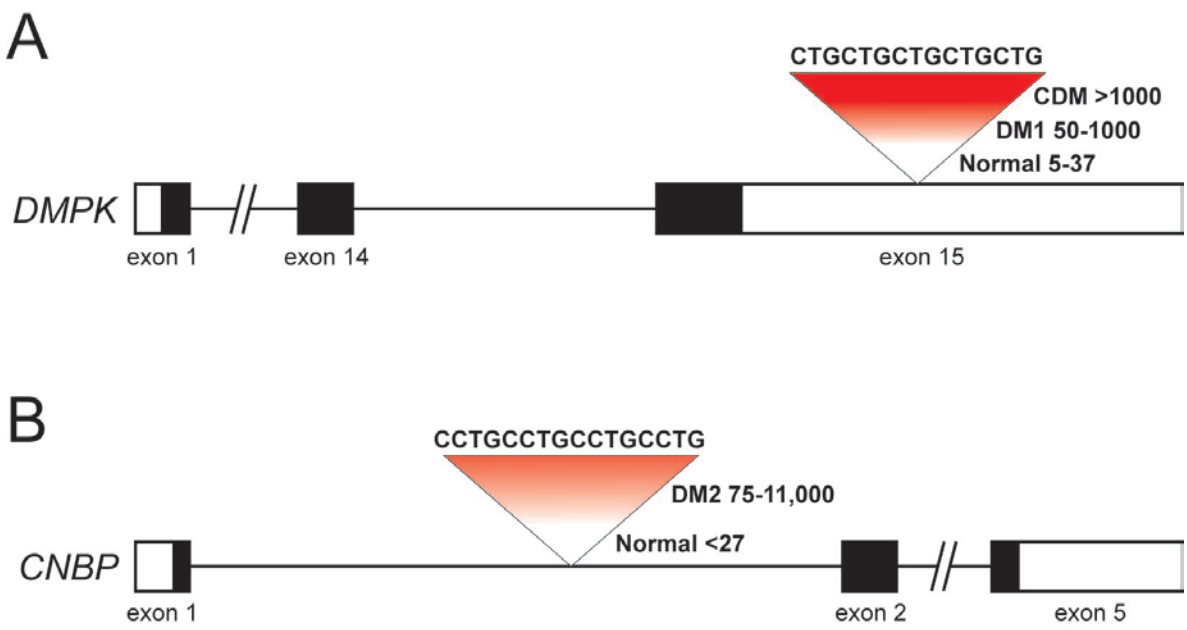

Figure 1-1. Non-coding microsatellite repeat expansions in myotonic dystrophy. (A) The *DMPK* gene with a $(CTG)_n$ expansion in the 3' UTR. *DMPK* exons (boxes) and introns (horizontal line) showing the position of $(CTG)_n$ repeats in the 3' UTR and expansion size correlating with disease (coding exons = black box; UTR = open box). (B) The *CNBP* gene with a $(CCTG)_n$ expansion. *CNBP* exons (boxes) and introns (horizontal line) showing the position of $(CCTG)_n$ repeats in the first intron and expansion size correlating with disease (coding exons = black box; UTR = open box).

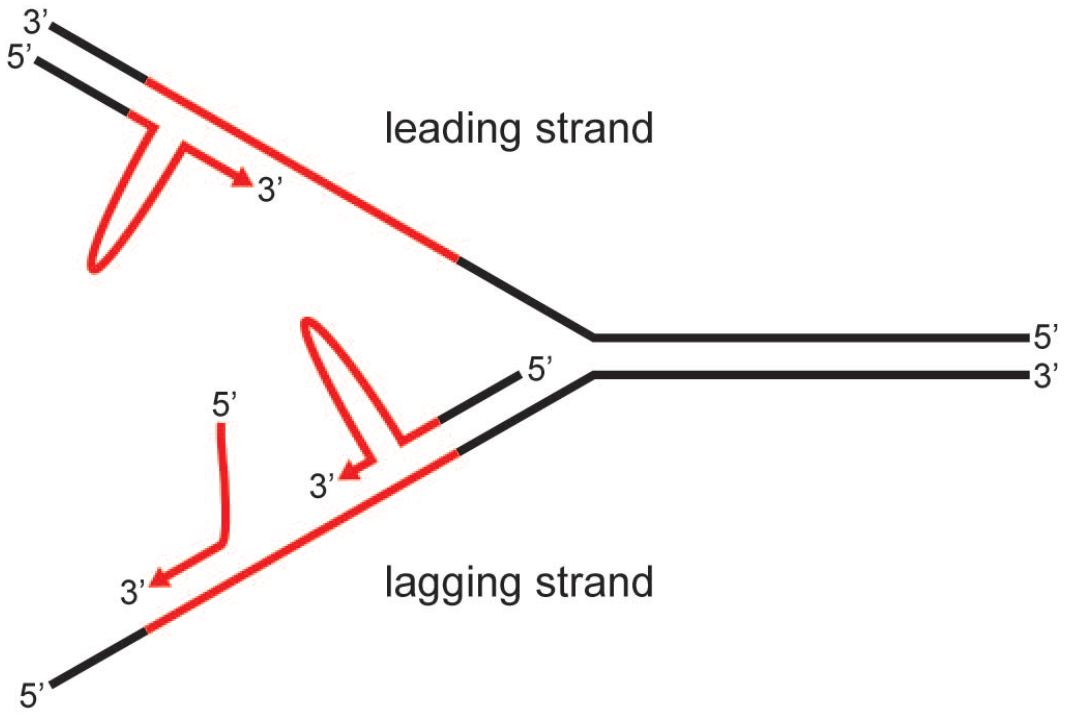

Figure 1-2. Genomic DNA replication errors promote microsatellite repeat expansions and contractions. $(CTG)_n$ DNA expansions (red) form stable hairpin structures on the replicating strand by intramolecular base pairing and promotes slippage on both the leading and lagging strand. Okazaki fragments located entirely within the CTG repeats on the lagging strand can shift out of register, leaving an unpaired 5’ end. Repair and replication promote further expansion.

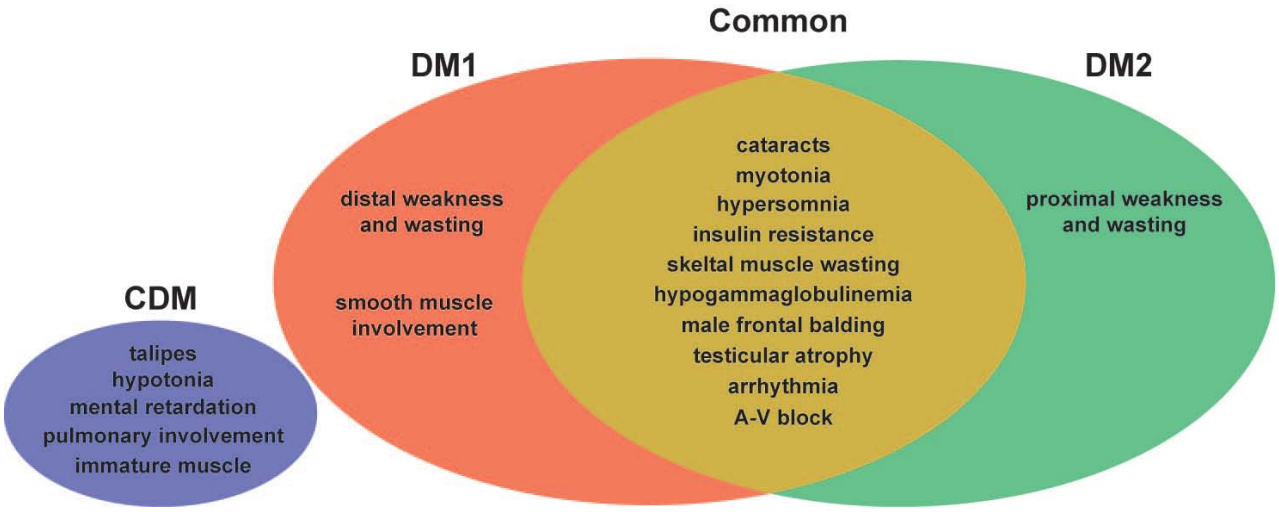

Figure 1-3. Comparison of myotonic dystrophy symptoms. DM1 (red) and DM2 (green) share a unique and overlapping (dark yellow) clinical presentation. CDM (purple) displays distinct neonatal characteristics.

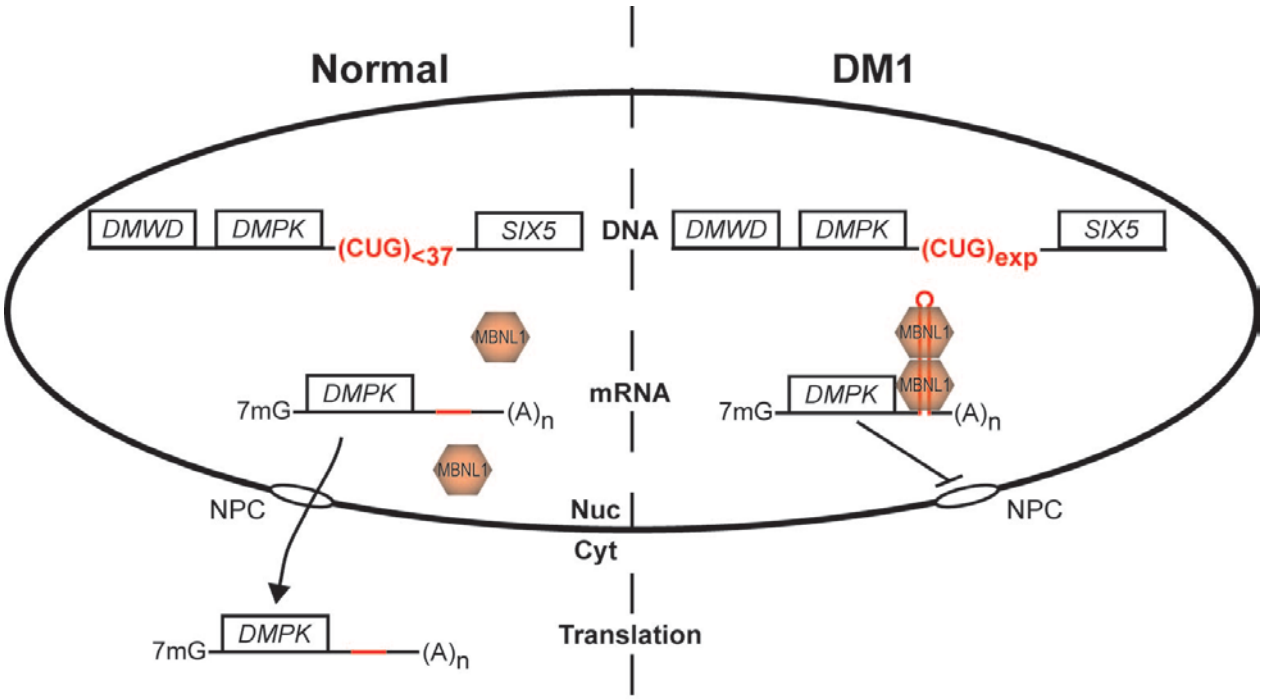

Figure 1-4. MBNL1 sequestration model for DM1. *DMPK* expression from the DM1 locus (including tightly linked genes *DMWD* and *SIX5*). Normal *DMPK* transcripts with $(CTG)_{>37}$ are processed and exported through the nuclear pore complex (NPC) and translated in the cytoplasm. MBNL1 steady-state levels remain diffusely nuclear. DM1 *DMPK* transcripts with expanded repeats, $(CTG)_{exp}$, are processed, blocked from transport, and sequestered into ribonucleoprotein (RNP) foci containing MBNL1. MBNL1 relocalization inhibits its normal function.

CHAPTER 2
MUSCLEBLIND PROTEINS REGULATE ALTERNATIVE SPLICING

Introduction

Introduction to Alternative Splicing

Alternative splicing is a remarkable process that allows an organism to extend its protein diversity from a limited gene pool by selectively including or excluding information in mature mRNA, which enables fine-tuning of protein function by expressing isoforms that are adapted for specific physiological requirements (Jin et al., 2008; Licatalosi and Darnell, 2010). In one instance, the Down syndrome cell adhesion molecule (*Dscam*) gene can potentially produce 38,016 protein isoforms (Schmucker et al., 2000). Unique *Dscam* isoforms expressed in each neuron promote repulsion of developing neurites from the same neuron by recognizing like isoforms. It has been estimated that a minimum of 4,752 isoforms are necessary to promote proper development of the Drosophila brain (Fuerst et al., 2009; Hattori et al., 2009; Matthews et al., 2007). Recent experimental approaches using deep sequencing of the entire human transcriptome reveals that nearly every gene is subjected to at least one alternative splicing event, highlighting the remarkable diversity that exists within the proteome itself (Pan et al., 2008; Pan et al., 2009; Wang et al., 2008). There are multiple patterns of alternative splicing that are capable of producing distinct mRNAs, including alternative promoters and untranslated regions (UTRs), mutually exclusive and cassette exons, retained introns, and cryptic 5' and 3' splice sites (Fig. 2-1). Each of these alternative splicing decisions is subject to multiple levels of regulation that are important for correctly identifying short exonic sequences within large pre-mRNAs that ultimately govern inclusion or exclusion in the mature transcript.

The bulk of pre-mRNA splicing is catalyzed by the major (or U2 dependant) spliceosome, an RNA:protein complex composed of five small nuclear ribonucleoprotein particles (snRNPs) U1, U2, U4/U6/U5, and a variety of auxiliary proteins (Black, 2003; Chen and Manley, 2009; Wahl et al., 2009). Although the spliceosome requires ~145 dynamically interacting proteins (Rappsilber et al., 2002; Zhou et al., 2002) important for functions including the coupling of splicing to transcription and the complex rearrangements of the core components, we will focus on the basal splicing machinery that is paramount for identifying prospective exons and completing the splicing reaction. Before splicing of pre-mRNA can occur, the prospective exons/introns must first be defined within the primary transcript. Constitutively spliced exons contain three consensus sequences that promote efficient recognition and recruitment of the spliceosome. Initially, U1 binds the canonical 5' splice site (CAG/guragu; uppercase = exon, lowercase = intron, r = any purine) through RNA-RNA base pairing. Next, splicing factor 1 (SF1) binds to the branch point (YNCURAY; Y = any pyrimidine, R = any purine, N = any nucleotide), followed by U2 auxiliary factor 65 and 35 (U2AF65 and U2AF35) recruitment to the consensus polypyrimidine tract and the 3' splice site ($u_3y_4$unyac/G; lowercase = intron, uppercase = exon, y = any pyrimidine, n = any nucleotide), forming the E-complex (Fig. 2-2). U2 interaction with the branch point through RNA-RNA base pairing displaces SF1, forming the A complex, a pre-mRNA in which exons and introns have been sufficiently defined to procede with assembly of the basal splicing machinery. Further, introduction of the U4/U6/U5 snRNP, replacing U2AF65 and U2AF35, establishes the B complex, which contains the necessary components to complete the splicing reaction. Extensive ATP dependant RNA-RNA

and protein-protein rearrangement of the B complex leads to the formation of the C complex. In the catalytic C complex, the 5' splice site undergoes a nucleophilic attack from the juxtaposed 2' OH of the branch point adenosine, forming a 5'-2' phosphodiester bond within the intron (forming the intron lariat). The 3' splice site subsequently undergoes a second nucleophilic attack from the newly formed 2'OH of the 5' splice site, creating a 3'-5' phosphodiester bond between the two spliced exons. Following the completion of splicing, the spliceosomal components are recycled and the 5'-2' intron lariat is debranched and degraded.

Alternatively spliced exons are included in mature mRNAs using the same mechanism as described for constitutive exons, however, differences in the three defining cis-elements used to identify these alternative exons are generally divergent from the canonical sequence. These non-consensus sequences inefficiently recruit the spliceosome and thus require additional protein co-factors to aid in the identification and definition of alternative exons. This subtle difference between constitutive and alternative exons allows for a layer of regulation that can be evolutionarily adjusted to promote tissue and developmental specific patterns of splicing by expressing proteins that assist in identifying or preventing recognition by the spliceosome. Two major families, the serine/arginine (SR) and heterogeneous ribonucleoproteins (hnRNP) proteins, are intimately involved in these alternative splicing decisions. The SR family of proteins contain multiple copies of an N-terminal RNA recognition motif (RRM) which generally recognize ssRNA motifs and C-terminal arginine/serine residues (RS) that often promote protein-protein interactions (Long and Caceres, 2009; Shepard and Hertel, 2009). SR proteins recognize exonic and intronic splicing enhancer (ESE and

ISE) elements and assist in the definition and recruitment of the basal splicing machinery to promote inclusion of alternatively spliced exons (Fig. 2-2). On the other hand, hnRNP proteins have been demonstrated to bind exonic and intronic splicing silencers (ESS and ISS) and inhibit interactions between alternatively spliced exons and the splicing machinery to promote exclusion of exons (Fig. 2-2) (He and Smith, 2009). Spatial and temporally restricted expression of these proteins can define a splicing environment, allowing cells to regulate the alternative splicing of genes involved similar pathways and functions (Chen and Manley, 2009). In contrast, misregulation of these factors can affect multiple downstream splicing decisions in trans, resulting in a widespread misregulation of alternative splicing and disease (Jensen et al., 2009; Ward and Cooper, 2010).

Alternative Splicing Misregulation in DM

The major molecular defect associated with adult-onset DM is the misregulation of alternative splicing for a specific subset of pre-mRNAs (Orengo and Cooper, 2007). The absence of a global misregulation in alternative splicing suggests that the defective component is independent of the basal splicing machinery, unlike the neuromuscular disorder Spinal Muscular Atrophy (SMA), which is caused by mutations in the survival of motor neuron (SMN) protein (Ogino and Wilson, 2004). SMN plays an integral part in snRNP biogenesis and mutations produce a global affect in splicing, manifesting primarily in motor neurons (Chari et al., 2009). Previous studies have indentified common misspliced genes in DM1 and DM2, including the insulin receptor (*INSR*), cardiac troponin T (*TNNT2,* formerly known as *cTNT*), fast skeletal muscle troponin T (*TNNT3*), and the muscle specific chloride channel (*CLCN1*) (Charlet et al., 2002; Mankodi et al., 2002; Philips et al., 1998; Savkur et al., 2001; Savkur et al., 2004).

Interestingly, the predominant pattern of misregulation in DM is the retention of a fetal splicing pattern in adult tissues, indicating that DM symptoms may result from fetal isoform expression in adult tissues that fail to meet the necessary physiological requirements.

Myotonia, a cardinal characteristic of DM, results from the inability to relax skeletal muscle after a voluntary contraction. To initiate a muscle contraction, acetylcholine is released from the motor neuron, binding nicotinic acetylcholine receptors on the muscle side of the neuromuscular junction, and activating Na^+ channels (*SCN4A*). Na^+ influx into the muscle triggers an action potential, activating Ca^{++} channels and initiating contraction (Barchi, 1995). After the initial contraction, CLCN1, a skeletal muscle specific voltage-gated chloride channel, is activated, allowing an influx if Cl^- into the muscle (Accardi and Pusch, 2000). The influx of Cl^- returns the membrane to its resting potential, preventing further contractions. Mutations in the muscle specific Na^+ and Cl^- channels have been shown to cause myotonia in both human disease and animal models (Hudson et al., 1995; Jurkat-Rott et al., 2002; Planells-Cases and Jentsch, 2009). RT-PCR analysis of DM1, DM2, HSA^{LR}, and $Mbnl1^{\Delta E3/\Delta E3}$ mice adult skeletal muscle revealed missplicing of *CLCN1*, including the retention of intron 2 and intron 6, as well as inclusion of cryptic exons 7a and 8a (Charlet et al., 2002; Kanadia et al., 2003a; Mankodi et al., 2002). Inappropriate inclusion of these exons and introns insert premature termination codons (PTCs) in the mature mRNA and promotes truncated isoforms, turnover of the transcript through nonsense-mediated decay (NMD), and loss of CLCN1 protein in the muscle sarcolemma in DM, HSA^{LR}, and $Mbnl1^{\Delta E3/\Delta E3}$ mouse models (Berg et al., 2004). While misregulation of the chloride channel alone is

sufficient to promote myotonia, previous reports have also implicated reduced activity of the Na^+ channel SCN4A in DM1 myocytes and *Dmpk*$^{-/-}$ mice (Benders et al., 1993; Mounsey et al., 2000; Reddy et al., 2002). However, there have been no documented changes in the expression of *SCN4A* in DM patients (Kimura et al., 2000), indicating that changes in Na^+ conductance may be an indirect affect of DM myopathy. To test the hypothesis that misregulation of the Cl^- channel alone is sufficient to cause myotonia in DM1, antisense morpholinos designed against the *CLCN1* cryptic exon 7a 3' splice site (to inhibit exon 7a inclusion in the mature mRNA) were injected into the tibialis anterior (TA) muscle of the *HSA*^{LR} mouse model and assayed for recovery of myotonia (Wheeler et al., 2007). Injected *HSA*^{LR} mice displayed normal *Clcn1* splicing patterns, full length protein correctly localized the sarcolemma, and a reversal of myotonia. This data provides evidence that missplicing of *CLCN1* is the primary cause of myotonia in DM patients. Additionally, it lends support to the idea that missplicing events in DM directly promote disease phenotypes associated with disease.

Another gene affected in DM, the *INSR*, is a tyrosine kinase receptor composed of two α subunits and two β subunits that localize to the plasma membrane and regulate glucose metabolism (Kahn and White, 1988). Interaction with its primary ligand, insulin, through the extracellular α subunit causes autophosphorylation of the intracellular β subunits, promoting uptake of glucose, glycogen synthesis, and glycolysis. Alternative splicing of *INSR* exon 11 in the α subunit results in two protein isoforms, IR-A (exon 11 exclusion) and IR-B (exon 11 inclusion) (Moller et al., 1989; Seino and Bell, 1989), which regulate the receptors sensitivity to insulin. The IR-B isoform displays increased sensitivity to insulin and a higher signaling capacity (Kellerer et al., 1992; McClain,

1991; Vogt et al., 1991). Tissues involved in maintaining glucose homeostasis primarily express IR-B, including skeletal muscle, liver, and adipose tissue. In skeletal muscle of DM1 and DM2 patients, *INSR* is misspliced, aberrantly excluding exon 11 which results in expression of the less sensitive isoform IR-A. DM patients also display insulin insensitivity in skeletal muscle, suggesting a direct link between *INSR* missplicing and disease (Moxley et al., 1978; Moxley et al., 1984; Vialettes et al., 1986).

Genes that regulate the protein machinery involved in generating muscle contractions are also misspliced in DM. Cardiac and skeletal muscle contractions are achieved through the regulated interaction of myosin with actin. During skeletal muscle contractions, myosin binds actin in an “open state”, releasing ADP and inorganic phosphate which promotes a conformational change in myosin to the “closed state”, resulting in force generation through the shortening of the sarcomere (Geeves and Holmes, 1999). Subsequent ATP hydrolysis returns myosin to the open position for ensuing contractions. During muscle relaxation, tropomyosin proteins negatively regulate skeletal and cardiac muscle contractions by occupying myosin binding sites on actin (Gunning et al., 2008). An increase in cellular Ca^{++} inhibits tropomyosin/actin binding through the trimeric troponin complex, allowing myosin/actin interactions and contraction (Ohtsuki and Morimoto, 2008). The troponin complex consists of three genes, troponin C, troponin I, and troponin T, encoded by *TNNC2*, *TNNI2*, and *TNNT3* in fast skeletal muscle and *TNNC1*, *TNNI3*, and *TNNT2* (*cTNT*) in cardiac muscle. Following the initiation of muscle contraction, troponin C binds increased cellular Ca^{++}, promoting conformational changes in troponin T, which allosterically inhibits tropomyosin/actin binding and allows myosin/actin interaction. *cTNT* and *TNNT3* both

contain fetal exons (Townsend et al., 1994; Yuan et al., 2007) that are excluded postnatally during normal development, but are included in adult heart and skeletal muscle in DM. Although the role of *cTNT* and *TNNT3* missplicing in DM is not as clear as *CLCN1* and *INSR* in disease, mutations in *cTNT* are associated with cardiac hypertrophy and sudden death (Moolman et al., 1997). Additionally, *TNNT3* may be involved in overall skeletal muscle weakness in DM due to inefficient contractions. However, unlike *CLCN1*, in which exon 7a inclusion displays a direct involvement in myotonia, the missplicing of *cTNT* and *TNNT3* most likely contribute to myopathy in conjunction with other misplicing events.

In DM, sequestration and inhibition of an RNA binding protein, MBNL1, by toxic $C(CUG)_n$ repeat expansions result in characteristic missplicing of alternative exons that promote disease. MBNL1 was originally identified based on its ability to interact with double stranded $(CUG)_{exp}$ RNA *in vitro* and loss of Mbnl1 protein in a mouse model recapitulated alternative splicing defects in DM (Kanadia et al., 2003a; Miller et al., 2000). One possibility to explain these observations is that MBNL1 is an RNA binding protein involved in promoting normal adult splicing patterns by directly interacting with pre-mRNA substrates and regulating the inclusion/exclusion of alternatively spliced exons (Fig. 2-3). In DM, MBNL1 is sequestered away from its targets, resulting in the loss of adult specific regulation of alternatively spliced fetal exons and the expression of aberrant isoforms.

Evolutionarily Conserved *Muscleblind* is Important for Muscle Development

MBNL1 belongs to a family of proteins, muscleblind, conserved from *Caenorhabditis elegans* to humans that share a common RNA binding motif, containing three cysteine, and one histidine, residues (C_3H) which are important for the

coordination of zinc ions and RNA-protein interactions (Pascual et al., 2006). Muscleblind (*mbl*) was originally described in a screen designed to identify genes involved in photoreceptor differentiation and development in *Drosophila*. *Sev-svp2* transgenic flies expressing *seven-up*, a hormone receptor transcription factor responsible for photoreceptor R3/4 and R1/6 differentiation, under the sevenless promoter which directs expression of photoreceptor R7, develop a rough-eye phenotype due to the inability to correctly pattern photoreceptor sub-type (Begemann et al., 1995; Hiromi et al., 1993). *Sev-svp2* flies were crossed to a collection of flies with UAS containing P-elements and screened for gain-of-function modifiers of the rough eye phenotype. *Mbl* was identified as a dominant suppressor of the rough-eye phenotype, suggesting a required role for terminal photoreceptor differentiation (Begemann et al., 1997). Additionally, *mbl* null mutants display larval lethality and skeletal muscle phenotypes, including disorganized Z-bands and reduced extracellular matrix at muscle attachment sites (Artero et al., 1998). Conservation of muscleblind homologs may suggest that they share a conserved molecular function important for the proper development muscle. If this observation is true, muscleblind should regulate alternative splicing in during both *Drosophila* and mammalian development.

Results

The *Drosophila mbl* gene and its mammalian homolog, *MBNL1*, play a pivotal role in the normal development of skeletal muscle (Pascual et al., 2006). Inhibition of MBNL1 function via dominant-negative interactions with toxic RNA contributes to the adult-onset diseases DM1 and DM2, including myotonia and a defect in alternative splicing of a specific subset of pre-mRNAs. *Mbnl1*$^{\Delta E3/\Delta E3}$ mice, designed to recapitulate

loss of MBNL1 function in DM1 and DM2, faithfully phenocopy myotonia and alternative splicing defects (Kanadia et al., 2003a). However, it is not clear if MBNL1 is directly involved in alternative splicing defects. One possibility is that MBNL1 is an alternative splicing factor that directly regulates normal alternative splicing decisions during development by interacting with those pre-mRNAs misspliced in disease. In this scenario, loss of MBNL1 by sequestration would directly affect missplicing by titrating MBNL1 away from its normal pre-mRNA substrates. Alternatively, loss of MBNL1 could promote downstream events that, in turn, result in disease phenotypes and missplicing. Therefore, we sought to test the hypothesis that MBNL1 directly regulates alternative splicing of pre-mRNAs affected in DM1 and DM2.

Muscleblind Protein Family: Evolutionary Conservation of Structure and Function

First, we explored the idea that muscleblind function is evolutionarily conserved in *Drosophila* and can regulate muscle-specific alternative splicing. *Mbl* can produce four distinct isoforms (A-D) varying at their C-termini (Fig. 2-4), but the functional distinction between these isoforms is unclear. If Mbl isoforms are in fact alternative splicing factors, a reasonable assumption is that they are localized to the nucleus. However, while MBNL1 has been shown to be predominantly nuclear, a lack of antibodies specific for Mbl isoforms has prevented similar analysis for the Drosophila homolog. Therefore, GFP-tagged MblA, MblB, MblC, and MblD isoforms were expressed in COSM6 cells and assayed for cellular localization. MblB and MblC localized predominantly to the nucleus while MblA displayed a more cytoplasmic pattern (Fig. 2-5). MblD was diffuse throughout the cell, possibly due to degradation (Fig. 2-5 and Fig. 2-7C). Interestingly, MblA and MblC also localized to cytoplasmic foci when exogenously expressed in

COSM6 cells, which were later identified as stress granules (Fig. 2-5 and data not shown). Mbnl1 and Mbnl3 also localize to stress granules in COSM6 cells while Mbnl2 and MblB were not present in the cytoplasmic foci (data not shown). To further determine if *mbl* functional interactions are conserved, GFP-tagged Mbl isoforms were coexpressed with $(CUG)_{300}$ repeats in COSM6 cells and assayed for interactions with toxic repeat RNA. MblA, MblB, and MblC colocalized in discrete nuclear foci with repeat RNA (Fig. 2-6), similar to MBNL1, MBNL2, and MBNL3 (Fardaei et al., 2002), while MblD remained diffuse throughout the cell. These observations demonstrate that the interactions and relocalization with toxic RNA repeats is conserved within the *muslceblind* family.

Drosophila α*-actinin* is a gene that participates in the organization of the sarcomere and undergoes a developmental alternative splicing switch (Fyrberg et al., 1990; Roulier et al., 1992). To test the hypothesis that *mbl* regulates alternative splicing during development, we coexpressed an α*-actinin* minigene (Fig. 2-7A) with different GFP-tagged Mbl isoforms in COSM6 cells and assayed for alternative splicing. MblB and MblC promoted exon 7 exclusion and adult muscle splicing patterns, while MblA was less efficient. MblD had no effect (Fig. 2-7B). *GFP-Mbl* expression was monitored by immunoblot (Fig. 2-7C). This result provides evidence that the developmental regulation of alternative splicing in muscle by the *muscleblind* family and interaction with toxic $(CUG)_n$ repeats is conserved from Drosophila to humans.

MBNL1 Proteins Directly Regulate Alternative Splicing of Gene Transcripts Misregulated in Myotonic Dystrophy

Many developmentally regulated genes are misspliced in adult-onset DM (Orengo and Cooper, 2007). Loss of Mbnl1 protein in the *Mbnl1*$^{\Delta E3/\Delta E3}$ mouse, model

reproduces the cardinal DM phenotypes, including missplicing events (Kanadia et al., 2003a). It is, however, unclear if MBNL1 is directly or indirectly involved in the alternative splicing of these genes. We sought to test the idea that MBNL1 directly regulates *INSR* and *cTNT*, two pre-mRNAs misspliced in DM. *INSR* and *cTNT* minigenes containing the alternatively spliced exons affected in DM (Fig. 2-8) where cotransfected into COSM6 cells with GFP-MBNL or siRNA specific for *MBNL1* and assayed for changes in alternative splicing by RT-PCR. Exogenous expression of GFP-tagged MBNL1, MBNL2, and MBNL3 promoted exclusion of *cTNT* fetal exon 5 and inclusion of *INSR* exon 11, two pre-mRNAs misspliced in DM, in a minigene assay in cell culture (Fig. 2-9A,B). Alternative splicing of a neuronal specific exon in *ClaLC*, a gene not affected in DM, was not influenced by MBNL expression (Fig. 2-9C). *GFP-MBNL* transgene expression was monitored by immunoblot analysis. GFP alone did not alter splicing. Interestingly, all three MBNL family members promote similar splicing patterns *in vitro*. This observation likely reflects the high degree of sequence identity shared by the MBNL family (Pascual et al., 2006). If exogenous expression of *MBNL1* promotes a splicing change in *cTNT* and *INSR*, then depletion of endogenous MBNL1 should recapitulate the DM splicing pattern. Two *MBNL1* specific siRNAs, directed against the coding sequence were used to knock down *MBNL1* expression >80-90% in HeLa cells (Fig. 2-10A,B and data not shown). Loss of endogenous MBNL1, via siRNA mediated knockdown, promoted the retention of *cTNT* fetal exon 5 and inclusion of *INSR* exon 11 (Fig. 2-10C), reproducing the missplicing pattern in DM. *ClaLC* was unaffected. siRNA directed against *GFP* was used as a knock down control and had no

affect on splicing. These data verifies that MBNL1 specifically regulates alternative of exons misspliced in DM.

MBNL1 Interacts Directly with Cis-Elements in *cTNT* pre-mRNA

To determine if MBNL1 interacts directly with *cTNT* pre-mRNA to regulate its splicing, an *in vitro* binding assay was performed using the alternatively spliced *cTNT* fetal exon 5 minigene which contains the upstream and downstream introns. Briefly, ^{32}P-labeled full length and truncated RNAs were incubated with recombinant GST-MBNL1, crosslinked with UV light, and resolved RNA:protein complexes by SDS-PAGE. RNA:protein interactions were visualized by label transfer. GST-MBNL1 was bound to a 41 nucleotide region directly upstream of the alternatively spliced exon (Fig. 2-11). Scanning mutagenesis further refined the MBNL1 binding site to two 8 nucleotide regions -18 to -26 and -36 to -44 upstream of exon 5 (Fig. 2-11). Two dinucleotide substitutions that eliminated MBNL1 binding (data not shown), but minimized disruption of the intron/basal splicing machinery interactions, were used in subsequent experiments to assay for splicing responsiveness *in vitro* (Fig. 2-11).

We next sought to determine if loss of the MBNL1 interacting cis-element in *cTNT* exon 5 inhibited MBNL1 responsiveness in a splicing assay. Coexpression of GFP tagged MBNL1, MBNL2, and MBNL3 and the wild type *cTNT* minigene promoted exon 5 exclusion in HeLa cells (Fig. 2-12). However, coexpression with the mutant *cTNT* minigene demonstrated a loss of responsiveness for MBNL overexpression (Fig. 2-12). *GFP-MBNL* transgene expression was monitored by immunoblot analysis and GFP alone did not alter splicing. Alignment of human *cTNT* and chicken *cTNT* MBNL1 binding sites (chicken *cTNT* experiments not shown) revealed a common YGCU(U/G)Y RNA motif for MBNL1 binding (Fig. 2-13). This data suggests that MBNL1 directly binds

to *cTNT* and *IR* pre-mRNAs and loss of this protein due to sequestration affects the alternative splicing of these genes.

$(CUG)_n$ and $(CAG)_n$ Repeats Relocalize MBNL1, but only $(CUG)_n$ Alter *cTNT* Splicing

In the RNA dominance model for DM, loss of MBNL1 function by toxic $C(CUG)_n$ RNA sequestration into nuclear foci results in disease. However, there are multiple trinucleotide repeat expansion diseases, the majority of which are coding region $(CAG)_n$ expansions, which fail to present DM-specific manifestations (Shao and Diamond, 2007). This observation led us to investigate the idea that MBNL1 sequestration is specifically dependent on $C(CUG)_n$ repeat RNAs. To address this question, a *DMPK* minigene construct with $(CTG)_{960}$, $(CAG)_{960}$, or 0 repeats (Fig. 2-14A) was transfected into COSM6 cells and assayed for repeat RNA and endogenous MBNL1 colocalization in nuclear foci. As expected, *DMPK* minigene expression with 0 repeats did not induce RNA foci and MBNL1 was localized diffusely throughout the nucleus (Fig. 2-14B). Suprisingly, MBNL1 colocalized with both $(CUG)_{960}$ and $(CAG)_{960}$ RNAs in nuclear foci (Fig. 2-14B). If $(CAG)_{960}$ repeats are capable of sequestering MBNL1, do they also alter pre-mRNA splicing of genes affected in DM? Increasing amounts of $(CUG)_{960}$, $(CAG)_{960}$, and control (0 repeats) plasmids were coexpressed with the *cTNT* minigene to assay for misregulation of alternative splicing. While $(CUG)_{960}$ promoted *cTNT* exon 5 inclusion at only 0.1 μg of transfected plasmid, increasing amounts of $(CAG)_{960}$ had only a minimal affect (Fig. 2-15). There were no statistical differences in RNA expression or in the number of foci produced from either repeat plasmid in cell culture (data not shown). Interestingly, MBNL1 relocalization into nuclear foci was not mutually exclusive with the misregulation of *cTNT* alternative splicing. This suggests that while

MBNL1 interacts with both $(CUG)_{960}$ and $(CAG)_{960}$ RNA, the nature of these interactions may be distinct.

The Stability of MBNL1:RNA complexes varies between $(CUG)_n$, $(CCUG)_n$ and $(CAG)_n$ Repeats

One possibility to explain the discrepancy between MBNL1 sequestration with $(CUG)_{960}$ and $(CAG)_{960}$ repeats and missregulation of alternative splicing is that both RNAs are capable of relocalizing nuclear MBNL1, but only $(CUG)_n$ repeats trap MBNL1 in foci resulting in their functional depletion. To test this hypothesis, ^{32}P-labeled $(CUG)_{54}$, $(CCUG)_{46}$, $(CAG)_{54}$, and T5.45 (an endogenous MBNL1 binding site from *TNNT3* intron 8,(Yuan et al., 2007)) RNAs were incubated with COSM6 whole cell lysate expressing myc-tagged MBNL1 in the presence of ATP, incubated for 30 minutes, UV crosslinked, RNase treated, and immunoprecipitated using an antibody specific for myc. Labeled RNAs were either challenged with 2000X fold excess of unlabeled RNA concurrently or 15 minutes following the initial labeled RNA:myc-MBNL1 incubation period. An unchallenged control was included to assay for the baseline level of RNA:protein crosslinking (Fig. 2-16A). Immunoprecipitated complexes were resolved by SDS-PAGE and autoradiographs of RNA:protein complexes were quantified using a phosphoimager. Myc-MBNL1 protein retained all of the $(CUG)_{54}$ RNA when challenged with 2000X fold excess unlabeled competitor, while $(CAG)_{54}$, $(CCUG)_{46}$, and T5.45 were less resistant to competition (Fig. 2-16B,C). MBNL1:$(CUG)_{54}$ interactions, once formed, effectively trapped MBNL1 protein and did not allow disassociation of the RNA:protein complex in 15 minutes. Interestingly, $(CCUG)_{46}$, a tetranucleotide repeat expansion that causes DM2, does not trap MBNL1 more effectively than non-pathogenic RNAs $(CAG)_{54}$ and T5.45. While all repeats tested are capable of relocalizing MBNL1 by interacting

with the protein, only $(CUG)_n$ repeats trap MBNL1 *in vitro*. Static MBNL1:$(CUG)_{54}$ interactions suggest that MBNL1 may form different complexes on $(CUG)_n$ repeats when compared to the more dynamic interactions with $(CCUG)_{46}$, $(CAG)_{54}$, and T5.45. This result predicts that $(CUG)_n$ repeat expansions are the most toxic RNAs in the sequestration model of myotonic dystrophy.

Discussion

Myotonic dystrophy is a neuromuscular disease caused by two microsatellite repeat expansions in unrelated genes that present with similar clinical manifestations. Related $C(CUG)_n$ toxic RNA expansions in DM1 and DM2 have been shown to compromise the function of MBNL1, an evolutionarily conserved protein family important for the terminal skeletal muscle development (Artero et al., 1998). Disruption of *Mbnl1* in a mouse model, mimicking a loss of function in disease, results in myopathy that phenocopies the primary symptoms of adult-onset DM and presents with characteristic molecular changes in alternative splicing (Kanadia et al., 2003a). Previous studies have demonstrated that the transcripts of numerous genes are misspliced in DM, including inappropriate exon inclusion in *cTNT* and exclusion in *INSR* (Philips et al., 1998; Savkur et al., 2001; Savkur et al., 2004). These misplicing events are particularly interesting because of their correlation to cardiomyopathy and endocrine abnormalities in DM (Belfiore et al., 2009; Fiset and Giles, 2008). However, it is unclear how loss of MBNL1 function contributes to alternative splicing defects in disease. In this study, we sought to test the hypothesis that muscleblind directly interacts with pre-mRNAs to regulate alternative splicing decisions and $(CUG)_n$ RNA expansions inhibit this function.

Muscleblind Proteins Directly Regulate Alternative Splicing

Drosophila mbl was originally identified in a genetic screen designed to identify modifiers of a rough eye phenotype. Interestingly, mutant flies that fail to express *mbl* develop severe embryonic muscle defects, including incorrect organization of the sarcomeric Z-line (Artero et al., 1998). The Z-line component, *α-actinin*, undergoes normal tissue and developmental alternative splicing transitions, including larval muscle, adult muscle, and non-muscle isoforms (Roulier et al., 1992). If the *muscleblind* family regulates developmental alternative splicing decisions, then it is reasonable to think that *mbl* promotes adult splicing patterns in Drosophila. Ectopic expression of Mbl isoforms A-D, which share the 63 N-terminal amino acids and $2XC_3H$ RNA binding motifs (Fig. 2-1), in COSM6 cells with an α-actinin minigene reporter (Fig. 2-4A) revealed that MblB and MblC preferentially promoted the adult muscle isoform (i.e. larval muscle exon 7 exclusion). Not surprisingly, Mbl isoforms B and C, which localized predominately to the nucleus, demonstrated more efficient exclusion of α-actinin larval exon 7 than MblA, which generally localized to the cytoplasm (Fig. 2-2). MblD, which contains only the N-terminal 84 amino acids, was diffusely distributed throughout the cell (Fig. 2-2 and 2-3) and was not stable when assayed for expression by immunoblot analysis (Fig. 2-4C). This observation may be a byproduct of expressing this protein in a mammalian cell. Interestingly, MBNL1 also shifts splicing of the α-actinin minigene in vitro, suggesting an overlap in function. Subsequent studies have also shown that *mbl* regulates adult specific alternative splicing of *Drosophila* troponin T, a homolog of *TNNT2* and *TNNT3*, genes known to be misspliced in DM. Additionally, Mbl isoforms A-C demonstrated an interaction with coexpressed $(CUG)_{300}$ RNA repeats in vitro, demonstrating an overlap

in RNA binding (Fig. 2-3) substrates with MBNL1. Moreover, transgenic flies expressing *MBNL1* also block the embryonic lethality associated with *mbl* mutant flies (Monferrer and Artero, 2006). This data provides evidence that *mbl* and *MBNL1* share a conserved function in promoting adult specific splicing patterns. However, these results do not make the distinction between a mechanism for Mbl promoting exclusion of alternative exons by directly interacting with the pre-mRNA or via an Mbl-mediated downstream event.

Recent studies using alternative splicing microarray analysis of HSA^{LR} and $Mbnl1^{\Delta E3/\Delta E3}$ mice reveal a >80% overlap in missplicing events between the two DM models, indicating that loss of MBNL1 function in disease is the primary cause of splicing defects. To test the idea that MBNL1 is responsible for regulating these splicing decisions by directly binding the pre-mRNAs affected in DM and promoting adult splicing patterns, we employed in vitro splicing and crosslinking assays. GFP-tagged MBNL1, MBNL2, and MBNL3 promote adult skeletal muscle splicing patterns in *cTNT* (exon 5 inclusion) and *INSR* (exon 11 exclusion) minigenes, while knockdown of endogenous MBNL1 in cell culture phenocopies aberrant DM splicing. Interestingly, all three MBNL family members promoted similar splicing patterns in vitro, indicating a potential overlap in function, which was also observed in the Mbl alternative splicing analysis. However, $Mbnl1^{\Delta E3/\Delta E3}$ mice and endogenous MBNL1 knockdowns reproduce DM splicing abnormalities, suggesting that other family members do not compensate this function *in vitro* or *in vivo*. The most likely explanation is that the MBNL family shares >90% amino acid identity while Mbl shares 100% identity in their respective C3H motifs, allowing dominant interactions with MBNL1 binding sites when the proteins are

exogenously overexpressed. More importantly, this data demonstrates that MBNL1 can regulate *cTNT* and *INSR* alternatively spliced exons in vitro, recapitulating disease splicing patterns and allowing for further analysis of the cis-elements responsible for mediating these decisions. Using UV crosslinking and mutagenesis of the *cTNT* alternatively spliced exon 5 and adjacent introns, we identified two MBNL1 binding sites immediately upstream of exon 5. Loss of either binding site through mutagenesis inhibited binding, suggesting that both are necessary for MBNL1 interaction with the pre-mRNA. In addition, two dinucleotide substitutions in MBNL1 binding sites repressed MBNL1 responsiveness in a *cTNT* minigene alternative splicing assay, demonstrating that inhibition of binding was sufficient to alleviate MBNL1 alternative splicing control of exon 5. Alignment of the human *cTNT* and conserved chicken *cTNT* MBNL1 binding sites revealed a common MBNL1 binding motif, YGCU(U/G)Y, that is responsive to MBNL1 *in vitro*. Other groups have verified additional MBNL1 binding sites using multiple approaches, including bioinformatics, systematic evolution of ligands by exponential enrichment (SELEX), and X-ray crystallography, which when compared with each other unmask a common YGCY sequence motif for MBNL1 binding (Du et al., 2010; Goers et al., 2010; Teplova and Patel, 2008). Interestingly, when Mbnl1 binding sites were mapped to mispliced exons in both HSA^{LR} and $Mbnl1^{\Delta E3/\Delta E3}$ mice, YGCY sequence motifs were enriched upstream of exons included in disease (normally excluded in adult tissues) and downstream of exons excluded in disease (normally included in adult tissues). This pattern is similar to alternative splicing factors Nova and Fox2, in which interaction with sequence motifs upstream of the alternatively spliced exon promotes exclusion and interaction with sequences downstream promote inclusion

(Licatalosi et al., 2008; Yeo et al., 2009). These observations suggest that these factors affect splicing decisions of alternative exons using a common mechanism. Although the MBNL1 sequence motif was not mapped to *INSR* pre-mRNA, this observation implies that the interaction would lie downstream of exon 11 and promote inclusion. How MBNL1 promotes exon inclusion is unclear. One possibility is that MBNL1 binds downstream of exon 11, altering the RNA structure of the pre-mRNA, and promoting favorable U1 snRNP binding at the exon 11 5' splice site. Alternatively, MBNL1 could compete with other negative regulators near the 5' splice site or could potentially assist in actively recruiting U1. These regulatory mechanisms are not mutually exclusive. The role of MBNL1 in exon inclusion may be a little more straight forward. MBNL1 interactions with *cTNT* exon 5 through the 3' slice site would potentially inhibit the binding of the basal splicing machinery, more specifically U2AF65/35 or SF1/U2 snRNA, and prevent the assembly of the spliceosome, masking the alternative exon from inclusion. In support of this, Berglund and coauthors demonstrated the MBNL1 binds a dsRNA structure directly upstream of *cTNT* exon 5 and prevents the ssRNA binding protein U2AF65 from interacting with the polypyrimidine tract, therefore inhibiting its ability to recruit U2 (Warf and Berglund, 2007). This mechanism is also likely responsible for promoting fetal exon exclusion of *Tnnt3*, in which MBNL1 binds a secondary structure directly upstream of the alternatively spliced exon (Yuan et al., 2007). This data provides evidence that MBNL1 is an alternative splicing factor responsible for modulating exon inclusion/exclusion by directly binding to sequences in the pre-mRNA. Loss of MBNL1 through sequestration would therefore remove this layer of regulation and promote missplicing of pre-mRNA targets affected in DM.

Microsatellite Repeat Expansions Display Variable MBNL1 Stability and Toxicity

Significant evidence supports a MBNL1 loss-of-function model underling DM pathogenesis, in which toxic RNAs that contain similar sequence motifs (YGCY) to normal MBNL1 substrates titrate the protein into insoluble complexes, inhibiting its activity. Colocalization of $(CUG)_{exp}$ RNA and MBNL1 protein in discrete nuclear foci has been observed in DM1 patient cells and tissues, *HSA*^{LR} mice, as well as cell culture in which repeat expansions are ectopically expressed (Kanadia et al., 2006; Mankodi et al., 2003; Mankodi et al., 2001). However, in an unexpected result, exogenously expressed $(CAG)_{960}$ repeats also relocalized endogenous MBNL1 in nuclear foci in COSM6 cells. Although $(CAG)_{960}$ RNA is predicted to form a thermodynamically stable dsRNA hairpin, it would not be predicted to interact with MBNL1 based on the deviation of its primary sequence from the consensus YGCY binding site in which purines have replaced pyrimidines. Despite the ability of $(CAG)_{960}$ to recruit MBNL1 into foci like $(CUG)_{960}$ repeats, the $(CAG)_{960}$ RNA expansion failed to appreciably alter splicing of exon 5 in *cTNT* minigene cotransfection experiments. This result is particularly intriguing considering that $(CUG)_{54}$ and $(CAG)_{54}$ shared similar affinities (MBNL1:CUG_{54} K_d=5.3 ± 0.6 nM and MBNL1:CAG_{54} K_d=11.2 ± 1.5 nM) in a filter binding experiments (Yuan et al., 2007), although this assay does not take into account variable RNA:protein binding modes. Interestingly, MBNL1 proteins form an oligomeric ring structure, with a prominent central hole that is large enough to accommodate dsRNA (Yuan et al., 2007). These structures are visible in an electron micrograph and show that MBNL1 stacks multiple rings on a $(CUG)_{136}$ dsRNA. However, these experiments have not been conducted with $(CAG)_n$ repeats. One explanation for this observation is that while MBNL1 interacts with both repeats, the nature of the RNA:protein interactions are

inherently different. Upon closer inspection, MBNL1:$(CAG)_{960}$ foci appeared larger and less compact than MBNL1:$(CUG)_{960}$ foci when visualized by RNA fluorescent *in situ* hybridization (FISH). This difference in appearance may be reflective of divergent molecular interactions.

Endogenous MBNL1 targets and $(CUG)_n$ repeats share similar characteristics, namely dsRNA substrates that contain paired GC dinucleotides with unpaired pyrimidine bulges. While it is likely that MBNL1 recognizes base-paired GC dinucleotides in $(CAG)_{960}$, the adenosine mismatches may prevent MBNL1 from stably interacting, via its ring structure, with the repeats like it does with $(CUG)_{960}$. In other words, MBNL1:CAG interactions may consist of monomer binding, while MBNL1:CUG repeats are "locked" in position through oligomeric binding and the subsequent stacking of MBNL1 rings. If this model is correct, then multiple mismatches in the DM2 $(CCUG)_n$ repeat could also potentially disrupt stable MBNL1 binding. To investigate these interactions, we performed competition assays in which MBNL1:RNA structures were allowed to form under splicing conditions and then challenged with 2000X fold excess competitor to assay for RNP stability. $(CUG)_{54}$ demonstrated highly stable MBNL1 interactions, while $(CCUG)_{46}$,$(CAG)_{54}$, and T5.45 (an endogenous MBNL1 binding site responsible for fetal exon exclusion in *TNNT3*) displayed less stable binding. This result provides additional evidence that MBNL1 interactions with microsatellite repeats are not the same. However, in this assay, MBNL1:$(CCUG)_{46}$ is even less stable than T5.45 and $(CAG)_{54}$. Unlike T5.45 and $(CAG)_{54}$, $(CCUG)_n$ repeats can cause DM symptoms and missplicing events. How do less stable $(CCUG)_n$ repeats promote disease while $(CAG)_n$ repeats do not? The answer may be a combination of variables that are necessary for

$(CCUG)_n$ to promote the onset of disease. The largest CCTG expansion in *CNBP* has been reported to be ~11,000 repeats, making the DM2 expansion ~3 times larger than the DM1 expansion (Ranum and Day, 2002). Although the CCUG repeat displays less stable MBNL1 interactions than CUG, the increase in repeat length may be sufficient to sequester enough MBNL1 to promotes onset of the disease. Additionally, *CNBP* expression levels have been estimated to be higher than that of *DMPK*, providing more toxic molecules to relocalize MBNL1 (Mankodi et al., 2003). Despite larger repeats and higher expression levels, the later age of onset and generally more mild symptoms characteristic of DM2 may be reflective of reduced CCUG repeat toxicity.

In conclusion, the dysregulation of alternative splicing has become a molecular hallmark of DM, affecting multiple cellular pathways and contributing to the clinical manifestations of disease. Our results provide evidence that MBNL1, a protein whose function is impaired by a gain-of-function at the RNA level in DM, is an alternative splicing factor that promotes adult-specific splicing by directly interacting with pre-mRNA. Additionally, MBNL1 displays variable stabilities when interacting with RNA repeat expansions. Underappreciated differences between $(CUG)_{exp}$ and $(CCUG)_{exp}$ RNAs and their interacting factors could potentially contribute to the immense variability of symptoms seen in DM patients.

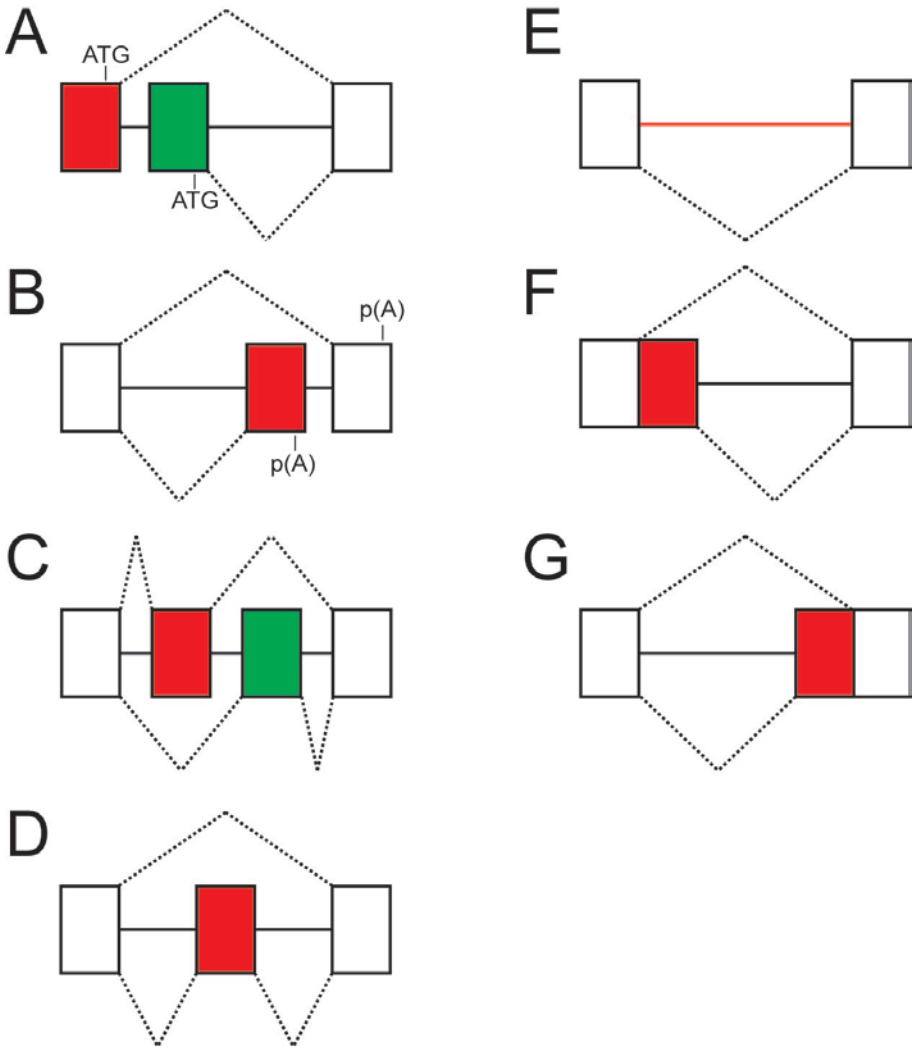

Figure 2-1. Patterns of alternative splicing. Alternatively spliced (red) and mutually exclusive (green) exons spliced into constitutive exons (open boxes) produce distinct mature mRNAs via (A) alternative promoters and (B) 3' exons, (C) mutually exclusive and (D) cassette exons, (E) intron retention, and (F) cryptic 5' and (G) 3' splice sites.

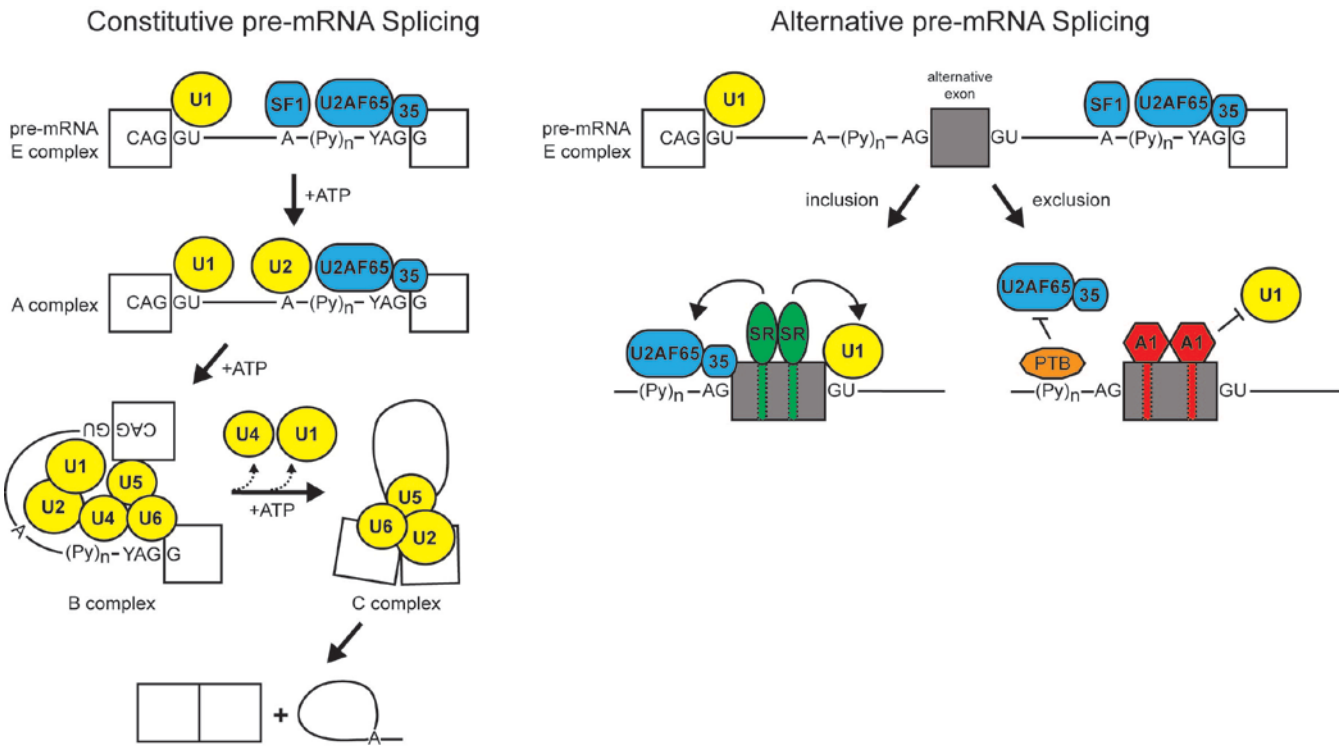

Figure 2-2. Schematic of constitutive and alternative pre-mRNA splicing mechanism. Constitutively spliced exons (open boxes) contain strong consensus 5' (CAGGU) and 3' (YAGG) splice sites that interact with snRNPs (yellow) and accessory proteins (blue) of the basal splicing machinery to define exons for splicing in the pre-mRNA. U1 snRNP (5' splice site), SF1 (branch point adenosine), U2AF65 (polypyrimidine tract), and U2AF35 (3' splice site) initially bind the pre-mRNA (E complex), followed by U2 substitution for SF1 (A complex), introduction of the tri-snRNP (U4/U6/U5), and ATP dependant rearrangements of the spliceosome (B comlex). Further ATP-dependant spliceosomal rearrangements, including disassociation of U4 and U1, are required to form the catalytically active C complex to promote transesterification reactions, join the exons and remove the intron lariat. Alternative spliced exons (grey boxes) often contain weak consensus 5' (GU) and 3' (AG) splice sites and smaller, less defined polypyrimidine tracts. Therefore, additional proteins are required to promote or inhibit the assembly of the sliceosome. SR proteins (green) interact with exonic splicing enhancers to define an alternatively spliced exon and assist in the recruitment of the sliceosome. Conversely, hnRNP proteins (e.g. hnRNP A1 – red) interact with exonic splicing silencers and polypyrimidine tract binding protein (PTB – orange) binds the polypyrimidine tract to block assembly of the spliceosome and prevent inclusion of the exon in the mature mRNA.

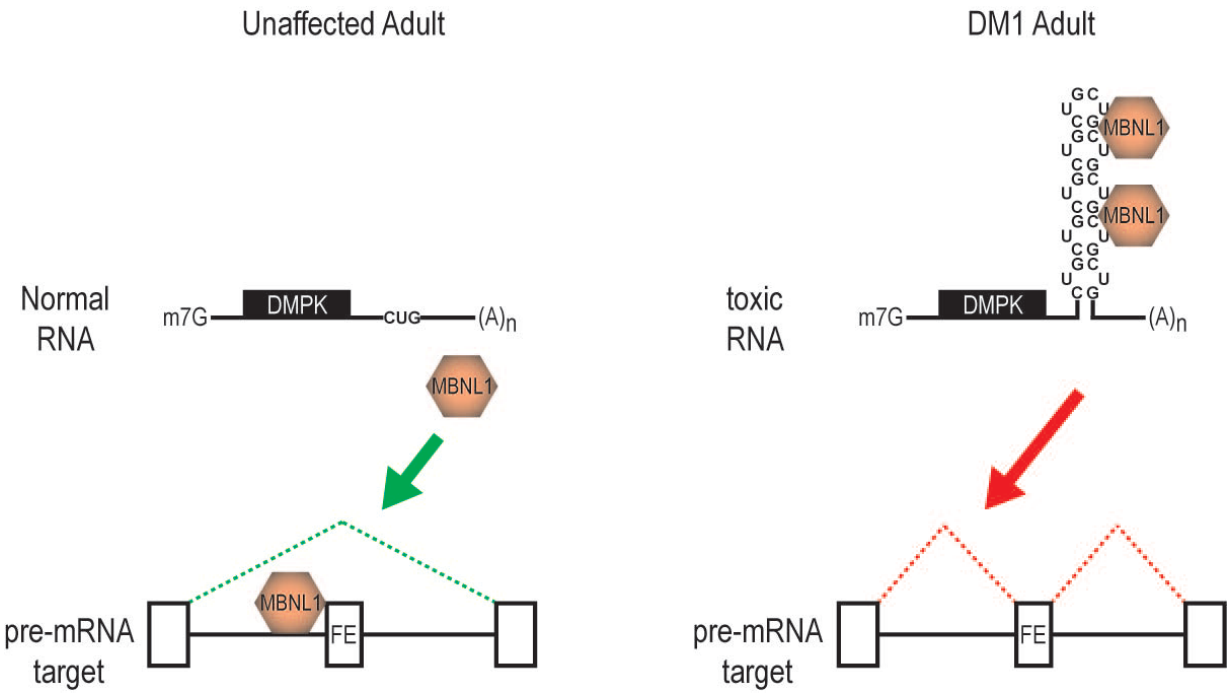

Figure 2-3. Model of MBNL1 sequestration promoting alternative splicing defects in DM1. In unaffected individuals, *DMPK*-$(CUG)_{5-37}$ mRNA does not interact with MBNL1, allowing MBNL1 to directly bind pre-mRNA substrates (exons = open boxes, introns = horizontal lines) and promote fetal exon (FE) exclusion (green). In DM1 affected individuals, *DMPK*-$(CUG)_{>50}$ sequester MBNL1, inhibiting MBNL1 interaction and promoting exon inclusion (red).

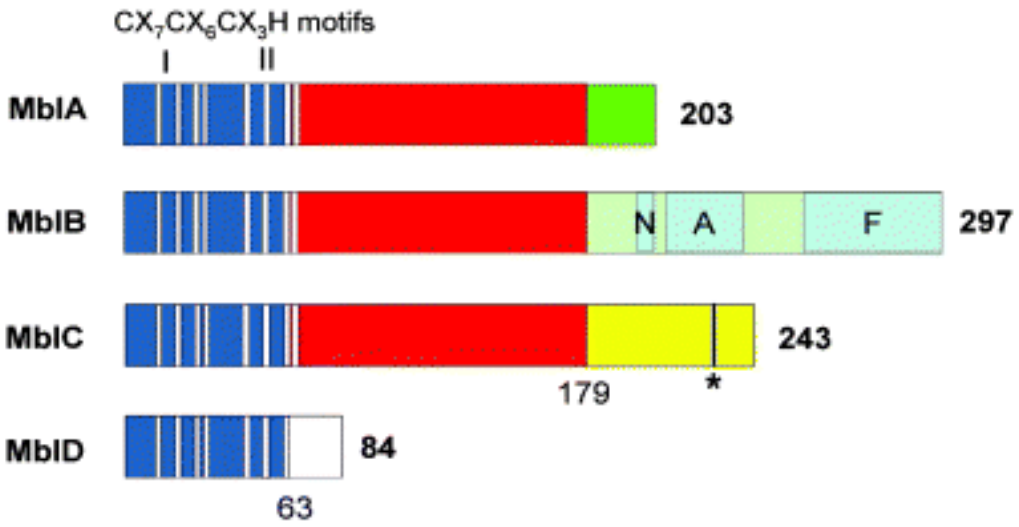

Figure 2-4. Alternative splicing produces four distinct Mbl isoforms in Drosophila. All isoforms share the N-terminal 63 amino acids (blue) which contain $2XC_3H$ RNA binding motifs (cysteine and histidine positions are indicated by vertical white lines). MblA, MblB, MblC share an additional 116 amino acids (red). All isoforms have unique C-termini. MblB contains arginine (N), alanine (A), and phenylalanine (F) rich regions. MblC has a putative (*) sumoylation site. Reproduced from *Muscleblind isoforms are functionally distinct and regulate α-actinin splicing*; Vicente M, Monferrer L, Poulos MG, Houseley J, Monckton DG, O'Dell KM, Swanson MS and Artero RD; Copyright 2007, with permission from John Wiley & Sons.

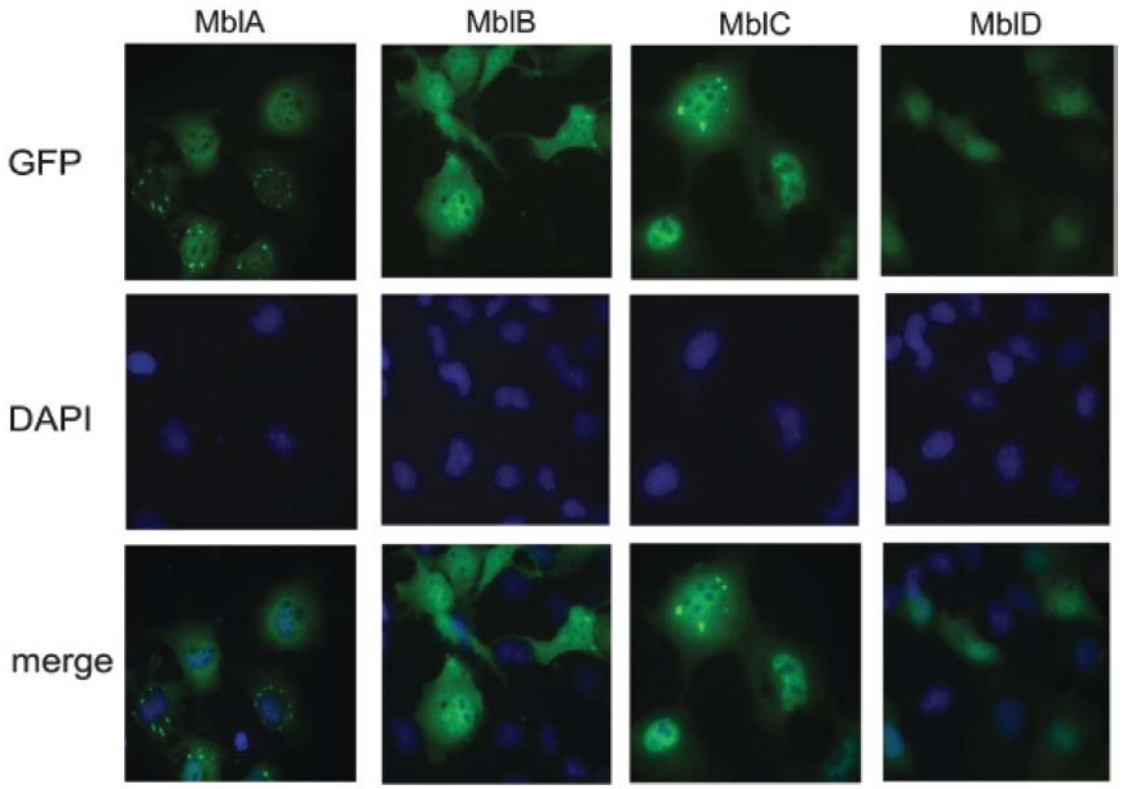

Figure 2-5. Mbl protein isoforms display different subcellular localizations. GFP-tagged Mbl isoforms were exogenously expressed in COSM6 cells. MblA appears cytoplasmic and nuclear while MblB and MblC are predominantly nuclear. MblD is diffuse throughout the cell. DAPI stain indicates nuclear location. Reproduced from *Muscleblind isoforms are functionally distinct and regulate α-actinin splicing*; Vicente M, Monferrer L, Poulos MG, Houseley J, Monckton DG, O'Dell KM, Swanson MS and Artero RD; Copyright 2007, with permission from John Wiley & Sons.

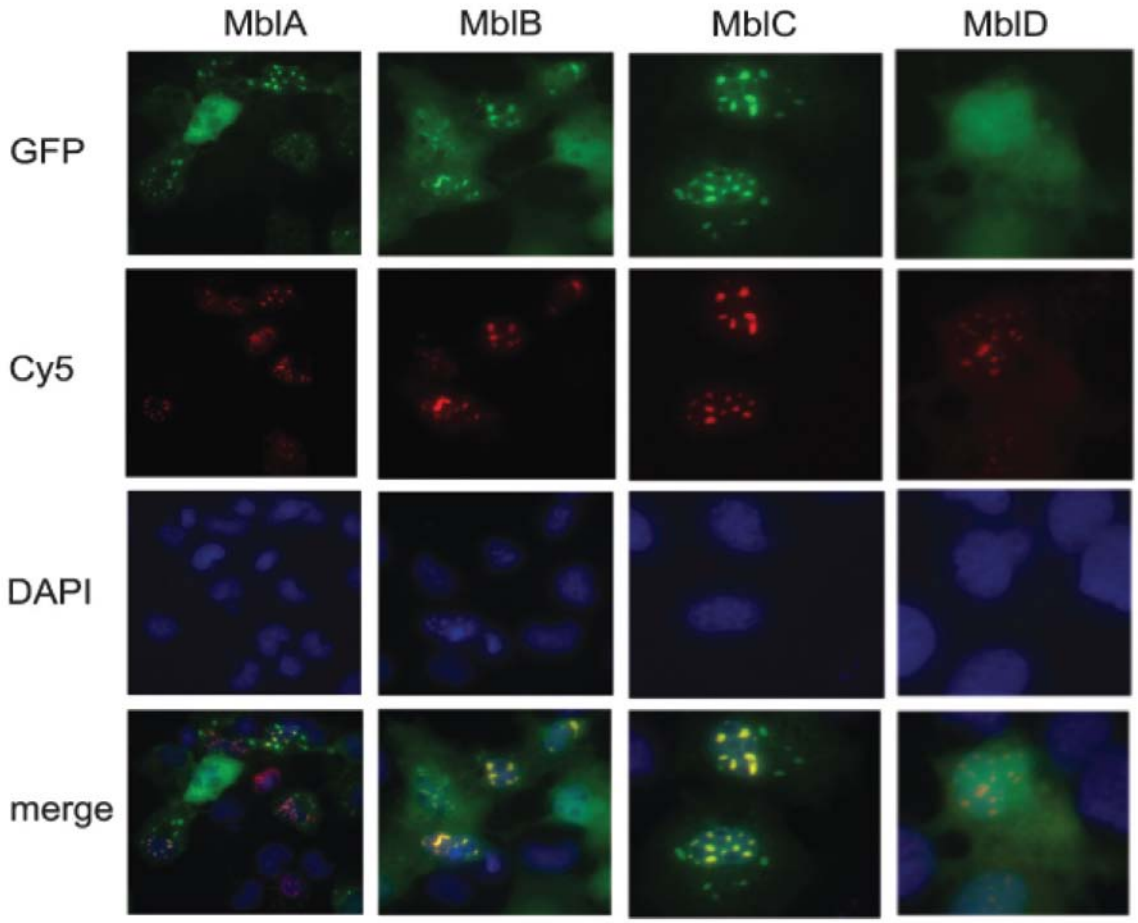

Figure 2-6. Mbl colocalizes with $(CUG)_{300}$ RNA in nuclear foci. GFP-tagged Mbl isoforms were cotransfected with $(CUG)_{300}$ in COSM6 cells. MblA, MblB, and MblC (green) colocalize with $(CUG)_{300}$ (red – labeled with Cy3-CAG_{10} oligonucleotide probe) in discrete nuclear foci. DAPI indicates nuclear location. Reproduced from *Muscleblind isoforms are functionally distinct and regulate α-actinin splicing*; Vicente M, Monferrer L, Poulos MG, Houseley J, Monckton DG, O'Dell KM, Swanson MS and Artero RD; Copyright 2007, with permission from John Wiley & Sons.

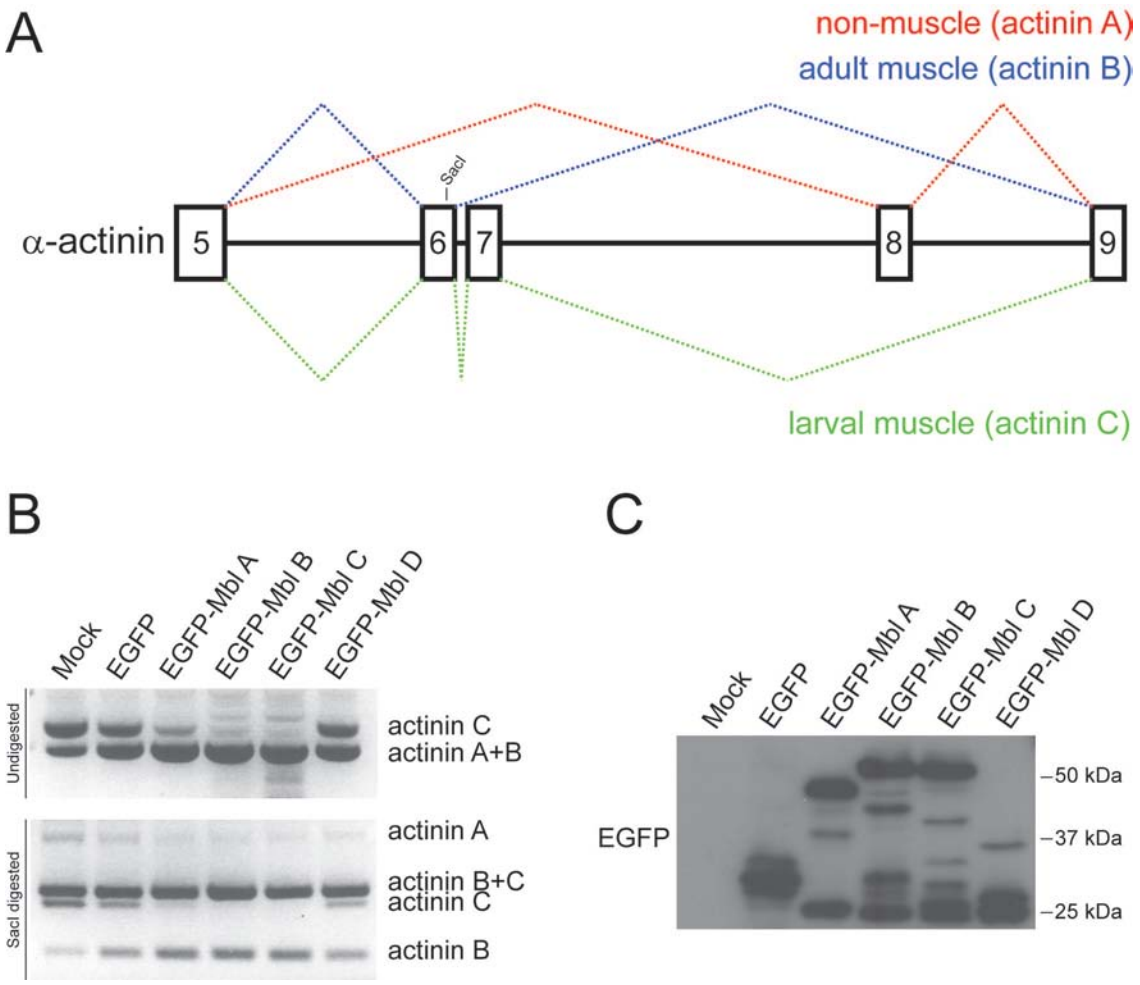

Figure 2-7. *Drosophila Mbl* regulates alternative splicing of *α-actinin*. An *α-actinin* minigene was coexpressed with EGFP tagged Mbl isoforms and assayed for alternative splicing by RT-PCR with primers positioned in constitutive exons 5 and 9. (A) The *α-actinin* minigene contains exons 5-9 (exons = open boxes; introns = horizontal lines). *α-actinin* is alternatively spliced to produce adult muscle (blue), larval muscle (green), and non-muscle (red) isoforms. (B) Mbl isoforms A, B, and C, but not D, promote an adult muscle splicing pattern. EGFP is used as a control. Because *α-actinin* adult muscle and non-muscle spliced mRNAs are the same size, a unique *Sac*I site is used to digest and resolve the resulting RT-PCR products. (C) An immunoblot using an EGFP specific antibody was used to control for *EGFP-Mbl* expression (expected sizes: MblA = 49 kDa, MblB = 58 kDa, MblC = 53 kDa, MblD = 35 kDa). Ponseau-S staining was used to control for loading (data not shown).

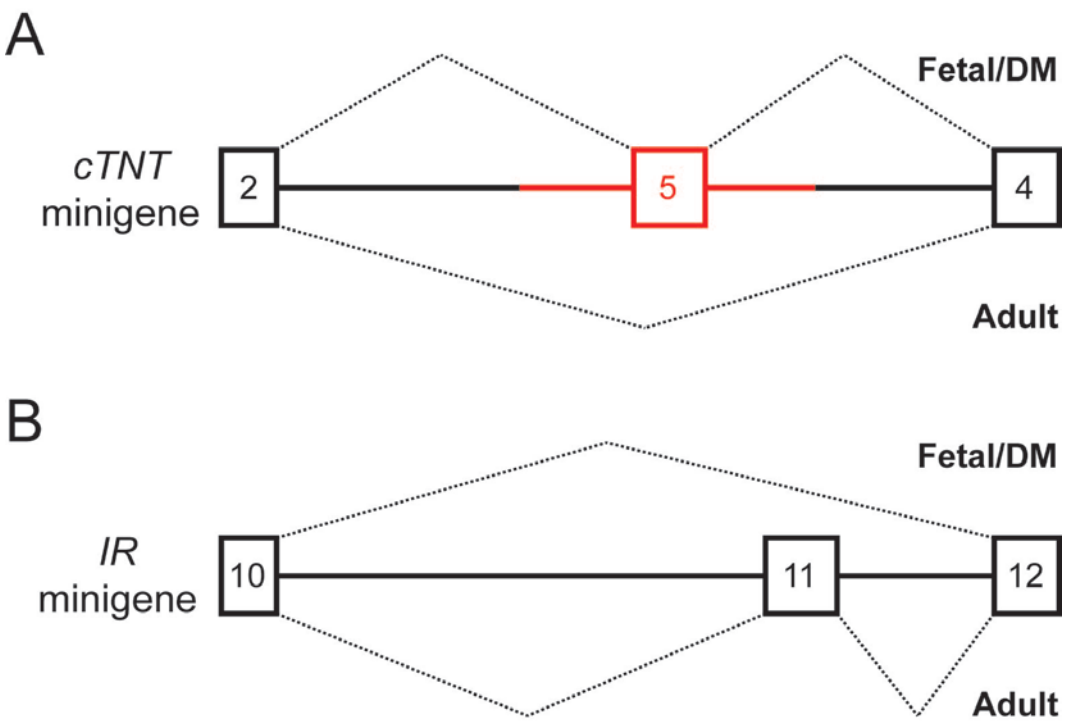

Figure 2-8. Schematic of *cTNT* and *INSR* alternative splicing minigene reporters. (A) cTNT minigene. A 730 nucleotide fragment including alternatively spliced human *cTNT* exon 5 (30 nucleotides) and adjacent upstream/downstream introns (red) cloned in between constitutively spliced exons 2 and 4 from the *Gallus gallus TNNI2* gene (black). Primers used for RT-PCR analysis are located in exons 2 and 4. (B) *INSR* minigene. A fragment from the human *INSR* locus containing alternatively spliced exon 11 and constitutive exons 10 and 12. Primers used for RT-PCR analysis are located in exons 10 and 12. Fetal/DM and adult splicing patterns are indicated by dashed lines. Minigenes are not drawn to scale.

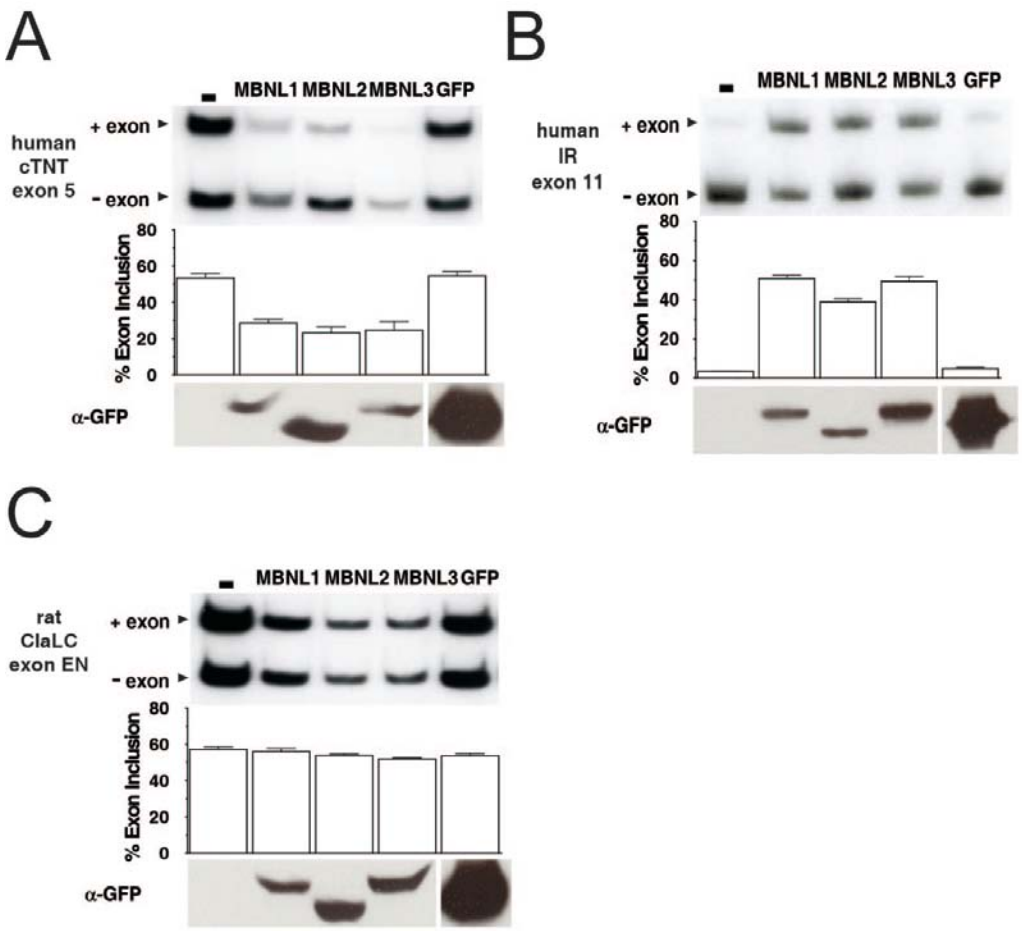

Figure 2-9. MBNL1, MBNL2, and MBNL3 regulate alternative splicing of *cTNT* and *INSR*. *cTNT* and *INSR* minigenes were coexpressed with GFP-tagged MBNL proteins in primary chicken myoblasts and assayed for inclusion of alternatively spliced exons by RT-PCR with primers positioned in constitutive exons. Bands were quantified by phosphoimager and exon inclusion was calculated as: [(exon inclusion)/(exon inclusion + exon exclusion) X100]. Transgene expression was monitored by immunoblot with an antibody specific for GFP. MBNL1, MBNL2, and MBNL3 promotes (A) fetal exon 5 exclusion of *cTNT* and (B) exon 11 inclusion of *INSR* minigenes while (C) a neuronal specific exon is unresponsive in the *ClaLC* minigene. Reproduced from *Muscleblind proteins regulated alternative splicing*. Ho TH, Charlet B, Poulos MG, Singh G, Swanson MS and Cooper TA; *EMBO*, Vol. 23, No 15, 3103-3112. Copyright 2004, with permission from The Nature Publishing Group.

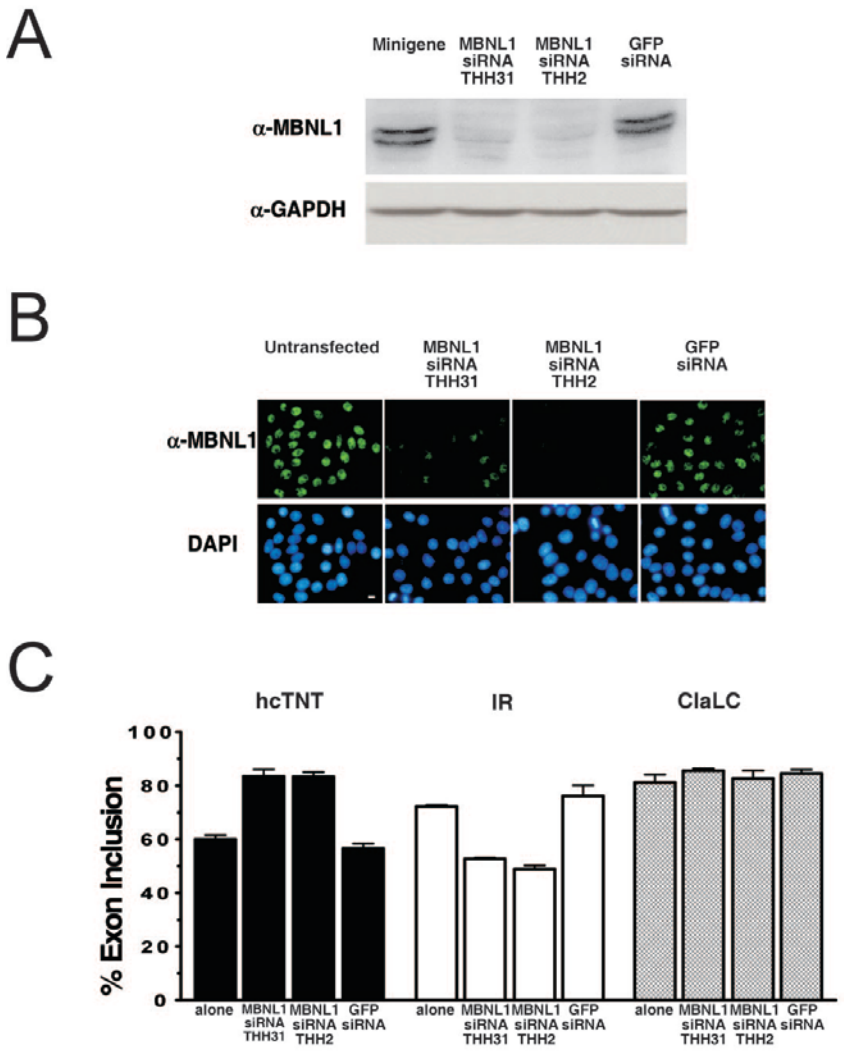

Figure 2-10. Endogenous MBNL1 regulates alternative splicing of *cTNT* and *INSR* minigenes. *cTNT* and *INSR* minigenes were cotransfected with siRNA directed against the *MBNL1* coding sequence (THH31 and THH2) and assayed for alternative splicing by RT-PCR (described in Figure 1-6). (A) siRNA knocks down endogenous MBNL1 41/42 kDa isoform expression. Immunoblot analysis of HeLa cells treated with THH31 and THH2 using antibodies recognizing MBNL1 and GAPDH (loading control) or (B) immunocytochemistry using antibodies recognizing MBNL1. A non-specific GFP siRNA has no affect on *MBNL1* expression. (C) Loss of endogenous MBNL1 recapitulates a DM splicing pattern in *cTNT* and *INSR* minigenes. Reproduced from *Muscleblind proteins regulated alternative splicing*. Ho TH, Charlet B, Poulos MG, Singh G, Swanson MS and Cooper TA; *EMBO*, Vol. 23, No 15, 3103-3112. Copyright 2004, with permission from The Nature Publishing Group.

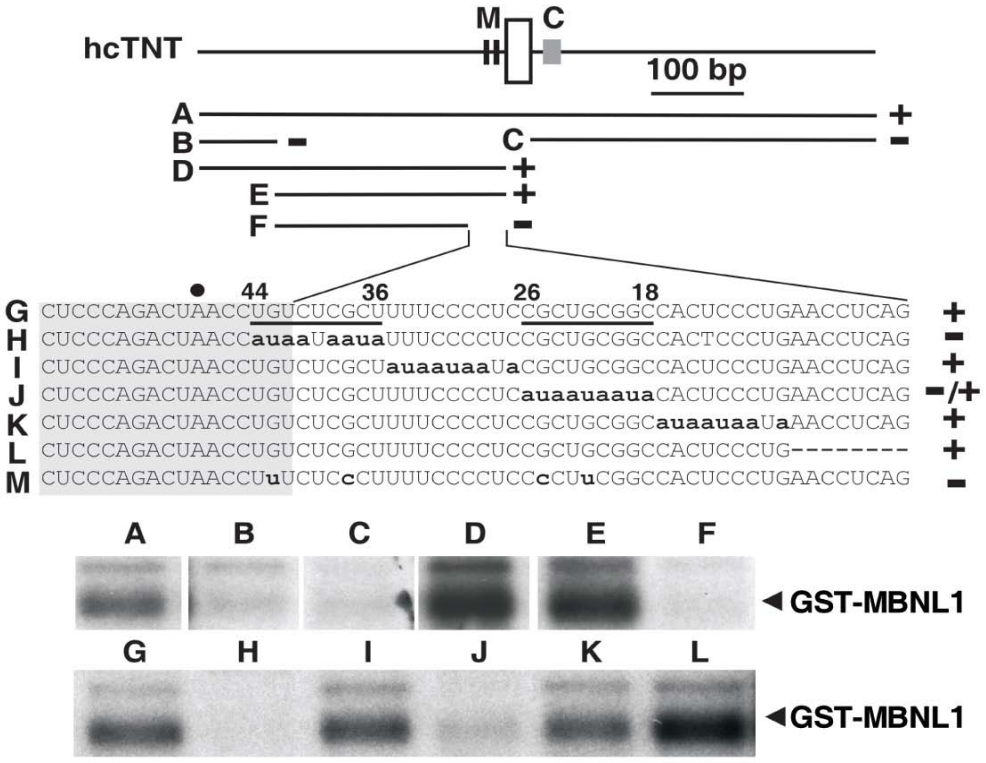

Figure 2-11. MBNL1 binds upstream of alternatively spliced *cTNT* exon 5. *cTNT* RNA from the alternatively spliced exon 5, with upstream and downstream introns, were uniformly body labeled with ^{32}P and incubated with recombinant GST-MBNL1, UV crosslinked and resolved on an non-denaturing acrylamide gel to assay for binding. Truncation and scanning mutagenesis of the RNA was performed to identify the MBNL1 binding site. (+) and (-) indicate GST-MBNL1 and RNA binding. The putative branch point adenosine (black circle) and RNA mutations (lowercase) are denoted. MBNL1 (two vertical lines) and antagonist CUGBP1 (grey box) binding sites are indicated on the *cTNT* RNA. Two dinucleotide substitutions (RNA M), which eliminate GST-MBNL1 binding, will be used for *in vitro* splicing assays. Reproduced from *Muscleblind proteins regulated alternative splicing*. Ho TH, Charlet B, Poulos MG, Singh G, Swanson MS and Cooper TA; *EMBO*, Vol. 23, No 15, 3103-3112. Copyright 2004, with permission from The Nature Publishing Group.

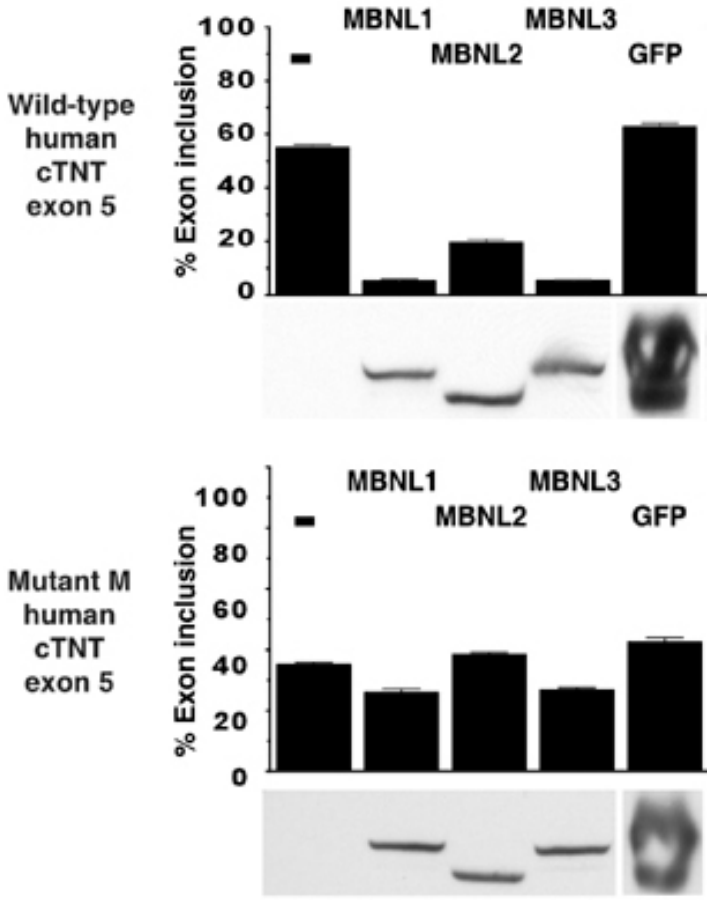

Figure 2-12. MBNL1 binding site mutations inhibit MBNL1, MBNL2, and MBNL3 responsiveness. Wild-type or mutant (RNA M in Figure 1-8) *cTNT* minigenes were cotransfected with GFP-tagged MBNL1, MBNL2, or MBNL3 and assayed for alternative splicing by RT-PCR (described in Figure 1-6). Transgene expression was monitored by immunoblot using an antibody specific for GFP. Reproduced from *Muscleblind proteins regulated alternative splicing*. Ho TH, Charlet B, Poulos MG, Singh G, Swanson MS and Cooper TA; *EMBO*, Vol. 23, No 15, 3103-3112. Copyright 2004, with permission from The Nature Publishing Group.

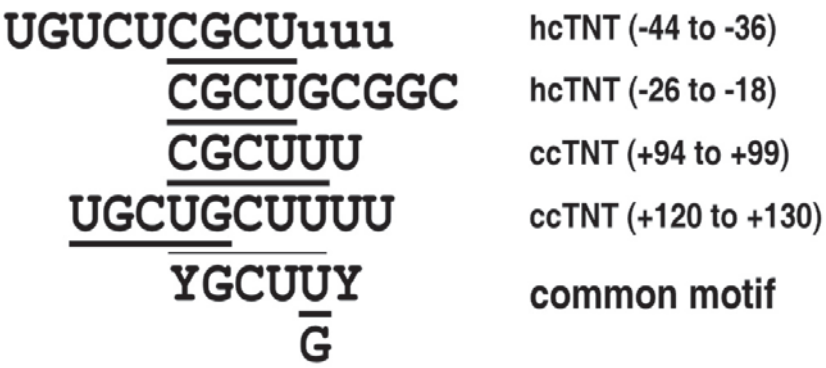

Figure 2-13. Alignment of the human and chicken *cTNT* MBNL1 binding sites reveals a conserved motif. Reproduced from *Muscleblind proteins regulated alternative splicing*. Ho TH, Charlet B, Poulos MG, Singh G, Swanson MS and Cooper TA; *EMBO*, Vol. 23, No 15, 3103-3112. Copyright 2004, with permission from The Nature Publishing Group.

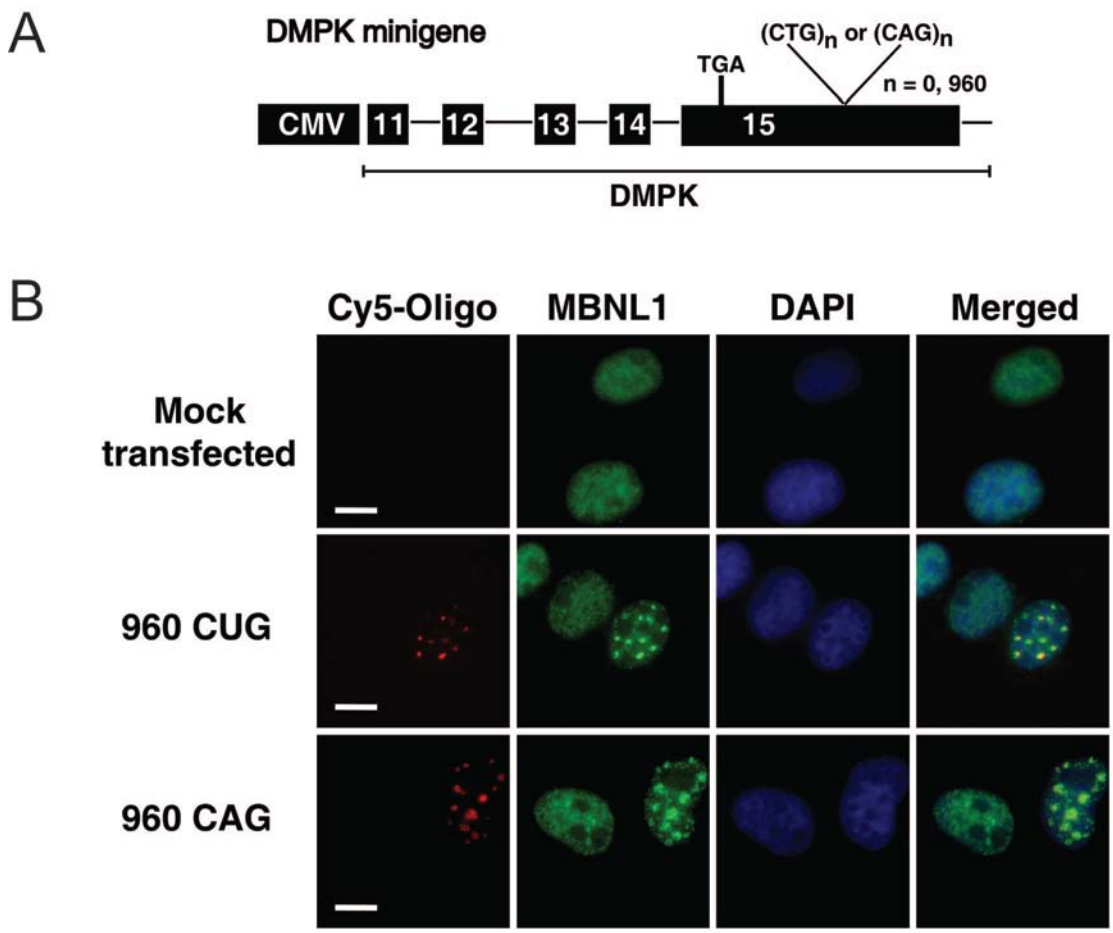

Figure 2-14. MBNL1 colocalizes with $(CUG)_{960}$ and $(CAG)_{960}$ RNA in nuclear foci. (A) *DMPK* minigene containing a *CMV* promoter and *DMPK* exons 11-15, with 960 interrupted CUG or CAG repeats in the 3'UTR. (B) *DMPK* $(CUG)_{960}$ or $(CAG)_{960}$ exogenously expressed in COSM6 cells. Immunocytochemistry with an antibody specific for MBNL1 (green) demonstrates colocalization with both CUG_{960} and CAG_{960} RNA (red – labeled with Cy5-CAG_{10} or Cy5-CUG_{10}, respectively). DAPI indicates nuclear location. Scale bar = 10 μm. Reproduced from *Colocalization of muscleblind with RNA foci is seperable from mis-regulation of alternative splicing in myotonic dystrophy*. Ho TH, Savku RS, Poulos MG, Mancini MA, Swanson MS and Cooper TA; *Journal of Cell Science*, 118, 2923-2933. Copyright 2005, with permission from The Company of Biologists Ltd.

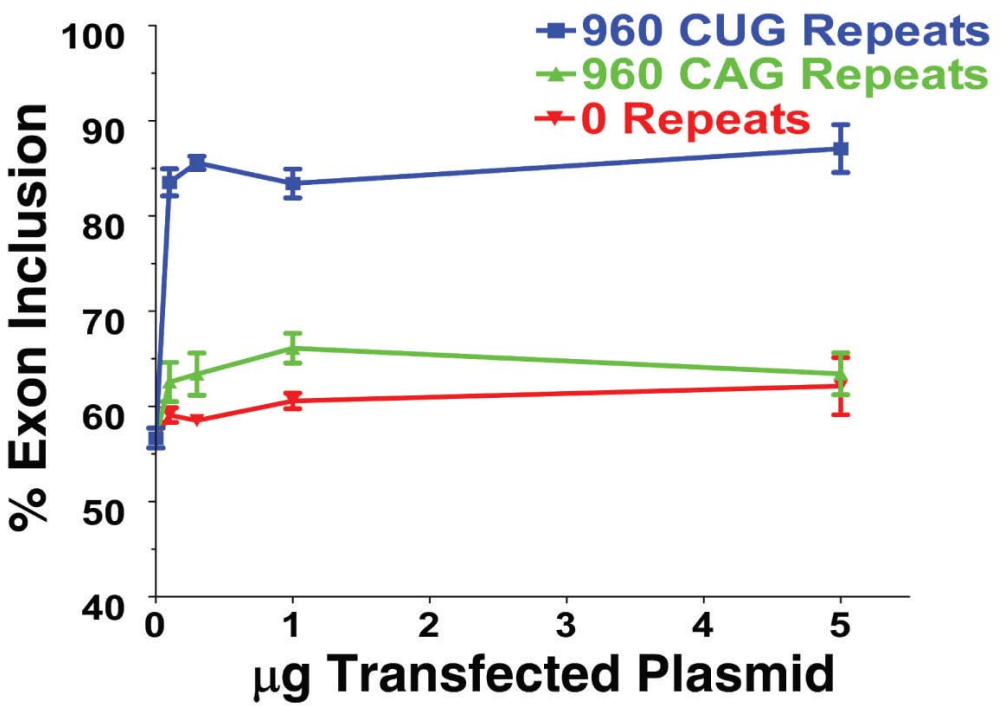

Figure 2-15. CUG_{960}, but not CAG_{960}, expression alters *cTNT* alternative splicing. A *cTNT* minigene, containing an alternatively spliced fetal exon 5 which is misregulated in DM, is cotransfected with either *DMPK* $(CUG)_{960}$ or *DMPK* $(CAG)_{960}$ and assayed for exon 5 inclusion using RT-PCR and primers positioned in constitutive exons. Three independent transfections were done and exon inclusion was calculated as: (exon inclusion)/(exon inclusion + exon exclusion). Increasing amounts of $(CUG)_{960}$ and $(CUG)_{960}$ do not significantly alter *cTNT* splicing. Reproduced from *Colocalization of muscleblind with RNA foci is seperable from mis-regulation of alternative splicing in myotonic dystrophy*. Ho TH, Savku RS, Poulos MG, Mancini MA, Swanson MS and Cooper TA; *Journal of Cell Science*, 118, 2923-2933. Copyright 2005, with permission from The Company of Biologists Ltd.

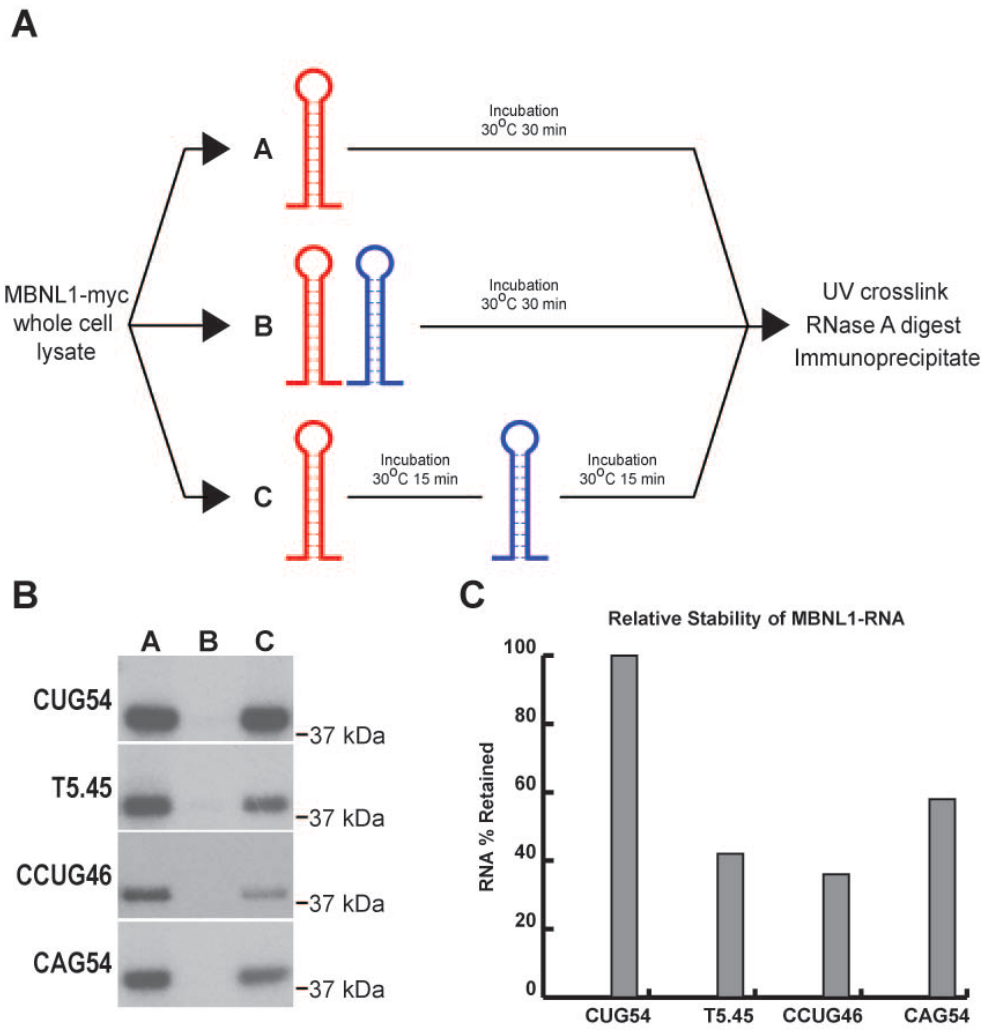

Figure 2-16. MBNL1 displays increased stability with $(CUG)_{54}$ RNA. (A) Schematic of the competition assay. MBNL1-myc whole cell lysate (expressed in COSM6 cells) is incubated with ^{32}P labeled RNA (red) alone for 30 minutes (A), concurrently with 2000X fold excess cold RNA (blue) for 30 minutes (B), or alone for 15 minutes and then challenged with 2000X fold excess cold RNA for 15 minutes (C). RNA:protein complexes are UV crosslinked, RNase digested, immunoprecipitated with an antibody specific for myc, and resolved by SDS-PAGE. (B) Autoradiographs of RNAs used in the competition assay: $(CUG)_{54}$; T5.45 (an endogenous MBNL1 binding site on the *TNNT3* mRNA responsible for alternative splicing); $(CCUG)_{46}$; $(CAG)_{54}$. (C) A representative bar graph indicating the percentage of ^{32}P labeled RNA retained when challenged. Bands were quantified by phosphoimager and retention was calculated as (C/A X100) from the autoradiograph.

CHAPTER 3
LOSS OF MBNL3 4XC$_3$H ISOFORMS ARE NOT SUFFICIENT TO MODEL CONGENITAL MYOTONIC DYSTROPHY

Introduction

Overview of Congenital Myotonic Dystrophy

Congenital myotonic dystrophy is perhaps the most severe form of DM1, although the disease likely manifests itself during embryonic development (Harper, 2001). Like the adult-onset disease, CDM affects a wide variety of tissues and displays variable penetrance at birth. The most prominent features of CDM include skeletal muscle, smooth muscle, pulmonary, and CNS involvement (Harper, 2001). Unlike adult-onset DM1, however, relatively little is known about the root cause of these symptoms at the molecular level. For example, pulmonary insufficiencies are responsible for the majority of infantile mortality in CDM, requiring prolonged ventilation at birth for survival (Campbell et al., 2004; Rutherford et al., 1989). While it is acknowledged that there is reduced efficiency of the neonatal lung, it is unclear whether respiratory distress is due to the inability of the underdeveloped diaphragm and intercostal muscles to promote breathing, inherent pulmonary defects, or a combination of both (Harper, 2001). Neonates with CDM also present with significant muscle involvement. Hypotonia is caused by a lack of baseline contractions in skeletal muscle that normally promote rigidity and posture (Bergen, 1985; Bodensteiner, 2008). CDM babies are characteristically “floppy” shortly after birth, lacking significant movements (Harper, 1975; Harper, 2001). Additionally, poor suckling may also be due to poorly developed tongue muscle. Histological analysis of CDM skeletal muscle reveals a delay in myogenic development, including the abnormal presence of immature myotubes, small bundles of fibers, thin fibers, and prominent centralized nuclei, while motor neurons

appear normal (Farkas-Bargeton et al., 1988; Sahgal et al., 1983a; Sahgal et al., 1983b; Tanabe and Nonaka, 1987; Tominaga et al., 2010). Difficulties in swallowing postnatally also indicate that the smooth muscle of the esophagus is affected, which may lead directly to another common CDM phenotype, hydramnios, in which an increase of amniotic fluid during pregnancy is caused by a lack of fetal swallowing (Schild et al., 1998; Wieacker et al., 1988). Remarkably, these symptoms, including pulmonary involvements and delays in muscle development, are generally resolved by early childhood and do not affect the prognosis of mobility or respiration later in life (Harper, 2001). CDM patients also display mental retardation and low IQs in adulthood, with one estimated mean of 66.1 +/- 16.2 , although there is little evidence as to the cause (Harper, 1974; Harper, 1975). Postmortem analysis of infants afflicted with CDM and mental retardation has yielded no evidence of gross cerebral morphological changes. However, magnetic resonance imaging (MRI) and computed tomography (CT) scans have indicated changes in ventricular size and an increased frequency in intraventricular haemorraging in neonates, although it is not evident if these changes are directly involved in disease (Regev et al., 1987). Finally, talipes, commonly known as club foot, is seen in approximately half of all CDM patients (Ray et al., 1984; Siegel et al., 1984). In CDM, talipes is thought to be caused by *in utero* muscle weakness (hypotonia) and decreased fetal movements. Surgical correction is commonly used to treat this disability. Ultimately, CDM patients, despite overcoming early muscle and pulmonary deficits, develop early-onset adult DM symptoms, often manifesting itself in early adolescence. While CDM appears clinically distinct from its adult-onset counterparts in DM1 and DM2, the causative mutation is the DM1 $(CTG)_{exp}$ that has

intergenerationally increased in size to greater than 1000 repeats (Tsilfidis et al., 1992). There is no evidence of a DM2 repeat expansion that contributes to the onset of a CDM phenotype, indicating a link between DM1 and disease. This observation begs the question, what causes CDM?

Potential Models of CDM

Two closely related $(CTG)_n$ rich repeat expansions cause very similar adult-onset diseases, DM1 and DM2. MBNL1 loss-of-function through interactions with either $(CUG)_{exp}$ or $(CCUG)_{exp}$ have proven to be the common link between these adult-onset disorders. How then does DM1 specifically promote additional CDM phenotypes during embryogenesis and neonatal life? The presence of a second form of DM that does not cause a congenital phenotype provides an opportunity to compare and contrast potential molecular mechanisms involved in congenital disease. Instead of focusing on the similarities of this disease, the answer likely lies in the differences between the two causative mutations.

The DM1 and DM2 mutations are located in non-coding portions of two seemingly unrelated genes, *DMPK* and *CNBP* respectively. The $(CTG)_{exp}$ promotes adult-onset DM1 between 50 and 1000 repeats, while expansions between 1000-4000 cause CDM. In DM2, $(CCTG)_{exp}$ greater than 75 and less than 11,000 cause the adult-onset disorder. One possibility to explain the observation that only $(CTG)_{exp}$ promote CDM is that embryonic expression patterns between the two genes do not overlap during development. In this scenario, *DMPK*-mediated RNA toxicity could promote the onset of CDM by being preferentially expressed in affected tissues, while *CNBP* expression is either absent or below a threshold of expression needed to promote disease. However, *Dmpk* and *Cnbp* have demonstrated no significant differences in embryonic expression

patterns (Swanson lab unpublished data). Using whole mount in situ hybridization, *Cnbp* is expressed ubiquitously during murine embryonic development while *Dmpk* displays a more restricted expression pattern. Additionally, *Cnbp* staining appears more intense than *Dmpk*, suggesting that *Cnbp* is expressed at a higher level. In an alternative scenario, a more restricted embryonic expression pattern for *DMPK* may promote the onset of CDM in affected tissues, while the ubiquitously expressed *CNBP* results in either an early termination of embryogenesis (i.e. before *DMPK* is expressed) or widespread expression compromises too many tissues which results in an inability to develop and subsequent lethality. Another possibility is that DM1 locus specific affects contribute to neonatal disease. Closely-linked *DMPK* genes, in which expression is inhibited by $(CTG)_{exp}$, may contribute to the onset of congenital disease through a combinatorial model of haploinsufficieny of DM1-linked genes and MBNL1 loss-of-function model (see General Introduction). Genetic crosses to create $Dmpk^{+/-}$; $Six5^{+/-}$; $Mbnl1^{\Delta E3/\Delta E3}$ mice (each individual knockout mouse did not develop a CDM phenotype) failed to recapitulate any CDM phenotypes (Swanson lab unpublished data). However, we cannot rule out the contributions of *DMWD* and any dominant-negative effects of the DM1 antisense transcript. The most likely explanation may be differences in the RNA repeat itself. We have previously demonstrated that $(CUG)_{exp}$ and $(CCUG)_{exp}$ display variable stabilities with MBNL1 *in vitro*, potentially altering its toxicity. Therefore, $(CUG)_{>1000}$ RNAs may be more inherently toxic. Another distinct difference between $(CUG)_{exp}$ and $(CCUG)_{exp}$ is the asymmetry in repeat expansion length. Somatic expansions of both $(CTG)_n$ and $(CCTG)_n$ result in a heterogeneous population of repeats in different tissues and time points throughout life that can contribute to multiple

aspects of adult-onset disease, including age-of-onset, severity of disease, and penetrance of symptoms (Ashizawa et al., 1993; Martorell et al., 1997; Martorell et al., 2000; Wong et al., 1995). On the other hand, both repeats reach a maximum expansion $(CTG)_{\sim 4000}$ and $(CCTG)_{\sim 11,000}$, suggesting that further expansions are incompatible with life. One possibility is that $(CUG)_n$ and $(CCUG)_n$ could both promote similar CDM phenotypes embryonically, however, unstable $(CCTG)_n$ expansions reach a critical length and promote embryonic lethality. This possibility would predict that longer $(CCTG)_n$ expansions are more unstable during transmission or development than expanded $(CTG)_n$ repeats. While these models are not mutually exclusive, they provide insight into the complexities involved in determining and modeling specific genetic contributions to the congenital onset of disease.

MBNL Family in DM

Considering that $(CUG)_{exp}$ contributes to adult-onset DM1 by sequestering a dsRNA binding protein, MBNL1, and inhibiting its function, the same toxic RNA could compromise other factors via dominant-negative interactions. Interestingly, *MBNL1* belongs to a family composed of three closely related genes, including *MBNL2* and *MBNL3* (Fig. 3-1). All three genes code for proteins that contain $4XC_3H$ motifs, which mediate RNA interactions (Teplova and Patel, 2008). MBNL shares >60% conserved amino acid identity in the full length proteins and on average ~90% in their C_3H motifs (Fig. 3-2). We have previously demonstrated that all three proteins can interact to suppress inclusion of *cTNT* exon 5 in cell culture (see Chapter 2 – Fig. 2-9), suggesting that their substrate specificity overlaps. Additionally, GFP-tagged MBNL1, MBNL2, and MBNL3 colocalize with $(CUG)_{exp}$ RNAs into discrete nuclear foci when ectopically expressed in cells (Fardaei et al., 2002). Moreover, MBNL1 and MBNL2 have been

shown to colocalize with $(CUG)_{exp}$ in the skeletal muscles of DM1 patients, further implicating the family in disease pathogenesis (Holt et al., 2009).

While *Mbnl1*$^{\Delta E3/\Delta E3}$ knockout mice recapitulate cardinal DM phenotypes, genetically altered *Mbnl2* mice have provided variable results. Two different mouse models with decreased levels of *Mbnl2* mRNA have been generated using genetrap technology in which a *Neo*^{R}/*LacZ* cassette traps expression using a novel 3' splice site to truncate the full length endogenous transcript. The first mouse contains a genetrap in *Mbnl2* intron 2 (*Mbnl2*$^{GT2/GT2}$) which produces a truncated protein containing the first 58 amino acids of Mbnl2 ($1XC_3H$ motif) fused with the reporter cassette and displays a significant decrease in overall *Mbnl2* mRNA (Hao et al., 2008). The *Mbnl2*$^{GT2/GT2}$ mouse displays multiple phenotypes, including skeletal muscle centralized nuclei, myotonic discharges, minor changes in *CLCN1* alternative splicing, and focal loss of CLCN1 localization to the sarcolemma. Interestingly, the second mouse, which contains a genetrap in *Mbnl2* intron 4 (*Mbnl2*$^{GT4/GT4}$) and produces a protein containing the first 180 amino acids of Mbnl2 ($2XC_3H$ motifs) fused with the reporter cassette, demonstrates normal adult skeletal muscle by histology and no changes in the alternative splicing of *CLCN1* or myotonia, despite a >90% loss of *Mbnl2* mRNA (Lin et al., 2006). One possibility to account for the observed differences in the two models is that while *Mbnl2*$^{GT4/GT4}$ mice have a greater reduction in endogenous *Mbnl2* mRNA than *Mbnl2*$^{GT2/GT2}$ mice, the pair of Mbnl2 N-terminal C_3H motifs may be sufficient to prevent the skeletal muscle pathology *in vivo*. Generation of a *Mbnl2* knockout mouse would help to clarify these results. Therefore, while the exact role of *MBNL2* in DM is unclear, *Mbnl2*$^{Gt/Gt}$ mice do not develop CDM or other unmodeled DM phenotypes, including progressive skeletal

muscle wasting. There are currently no *in vivo* genetic models for *Mbnl3* involvement in disease. Given the integral involvement of MBNL1 in adult-onset DM, compromised function of MBNL2 and MBNL3 are promising candidates for contribution to the onset and progression of disease. However, given the genetic contribution of the many genes that are potentially affected by the $(CTG)_{exp}$, a comprehensive model may be necessary to fully dissect the constellation of DM phenotypes (Fig. 3-3).

Overview of MBNL3

MBNL3 was first identified in a bioinformatic search for proteins that shared a high degree of identity to MBNL1 and later in a suppression subtractive hybridization (SSH) experiment designed to screen for genes preferentially expressed during myogenic proliferation (Fardaei et al., 2002; Squillace et al., 2002). *MBNL3* is an interesting candidate gene in CDM for a variety of reasons. First, *Mbnl3* is primarily expressed embryonically. Northern blot analysis of poly A selected RNA from adult murine tissues reveals very low *Mbnl3* expression in all adult tissues tested, despite a ten-fold higher exposure time than *Mbnl1* and *Mbnl2*, which displayed near ubiquitous expression (Kanadia et al., 2003b). *Mbnl3* expression was detectable by RT-PCR in adult murine lung, spleen, and testis, however, this analysis is not quantitative and likely is reflective of either low level transcription throughout the tissues tested or a subpopulation of cells in these tissues that express *Mbnl3* preferentially (Lee et al., 2007). Whole embryo total RNA northern blotting from various murine embryonic stages indicate a high level of *Mbnl3* expression, comparable to *Mbnl1* and *Mbnl2* (Kanadia et al., 2003b). However, the RNA for this analysis was collected from total embryos and does not offer any insight spatially during development. Additionally, *Mbnl3* is known to be expressed highly in the placenta, which was included with the embryos, making interpretation of

actual *Mbnl3* levels attributable to embryonic expression difficult. Whole mount RNA *in situs* demonstrate that *Mbnl3* mRNA is present in the limb bud, first brachial arch, and the neural tube at E9.5 (Kanadia et al., 2003b). This data indicates that *Mbnl3* expression, unlike that of *Mbnl1* and *Mbnl2*, is limited to embryonic development. This data is based solely on mRNA expression. Interestingly, *Mbnl3* mRNA contains a unusually large 3' UTR (~ 8 kb), making it subject to post-transcriptional regulation. In other words, Mbnl3 protein levels may not accurately reflect the corresponding mRNA levels *in vivo*. Nonetheless, an overlap in *Mbnl3* expression and tissues affected in CDM is consistent with an MBNL3 loss-of-function model for embryonic and neonatal disease. Second, *Mbnl3* expression has been shown to affect myogenic differentiation *in vitro*. *Mbnl3* mRNA is expressed in proliferating C2C12 cells and is down-regulated after differentiation (Lee et al., 2007; Squillace et al., 2002). Overexpression of *Mbnl3* in stably transfected C2C12 cells inhibits differentiation, while 50% knockdown of Mbnl3 using an antisense morpholino promotes the opposite effect, an increase in differentiation. This data suggests a critical role for *Mbnl3* during normal myogenesis. Therefore, inhibition of this function due to sequestration may play an important role in the misregulation of embryonic skeletal muscle development seen in CDM.

Results

Mbnl3 mRNA Expression and Alternative Splicing are Spatially and Temporally Regulated

Mbnl3 has been previously reported to be expressed during embryonic development with the highest levels in the placenta (Fardaei et al., 2002). However, this analysis focused on whole embryos and non-quantitative methods. Therefore, we chose to detail *Mbnl3* expression throughout development both spatially and temporally

using Northern blotting (Fig. 3-4). *Mbnl3* mRNA is expressed at its highest levels at embryonic day 15 (E15) in placenta, forelimb and tongue with levels reduced in forelimb and tongue by E18 and undetectable in adult tissues. *Mbnl3* is also present in E18 and postnatal day 1 lung (P1) and absent in adult. Expression is also noted in P1 and adult spleen. *Mbnl3* family member, *Mbnl1*, is expressed in more tissues during development. This analysis confirms and extends previous reports of *Mbnl3* expression to specific time points and tissues during development. It is worth noting that *Mbnl3* mRNAs are expressed during embryogenesis in those tissues affected in the congenital form of myotonic dystrophy.

Although *Mbnl3* Northern blotting reveals expression, it fails to detail the mRNAs that compose the population. *Mbnl1* and *Mbnl2* are extensively alternatively spliced, producing multiple isoforms (Pascual et al., 2006) (Swanson Lab, unpublished data). In order to assay for the alternative splicing patterns of *Mbnl3*, we amplified isoforms using RT-PCR from P1 and adult tissues with primers positioned in constitutively spliced non-coding exons 1 and 8 (5' and 3' untranslated regions, respectively). Subsequent amplicons were sub-cloned, sequenced, and annotated (Fig. 3-5). Analysis revealed previously unreported exons 7a, 7b, and cryptic 3' splice sites in exons 6 and 7 (exon 7 is referred to as exon 7c from here in). Exons 7a, 7b, and 7c are alternatively spliced producing unique Mbnl3 C-termini. Exon 2 is also alternatively spliced producing a hypothetical N-terminal truncation which could produce a protein with either four C_3H or two C_3H RNA binding motifs (Fig. 3-6). Although *Mbnl3* undergoes alternative splicing events which produce multiple isoforms, we sought to identify the relative use of these alternative exons in the mature mRNA. To examine *Mbnl3* alternative splicing, RT-PCR

with a forward primer positioned in constitutively spliced exon 3 and reverse primers specific to alternatively spliced exons 7a, 7b, and 7c were used (Fig. 3-7A). Embryonic and P1 brain, as well as placenta, included exons 7a and 7b to the greatest extent while exon 7c isoforms were the most abundant in all tissues examined (Fig. 3-7B). *Mbnl3* exon 2 assumes a default level of inclusion of approximately 30% (data not shown). Taken together with the sequencing data, *Mbnl3* exon 7c containing mRNAs are the predominant isoforms in all tissues while inclusion of exons 7a and 7b vary between tissues. This is particularly interesting considering *Mbnl1* utilizes its C-terminus for self-interactions, which suggests that Mbnl3 may interact with multiple proteins in different tissues.

Polyclonal Antisera Raised Against the C-terminus of Mbnl3

Mbnl3 contains an unusually large 3' UTR in the mature mRNA (~7800 nucleotides) which contains many predicted cis elements that have been shown to post-transcriptionally regulate mRNA fate (data not shown). Of course, *Mbnl3* mRNA levels may not accurately reflect the level of protein produced in the tissues assayed. Unfortunately, there is no antibody available that is cross reactive with murine Mbnl3 which is useful for tissue analysis. The Mbnl family of proteins also share a high degree of amino acid identity, which limits the unique regions available that may be targeted for antibody production. Therefore, we prepared polyclonal antibodies against the most immunogenic and unique region of the protein, the C-terminus (Fig. 3-7A). The majority of Mbnl3 isoforms should be recognized by this antibody (Table 3-1). A 15 amino acid peptide (NVPYVPTTTGNQLKY) was synthesized and conjugated to KLH, four rabbits were immunized, and bleeds were taken to assay for Mbnl3 reactivity. Placenta whole cell lysate (E15) was immunoblotted with either Mbnl3 antisera or pre-immune sera from

immunized rabbits (Fig. 3-8A). Antisera from rabbits A and C recognized two distinct bands at approximately 37 kDa and 27 kDa that were not seen in the pre-immune sera. Further evaluation revealed Mbnl3 anti-sera A and C recognized these bands with a similar affinity when E15 placenta lysate concentration was varied (Fig. 3-8B). Bleeds A and C were chosen for affinity purification and subsequent studies. Mbnl3 polyclonal antibody C (α-Mbnl3) will be used for the remainder of this study.

Due to the high degree of identity shared between the Mbnl proteins, it is important to exclude the possibility that α-Mbnl3 cross reacts with other family members. An α-Mbnl3 immunoblot shows that antisera reacts with exogenous myc-tagged Mbnl3 in COSM6 cells, but not Mbnl1 or Mbnl2 (Fig. 3-9A). Mbnl3 was also efficiently immunoprecipitated (Fig. 3-9B) as well as visualized by immunocytochemistry (Fig. 3-10) in exogenously expressing COSM6 cells. To ensure that α-Mbnl3 is specifically recognizing Mbnl3 in placenta, protein was immunoprecipitated from placenta whole cell lysate using α-Mbnl3 and subjected to MALDI-TOF mass spectrometry, which correctly identified Mbnl3 peptides (data not shown). Therefore, we have developed a new polyclonal antibody that recognizes Mbnl3 and which performs well for a wide range of applications.

Mbnl3 Isoforms Localize to both the Nucleus and Cytoplasm

Mbnl3 has been previously reported to be expressed in the proliferating myoblast cell line C2C12 and is down regulated in response to myogenic differentiation (Lee et al., 2007; Squillace et al., 2002). This expression pattern suggests an important role for *Mbnl3* during myoblast proliferation which must be down regulated in response to external stimuli to promote proper myotube maturation. Using our α-Mbnl3 antibody, we

sought to confirm *Mbnl3* expression in proliferating C2C12 cells. Interestingly, we identified two Mbnl3 proteins at 37 kDa and 27 kDa corresponding to the predicted Mbnl3 isoforms encoded by mRNAs (Table 3-1 and data not shown) that either included or excluded exon 2. Expression of both isoforms are inhibited by *Mbnl3* siRNA, but not by non-specific siRNA (Fig. 3-11). This is the first evidence of an N-terminal truncation producing a lower isoform in the Mbnl family of proteins. Moreover, a human hepatocarcinoma cell line, Huh7, exclusively expresses the MBNL3 lower isoform at 29 kDa (Fig. 3-11). siRNA directed against *MBNL3* efficiently knocks down MBNL3 expression while *SMN* siRNA does not affect expression. The human MBNL3 isoform is 2 kDa larger due to the inclusion of a human-specific exon 8 that is not conserved in mice (data not shown). The appearance of distinct Mbnl3 isoforms suggests that Mbnl3 possesses multiple functions.

Examination of MBNL3 localization using immunocytochemistry reveals discrete cytoplasmic foci in Huh7 cells, while C2C12 displays both nuclear and cytoplasmic foci (Fig. 3-12). This result is in direct contrast with Mbnl1 and Mbnl2, which appear predominately nuclear (Fardaei et al., 2002) (data not shown). Cytoplasmic foci in Huh7 and C2C12 cells do not localize with known structures, including P-bodies and other RNA binding proteins (Fig. 3-12 and data not shown). In support of the previous observation, lysates from nuclear and cytoplasmic fractionations of Huh7 and C2C12 cells were immunoblotted to assay for cellular localization. Interestingly, the MBNL3 isoform excluding exon 2 localized exclusively to the cytoplasm in both Huh7 and C2C12 cells (Fig. 3-13A,C). The Mbnl3 isoform including exon 2 localized to both the nucleus and cytoplasm in C2C12 cells. CUGBP1 and lactate dehydrogenase A (LDHA)

were used to assay for nuclear and cytoplasmic fractionation, respectively, while MBNL1 served as a family control. We next sought to determine if Mbnl3 localization *in vitro* accurately reflected it localization *in vivo*. Considering *Mbnl3* is expressed in a proliferating myoblast cell line, we chose to assay for Mbnl3 localization during embryonic myogenesis. Localization of Mbnl3 isoforms from E15 forelimbs was remarkably similar to C2C12 (Fig. 3-13B). Unlike Mbnl1, which is predominately a nuclear RNA binding protein involved in alternative splicing, Mbnl3 encodes several isoforms whose variable cellular distribution implies a non-overlapping function with other family members.

Mbnl3 Proteins are Primarily Expressed during Embryogenesis

To be a viable candidate for sequestration according to the RNA dominance model, Mbnl3 protein expression must overlap with DMPK expression. To address this question, brain, skeletal muscle/limb, tongue, and lung tissues were isolated from multiple time points during mouse development and immunoblotted for *Mbnl3* expression. Mbnl3 protein isoforms were readily detectable in E15-E18 placenta. Forelimb and tongue displayed more moderate levels at E15 and were undetectable by E18, while lung showed expression at E18 which persisted until P1 (Fig. 3-14). *Mbnl3* expression was also present in forelimb buds as early as E11.5 (data not shown). The single band recognized by α-Mbnl3 in adult brain was immunoprecipitated, subjected to MALDI-TOF mass spectrometry and determined to be glutamine synthetase (Fig. 3-14 and data not shown). Since Mbnl3 protein is not detectable in the adult brain, this cross-reaction did not affect our analysis. Interestingly, *Mbnl3* expression in the forelimb and tongue closely overlaps with embryonic myogenesis (Fig. 3-15) and is

down regulated after completion of secondary myogenesis, indicating a potential role for *Mbnl3* during skeletal muscle development.

If Mbnl3 is involved in the embryonic development of skeletal muscle, Mbnl3 might also be important for adult myogenesis (i.e. skeletal muscle regeneration). To test this hypothesis, we induced regeneration *in vivo* by injecting mouse TA muscles with Notexin, a venom peptide isolated from the Australian viper Notechis scutatus. Notexin is a 119 amino acid peptide with phospholipase A2 activity which, when injected into skeletal muscle, has potent myotoxic affects and promotes muscle necrosis, satellite cell activation and muscle regeneration (Dixon and Harris, 1996; Harris and MacDonell, 1981). Notexin was injected into the TA muscles of C57BL6/J mice (10-12 weeks of age), where they were isolated at two day intervals to assay for regeneration (Fig. 3-16A,B). Hematoxylin and Eosin (H&E) staining of 10 μm transverse cryosections was used to track skeletal muscle regeneration (Fig. 3-16C). Surprisingly, Mbnl3 demonstrated a sharp expression peak at Day 3 post-injection simultaneously with a gene involved in myoblast fusion, myogenin (Fig. 3-17). *Mbnl3* expression at this time point during regeneration is consistent with proliferating myoblasts (Fig. 3-18), but does not rule out the contribution from other cell types. However, this pattern of expression is similar to embryonic forelimb/tongue and C2C12 expression, suggesting that Mbnl3 may play an essential role during myogenesis.

Loss of Mbnl3 Inhibits Myogenic Differentiation in a C2C12 Model

Congenital myotonic dystrophy patients present with immature myotubes perinatally (Farkas-Bargeton et al., 1988; Iannaccone et al., 1986; Sarnat and Silbert, 1976). If *Mbnl3* plays an important role during myogenesis, then conceivably loss of

Mbnl3 would potentially inhibit this process. Using the C2C12 myogenesis model, in which immortalized proliferating myoblasts can be induced to differentiate upon the withdrawal of growth factors, we knocked down *Mbnl3* expression with siRNA and assayed for a delay in myogenesis (Fig. 3-19). In a control differentiation, *Mbnl3* expression is down regulated during the time course of myogenic differentiation at the same period when myotubes and myosin heavy chain (Mhc – a terminal differentiation marker) appear. However, when *Mbnl3*, expression is inhibited using siRNA, there is a delay in Mhc expression and fewer myotubes at corresponding time points during differentiation (data not shown). Therefore, inhibition of *Mbnl3* is sufficient to delay myogenesis *in vitro*; reaffirming that *MBNL3* is a viable candidate to contribute to a delayed myogenesis phenotype seen in CDM infants.

Targeting *Mbnl3* to Generate a Conditional Allele in Embryonic Stem Cells

If loss of MBNL3 function by sequestration contributes to the onset of CDM, then removing Mbnl3 *in vivo* would model the onset of disease. To eliminate those *Mbnl3* isoforms that interact with toxic $(CUG)_n$ RNAs (Fardaei et al., 2002), we focused our attention on *Mbnl3* exon 2. This exon encodes the first translational initiation codon which is responsible for producing the full length four C_3H motif isoform (Fig. 3-6), which has been previously implicated in disease. The *Mbnl3* exon 2 was replaced with a *Mbnl3* exon 2 flanked by loxP sites in C57BL6 embryonic stem cells (ESC) using standard targeting techniques (Fig. 3-20A). LoxP sites were used to eliminate intervening DNA sequence *in vivo* by expressing *Cre*, a site specific DNA recombinase. This targeting strategy allows us to remove *Mbnl3* exon 2 in tissues and time points of interest by crossing $Mbnl3^{cond/+}$ mice to mice expressing *Cre* under a gene-specific promoter. Probes located outside the arms of homology were used to identify correctly

targeted ESC clones. Out of 120 clones screened, 13 were positive (10.8%) for recombination at the 3' end and 5 were positive (4.2%) for recombination at the 5' end (Fig. 3-20B). Less efficient 5' targeting was due to recombination of the conditional exon (instead of the 5' arm of homology), excluding the 5' loxP site from the allele. Considering *Mbnl3* is an X-linked gene targeted in male ESCs, only one allele is detectable by southern analysis.

To verify successful targeting of *Mbnl3*, *Cre* was ectopically expressed (Fig. 3-21A) in targeted ESCs and assayed for loss of exon 2. The reduction in size of the *Mbnl3* KpnI fragment by southern blot analysis coincides with removal of exon 2 from the locus (Fig 3-21B). RT-PCR analysis with a forward primer in *Mbnl3* exon 1 and reverse primer in exon 8 demonstrate the loss of exon 2 containing transcripts in $Mbnl3^{\Delta E2/Y}$ ESCs, while $Mbnl3^{cond/Y}$ ESCs maintain exon 2 splicing (exon 2 inclusion = 1235 bp; exon 2 exclusion = 988 bp). Primary neomycin-resistant (Neo^R) mouse embryonic fibroblasts (MEFs) serve as a positive control for *Mbnl3* splicing. One concern about leaving the Neo^R cassette in the conditional *Mbnl3* locus is the possibility that it is included in the mature transcript by utilizing cryptic 3' splice sites (Nagy et al., 1998; Ren et al., 2002). This scenario could result in hypomorphic *Mbnl3* expression instead of a conditional allele designed to express at wild type levels. However, RT-PCR did not detect any cryptic splicing products into the Neo^R cassette (Fig. 3-21C). $Mbnl3^{cond/Y}$ ESCs were injected in C57BL/6J-Tyr^{c-2J} blastocysts to obtain chimeric mice. Male chimeras were bred to female C57BL6/J mice to attain germline transmission; the resulting F1 mice were C57BL/6J congenic $Mbnl3^{cond/+}$ females (data not shown).

Mbnl3$^{\Delta E2/Y}$ Mice Fail to Express Mbnl3 4XC$_3$H Isoforms

To test the hypothesis that loss of Mbnl3 full length isoforms contribute to CDM phenotypes, *Mbnl3*$^{cond/+}$ female mice were crossed to male B6.C-Tg(*CMV-cre)*1Cgn/J mice to remove *Mbnl3* exon 2 in the germline (Schwenk et al., 1995; Utomo et al., 1999). The resulting *Mbnl3*$^{\Delta E2/+}$ females were crossed to wild type males to obtain *Mbnl3*$^{\Delta E2/Y}$ males. *Mbnl3*$^{\Delta E2/Y}$ males were used for analysis for the remainder of this study. Because *Mbnl3* is primarily expressed during embryogenesis, it is possible that the loss of *Mbnl3* exon 2 isoforms during this stage of development result in an embryonic lethal phenotype. However, the numbers of wild type and *Mbnl3*$^{\Delta E2/Y}$ males (wild type, n=36 and *Mbnl3*$^{\Delta E2/Y}$, n=29) genotyped suggests that *Mbnl3*$^{\Delta E2/Y}$ has not deviated from the expected Mendelian ratio of 1:1 (wild type male:*Mbnl3*$^{\Delta E2/Y}$ male). Therefore, loss of *Mbnl3* exon 2 isoforms do not result in an embryonic lethality.

To confirm that *Mbnl3*$^{\Delta E2/Y}$ male mice no longer express full length Mbnl3 37 kDa isoforms, E15 forelimb and P1 lung were examined. RT-PCR comparing *Mbnl3* transcripts in wild type and *Mbnl3*$^{\Delta E2/Y}$ males indicate a loss of exon 2 transcripts, but an increase in exon 2 excluded transcripts (Fig. 3-22A). This is likely reflective of the steady state levels of total *Mbnl3* mRNA in the absence of *Mbnl3* exon 2 transcripts. Moreover, *Mbnl3*$^{\Delta E2/Y}$ males show an increase in the expression of the Mbnl3 27 kDa isoform by immunoblot analysis (Fig. 3-22B). P1 lung also shows and increase in the Mbnl3 27 kDa isoform. However, this increase does not correlate with the total protein in the wild type P1 lung. This suggests that the expression of the Mbnl3 27 kDa isoform is regulated post-transcriptionally and independently of the Mbnl3 37 kDa isoform.

Therefore, *Mbnl3*$^{\Delta E2/Y}$ mice do not express Mbnl3 37 kDa isoforms and upregulate 27 kDa isoforms during development.

Mbnl3$^{\Delta E2/Y}$ Mice do not Recapitulate Cardinal Phenotypes of Congenital Myotonic Dystrophy

The clinical manifestations of CDM at birth include poor suckling, movement deficits (hypotonia), immature skeletal muscle, skeletal deformities in the extremities (talipes), pulmonary insufficiencies, and a failure to reach developmental milestones (Harper, 2001). However, *Mbnl3*$^{\Delta E2/Y}$ mice appear visibly normal at birth (Fig. 3-23A), including indistinguishable neonatal movements and milk spots by P2, indicating normal feeding. Talipes may result from delayed muscle development during embryogenesis. To determine if *Mbnl3*$^{\Delta E2/Y}$ mice displayed talipes at birth, skeletal preps were performed on P1 pups; wild type and *Mbnl3*$^{\Delta E2/Y}$ hind limbs and forelimbs were compared, but no differences in the skeletal structures of the limb were observed (Fig. 3-23B). CDM patients also present with immature skeletal muscle at birth, highlighted by smaller myofibers and centralized nuclei, indicating a delay in embryonic myogenesis (Farkas-Bargeton et al., 1988). To assay for immature muscle fibers, H&E staining of 7 μm transverse paraffin sections from *Mbnl3*$^{+/Y}$ and *Mbnl3*$^{\Delta E2/Y}$ mice forelimbs were compared. No significant differences in either centralized nuclei or myofiber size were noted (Fig. 3-24). Normal myofiber appearance was confirmed by wheat germ agglutinin, a fluorescently conjugated lectin that outlines muscle fibers by interacting with glycosylated proteins in the extracellular matrix (Shaper et al., 1973), and immunohistochemistry using a terminal skeletal muscle marker, muscle specific myosin heavy chain (Fig. 3-24). No differences were observed postnatally, as *Mbnl3*$^{\Delta E2/Y}$ mice to continue gain weight normally past sexual maturity (Fig. 3-25). The major cause of

mortality in infants with CDM is respiratory distress, most likely from a combination of immature diaphragm and intercostals muscles as well as poorly understood pulmonary insufficiencies (Campbell et al., 2004; Rutherford et al., 1989). However, no *Mbnl3*$^{\Delta E2/Y}$ mice died (n=65) postnatally, indicating adequate pulmonary function. Taken together, this data suggests that loss of Mbnl3 37 kDa isoforms (i.e. isoforms including exon 2) alone are not sufficient to reproduce the onset of CDM in a mouse model.

Loss of *Mbnl3* 4XC$_3$H Isoforms does not Inhibit Skeletal Muscle Regeneration in an Injury Model

Individuals with congenital myotonic dystrophy that survive past childhood go on to develop the adult onset symptoms of myotonic dystrophy. One of the key characteristics of the adult-onset neuromuscular disorder is the progressive wasting of skeletal muscle. Although mouse models have successfully recapitulated the majority of highly penetrate disease characteristics, they have failed to address the cause of muscle wasting in adult-onset DM. If *Mbnl3* expression in activated satellite cells is required for proper maintenance or repair of skeletal muscle following injury, one interesting possibility is that loss of Mbnl3 due to sequestration by toxic C(CUG)$_n$ repeats could inhibit repair/regeneration of muscle. To test this hypothesis, adult skeletal muscle was induced to regenerate by injecting the TA muscle of wild type and *Mbnl3*$^{\Delta E2/Y}$ mice (10-12 weeks of age) with Notexin; TA muscle was pulled at two day intervals to assay for inhibition of regeneration (Fig. 3-26A,B). RT-PCR and immunoblot analysis confirmed the loss of Mbnl3 37 kDa isoform expression during regeneration (Fig. 3-26C,D). Expression levels of myogenin, a transcription factor involved in the commitment and fusion of proliferating myoblasts (Buckingham et al., 2003), was not significantly altered during regeneration in *Mbnl3*$^{\Delta E2/Y}$ mice compared to wild type

controls. *Mbnl1* RNA levels at Day 3 post injection in *Mbnl3*$^{\Delta E2/Y}$ mice were slightly lower, but protein levels remained unaffected (Fig. 3-26C,D). Interestingly, Mbnl3 27 kDa isoform was not upregulated at Day 3 post Notexin injection in *Mbnl3*$^{\Delta E2/Y}$ mice. H&E staining of 10 μm transverse cryosections to assay for regeneration did not show a reduction in the regenerative capacity of *Mbnl3*$^{\Delta E2/Y}$ mice (Fig. 3-27A). Wheat germ agglutinin, a lectin that outlines myofibers by interacting with glycosylated proteins of the extracellular matrix and plasma membrane, does not show any difference between wild type and *Mbnl3*$^{\Delta E2/Y}$ mice during regeneration (Fig. 3-27B). Desmin, a marker of terminal skeletal muscle differentiation, is unaffected during regeneration (Fig. 3-28). This data suggests that loss of Mbnl3 37 kDa isoforms (i.e. isoforms including exon 2) alone do not inhibit the regeneration of adult skeletal muscle and are not sufficient to cause a wasting phenotype.

Discussion

Mbnl is a family of highly conserved genes that are implicated in developmental maturation of tissues through the regulation of RNA metabolism (Pascual et al., 2006). Most notably, MBNL1 promotes a change in alternative splicing patterns of specific genes during the fetal to adult transition in terminally differentiating tissues, which results in protein isoforms that optimally support that tissues physiological needs (Lin et al., 2006). When MBNL1 function is inhibited by toxic C(CUG)$_n$ RNA expression in DM, impairment of this splicing transition results in the misregulation of alternative splicing and disease. Remarkably, the *Mbnl1*$^{\Delta E3/\Delta E3}$ mouse, which mimics loss of MBNL1 protein in DM by removing the exon responsible for initiating translation of full length Mbnl1 protein, recapitulates the majority of highly penetrant DM phenotypes (Kanadia et

al., 2003a). This observation is even more extraordinary considering that toxic C(CUG)$_n$ repeats likely contribute to disease pathogenesis by compromising not only MBNL1, but also other family members, altering downstream phosphorylation of PKC target CUGBP1, as well as altering expression of genes linked to the *DMPK* locus (Fu et al., 1993; Inukai et al., 2000; Kuyumcu-Martinez et al., 2007; Novelli et al., 1993; Sabouri et al., 1993). Although the *Mbnl1*$^{\Delta E3/\Delta E3}$ mouse has provided important insights into adult-onset DM and the molecular pathways affected, many questions remain. For example, what causes deficits in the central nervous system and progressive wasting of skeletal muscle in adults, and what molecular events lead to CDM? Modeling these elusive phenotypes will prove critical to determining the underlying mechanism of disease which should allow the development of effective therapies.

Mbnl3 is Expressed in Developing Tissues that are Affected in CDM

Previous studies have reinforced the idea that other MBNL family members are likely comprised in CDM/DM, including evidence that MBNL1, MBNL2, and MBNL3 interact with expanded (CUG)$_n$ repeats both *in vitro* and *in vivo* (Fardaei et al., 2002; Holt et al., 2009). While *Mbnl1* and *Mbnl2* expression overlaps in many tissues throughout development, in situ analysis, tissue RT-PCR, and whole embyo RNA Northern blot analysis of *Mbnl3* has led to the prevailing view point that *Mbnl3* expression is limited to embryogenesis. However, this interpretation relies on the assumption that *Mbnl3* is not post-transcriptionally regulated and protein levels are reflective of mRNA levels. Our data, using a new α-Mbnl3 antibody, provides evidence for embryonic Mbnl3 expression in those tissues affected in CDM (tongue, lung, forelimb) as well as regenerating adult skeletal muscle. Mbnl3 protein is not detectable

in differentiated adult tissues. While we were unable to use the α-Mbnl3 antibody for immunohistochemistry and localization of Mbnl3 to specific cell types *in vivo*, it appears that *Mbnl3* is most likely expressed in myoblasts and not differentiated myofibers. *Mbnl3* is expressed in proliferating C2C12 cells and downregulated upon differentiation. *Mbnl3* mRNA is also expressed in limb buds by E9.5 (Kanadia et al., 2003b) and protein is present in developing forelimb from E11.5-E16.5 until the completion of secondary myogenesis (Fig. 3-14 and 3-15 and data not shown) and during adult skeletal muscle regeneration until cell cycle arrest and terminal differentiation (Fig. 3-17 and 3-18). An interesting possibility is that Mbnl3 is important for providing alternative splicing, or another mRNA regulating event, that is necessary for the embryonic developmental program, similar to Mbnl1 is in adult tissues.

Mbnl3 alternative splicing produces multiple isoforms varying at both the N- and C-termini. This is not uncommon in the *Mbnl* family, as *Mbnl1* and *Mbnl2* both undergo alternative splicing of their C-termini which can modulate localization and self-interaction (Swanson Lab unpublished data). Like Mbnl1 and Mbnl2, Mbnl3 may utilize alternative splicing at its C-terminus to modulate protein-protein interactions and function. However, the Mbnl3 N-terminal truncated isoform which contains $2XC_3H$ motifs (Fig. 3-2 and Fig. 3-6) is unique among family members. *Mbnl1* also produces mRNAs that lack exon 3, which contains the *Mbnl1* translation initiation codon, but no truncated proteins have been detected in tissues (Swanson lab unpublished data). N-terminal truncated isoforms with $2XC_3H$ motifs may represent an Mbnl3 protein that interacts with a different set of RNAs and plays a different role than Mbnl3 full length $4XC_3H$ protein. Temporal and spatial expression, as well as localization differences between Mbnl3

$4XC_3H$ and $2XC_3H$ isoforms further suggest a divergent role for these proteins during development.

Mbnl3 is Required for C2C12 Differentiation

While it has been established that *Mbnl3* is expressed in proliferating C2C12 cells (Fig. 3-11), previous studies designed to determine the role of Mbnl3 during myogenesis have provided counterintuitive results. Neonates with CMD present with immature myotubes, which suggests a delay in myogenesis (Farkas-Bargeton et al., 1988). Previous studies where a myc-tagged $4XC_3H$ *Mbnl3* transgene was overexpressed in stably transfected C2C12 demonstrated a delay in differentiation as assayed by *Mhc* expression (Squillace et al., 2002). Moreover, a morpholino antisense oligonucleotide (designed against the 5' UTR/translation initiation codon) that inhibits translation of *Mbnl3* also promotes differentiation in C2C12 cells. If loss of Mbnl3 function by $(CUG)_n$ sequestration contributes to CDM, then the expected result of inhibiting translation in C2C12 would be to delay differentiation. However, there are critical controls absent from these experiments. First, the *Mbnl* family shares a high degree of amino acid identity and have a previously demonstrated redundant function in an *in vitro* splicing assay (Fig. 2-9). Therefore, overexpression of *Mbnl3* would potentially have downstream affects on not only its own targets, but on *Mbnl1* and *Mbnl2* as well. The lack of control C2C12 cell lines expressing either exogenous *Mbnl1* or *Mbnl2* prevent proper interpretation of this data. Second, the polyclonal antisera used for these studies was generated using the C-terminal 83 amino acids of Mbnl3. The immunizing peptide shares a greater than 60% identify with Mbnl1 and Mbnl2 which greatly increases the chances that this polyclonal antisera reacts against multiple members of the *Mbnl* family. The Mbnl3 protein also runs at a higher molecular weight by SDS-PAGE and

immunoblotting than anticipated, so there may be a cross-reaction with Mbnl1 or Mbnl2. No further antibody validation studies were performed. Third, the *Mbnl3* antisense morpholino experiment to knock down Mbnl3 4XC$_3$H expression was inefficient and was not verified in any other experiments. In addition, the *Mbnl3* morpholino antisense oligonucleotide is directed against only the full length Mbnl3 4XC$_3$H isoform and does not target the Mbnl3 2XC$_3$H isoform.

To clarify these results, we inhibited *Mbnl3* expression using siRNAs directed against the *Mbnl3* coding sequence/3' UTR, which target all isoforms, and assayed for differentiation of C2C12 cells. Knockdown of *Mbnl3* expression inhibited myogenic differentiation of C2C12 cells, which indicates that loss of Mbnl3 alone was adequate to delay myogenesis *in vitro*. This result is consistent with a delay in myogenesis seen in CDM patients. While it is not clear if the delay seen in C2C12 differentiation is sufficient to cause a delay in embryonic myogenesis *in vivo*, it is important to note that loss of *Mbnl3* expression in C2C12 does not promote premature differentiation as reported.

While previous studies have established interactions of the Mbnl3 4XC$_3$H isoform with (CUG)$_n$ repeat RNA, it is unclear whether the Mbnl3 2XC$_3$H isoform is sequestered. Therefore, the Mbnl3 2XC$_3$H isoform is a potential candidate in the RNA dominance model of CDM/DM. Interestingly, Drosophila Mbl proteins containing only 2XC$_3$H motifs localize with (CUG)$_n$ repeat RNAs *in vitro* and *in vivo*, suggesting the Mbnl3 N-terminal truncated isoform may be important for disease pathogenesis.

Loss of Mbnl3 4XC3H Isoforms are not Sufficient to Phenocopy CDM or Skeletal Muscle Wasting

The *MBNL* family was originally proposed to be involved in the pathogenesis of myotonic dystrophy by sequestration and loss of protein function based on its length

dependent interaction with expanded $(CUG)_n$ RNA *in vitro* and *in vivo* (Miller et al., 2000). *Mbnl3* demonstrates a spatially and temporally restricted expression pattern in tissues affected in CDM and DM. Knockdown of *Mbnl3* expression via siRNA in a C2C12 model of myogenic differentiation delayed myogenesis *in vitro*. Therefore, we tested the hypothesis that CDM and adult-onset muscle wasting is caused by sequestration of Mbnl3 $4XC_3H$ isoforms, which have been previously shown to interact with $(CUG)_n$ repeat RNA. Although all *Mbnl3* exon 2 containing isoforms where eliminated, *Mbnl3*$^{\Delta E2/Y}$ mice did not phenocopy the onset of CDM or demonstrate an impaired capacity for skeletal muscle regeneration. *Mbnl3*$^{\Delta E2/Y}$ mice live to ~6 months of age and appear phenotypically normal at the time of this study. While this data does not rule out a lack of a molecular phenotype in *Mbnl3*$^{\Delta E2/Y}$ mice, it does provide evidence against the idea that loss of function of Mbnl3 $4XC_3H$ isoforms alone by toxic RNA sequestration result in CDM or skeletal muscle wasting. What other factors could be potentially involved in this disease?

One possibility is that other proteins partially compensate for loss of Mbnl3 $4XC_3H$ isoforms *in vivo*, including Mbnl1, Mbnl2, and upregulated Mbnl3 $2XC_3H$ isoforms. Alternatively, multiple *Mbnl* pathways may need to be comprised to promote the onset of CDM or skeletal muscle wasting. As $(CUG)_n$ repeats expand to >1000, they may titrate out all Mbnl proteins, and thus compromise both independent and compensatory functions. If so, *Mbnl1*, *Mbnl2*, and *Mbnl3* conditional and knockout mice must be generated and crossed to test the hypothesis that sequestration of the *Mbnl* family contributes to CDM and skeletal muscle wasting. One argument against this hypothesis is that *MBNL* family members can interact with both $(CUG)_n$ and $(CCUG)_n$ repeat RNAs

in vitro, but only *DMPK* $(CUG)_{>1000}$ promote the onset of CDM (Fardaei et al., 2002). In fact, *DMPK* $(CUG)_n$ and *CNBP* $(CCUG)_n$ repeats can expand to >4,000 and 11,000, respectively (Ranum and Day, 2002). While *DMPK* is expressed at its highest levels in the CNS and skeletal muscle, *CNBP* is expressed ubiquitously and at higher levels than *DMPK* (Mankodi et al., 2003). This provides evidence for the idea that either $(CUG)_n$ repeat RNA is inherently more toxic than $(CCUG)_n$ or repeat expansions in the *DMPK* locus, alter neighboring gene expression and promote a locus-specific contribution to disease .

However, a combinatorial approach using genetic crosses which eliminate *Dmpk*, *Six5*, and *Mbnl1* exon 3 containing isoforms failed to phenocopy CDM in a mouse model (Swanson lab unpublished data). We cannot, however, rule out that another closely linked *Dmpk* gene, *Dmwd*, is involved in disease. We have also previously demonstrated that MBNL1:$(CUG)_n$ RNA form more stable complexes than MBNL1:$(CCUG)_n$ RNA in a competition assay. This implies that there is an inherent difference in the interactions between MBNL proteins and $(CUG)_n$/$(CCUG)_n$ RNAs. This observation may explain differences in disease manifestation between DM1 and DM2.

In conclusion, CDM and DM are complex diseases that affect nearly every organ of the body. Our results provide evidence against a singular role for Mbnl3 $4XC_3H$ isoforms in sequestration and onset of CDM and adult-onset skeletal muscle wasting. This study highlights the importance of developing a combinatorial genetic model of CDM to elucidate the molecular etiology of this disease. It will be intriguing to unfold the complex genetic contributions and pathways that underlie this disorder.

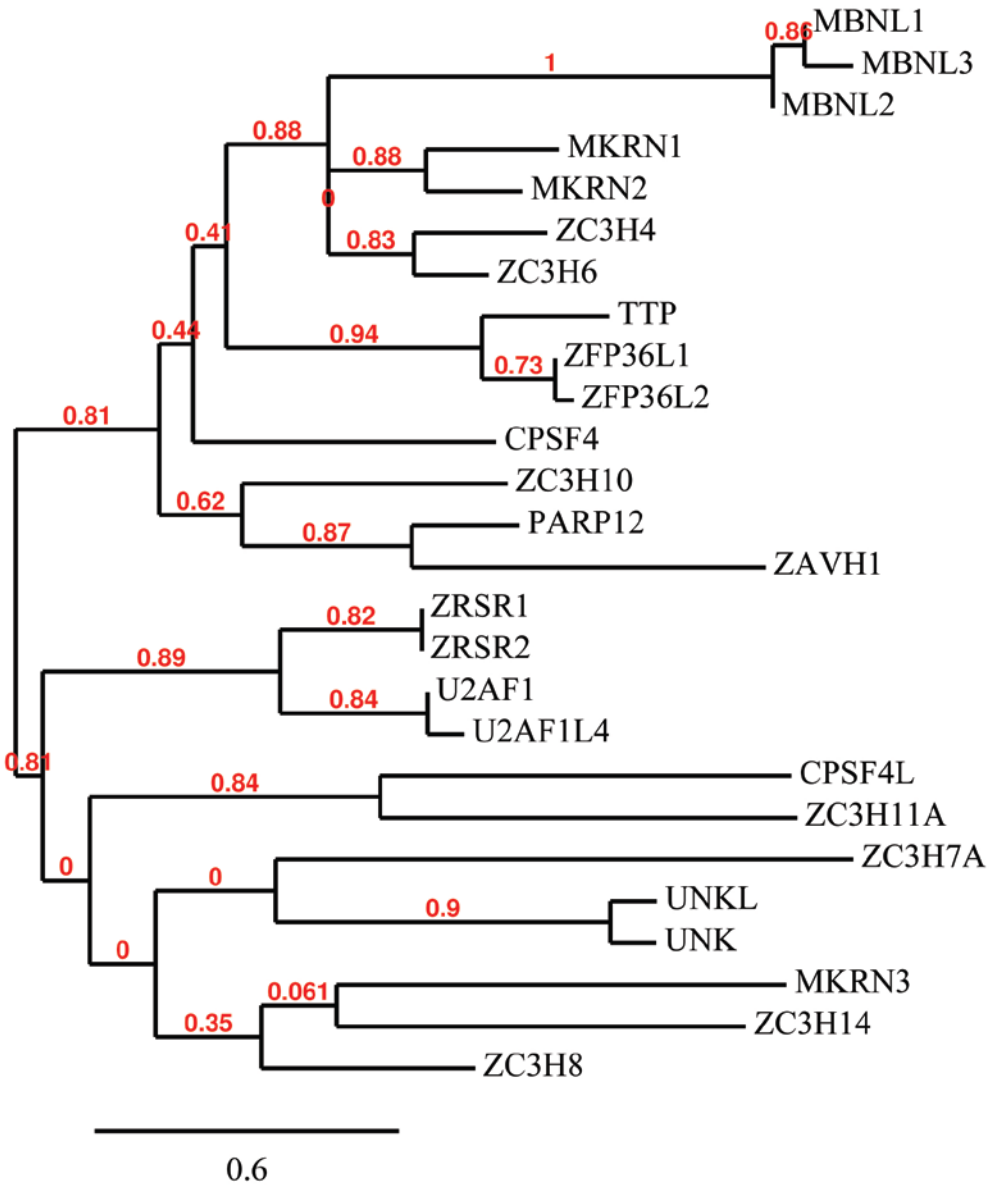

Figure 3-1. Phylogram shows evolutionary proximity of human C_3H zinc finger motif containing proteins. Protein sequences were derived from EMBL-EBI database and C_3H pairs were analyzed by multiple alignment using MUSCLE (multiple sequence comparison by log-expectation). Phylogenetic tree was created using PhyML (phylogeny analysis using maximum likelihood method). Bootstrap confidence levels (red) measure robustness of support for a given clade. (0.95 indicates a 95% reproducibility in analysis). MUSCLE and PhyML located at www.phylogeny.fr.

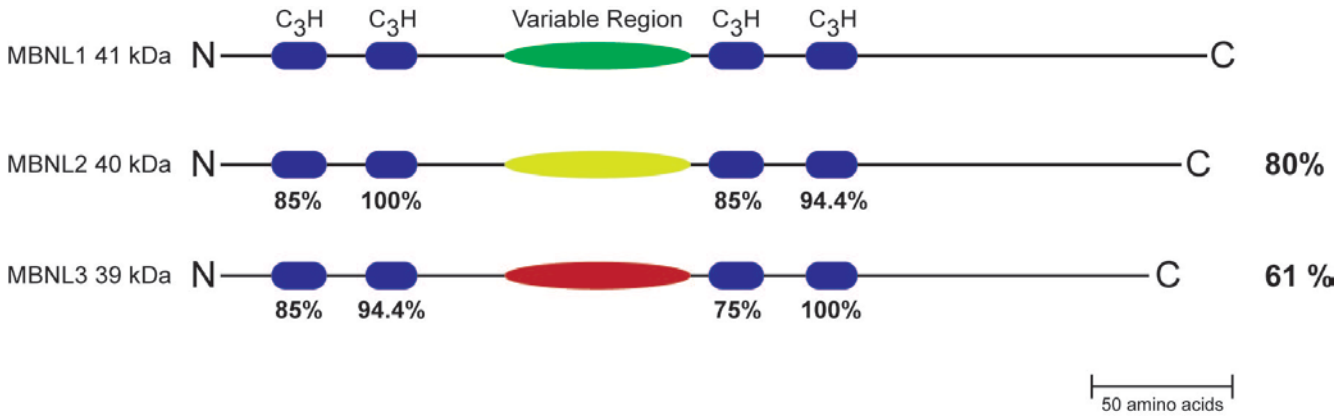

Figure 3-2. The MBNL family is composed of three closely related paralogs. MBNL proteins have conserved C_3H RNA binding domains (blue) clustered in two pairs. Percentages below the C_3H domains indicate conserved amino acid identity with MBNL1 C_3H. Percentages to the right of MBNL2 and MBNL3 indicate amino acid identity of the total protein when compared with MBNL1. The hyper-variable linker region (green, yellow, red) is the most divergent sequence among the proteins.

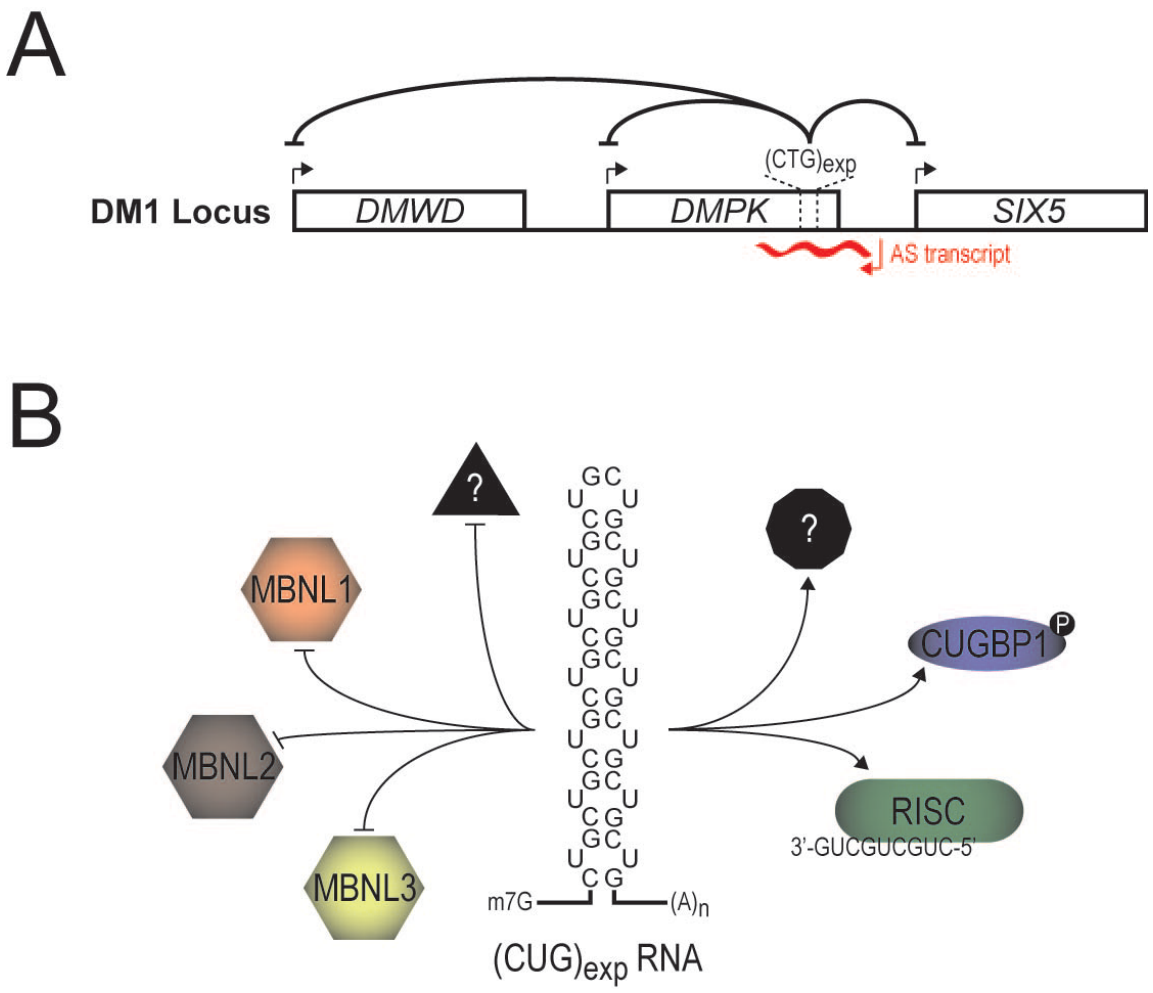

Figure 3-3. Schematic of potential genetic contributions to DM1. (A) *DMPK* $(CTG)_{exp}$ inhibits transcription of closely linked genes *DMWD* and *SIX5*. $(CTG)_{exp}$ can act as a strong nucleosomal positioning element, leading to the formation of heterochromatin at the DM1 locus. Antisense transcription through the repeat promotes further chromatin condensation by recruiting heterochromatin associated factors. (B) *DMPK* $(CUG)_{exp}$ can inhibit or activate downstream protein function. The MBNL family of proteins are sequestered by $(CUG)_{exp}$, impeding normal cellular functions. Additionally, increased CUGBP1 stability (through hyperphosphorylation) and activation of the RISC complex containing short CUG siRNAs may play roles in disease.

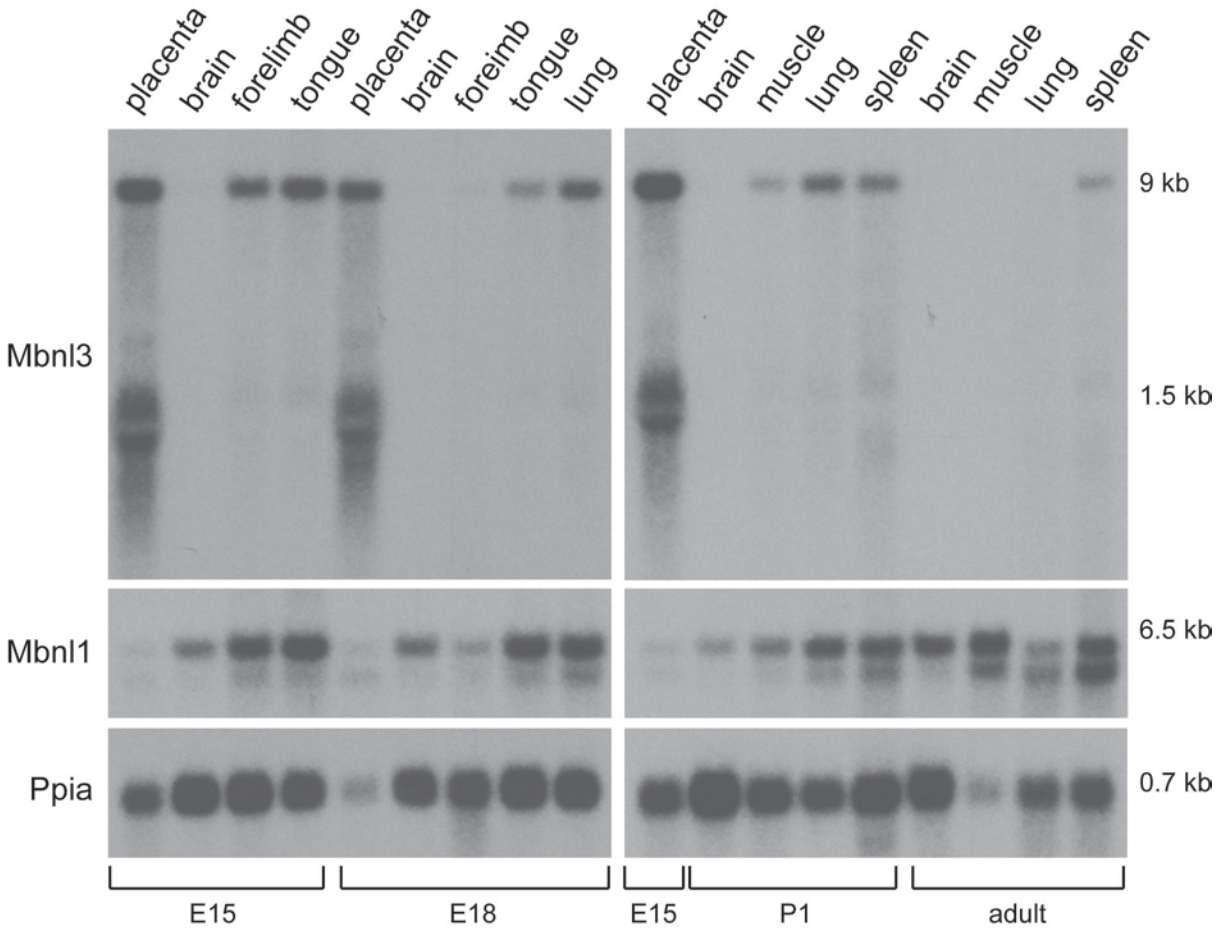

Figure 3-4. *Mbnl3* expression is restricted spatially and temporally. RNA blot analysis at different murine developmental time points (embryonic day 15/18 = E15/E18; postnatal day 1 = P1). Coding sequence probes are used to detect *Mbnl3*, *Mbnl1* (family control), and *Ppia* (loading control). *Mbnl3* expression in E15 placenta is used as an exposure control between blots. Expected sizes of mRNAs; *Mbnl3* (9 kb, 1.5 kb, 1.3 kb), *Mbnl1* (6.5 kb, 5.3 kb), and *Ppia* (0.7 kb).

cDNA	Band	Seq. Clone #	Isoform	Notes	Recognized by Mbnl3 pAb
P1 Brain	A - Upper	1	37 kDa		Yes
P1 Brain	A - Upper	2	37 kDa		Yes
P1 Brain	B - Lower	3	27 kDa (-E2)		Yes
P1 Brain	B - Lower	4	37 kDa	Topo Cloned Upper Band	Yes
P1 Lung	A - Upper	5	37 kDa		Yes
P1 Lung	A - Upper	6	37 kDa		Yes
P1 Lung	A - Upper	7	37 kDa		Yes
P1 Lung	A - Upper	8	37 kDa		Yes
P1 Lung	B - Lower	9	27 kDa (-E2)		Yes
P1 Lung	B - Lower	10	27 kDa (-E2)		Yes
P1 Lung	B - Lower	11	27 kDa (-E2)		Yes
P1 Lung	B - Lower	12	27 kDa (-E2)		Yes
P1 Spleen	A - Upper	13	27 kDa (-E2)	Topo Cloned Lower Band	Yes
P1 Spleen	A - Upper	14	37 kDa		Yes
P1 Spleen	A - Upper	15	37 kDa		Yes
P1 Spleen	A - Upper	16	37 kDa		Yes
P1 Spleen	B - Lower	17	27 kDa (-E2)		Yes
P1 Spleen	B - Lower	18	27 kDa (-E2/Cryptic E7 3' SS)		Yes
P1 Spleen	B - Lower	19	27 kDa (-E2)		Yes
P1 Spleen	B - Lower	20	27 kDa (-E2)		Yes
Adult Brain	A - Upper	21	37 kDa (-E7)	Shift in reading frame - New C-terminus	No
Adult Brain	A - Upper	22	37 kDa		Yes
Adult Brain	B - Lower	23	27 kDa (-E2/-E7/+E7a/+E7b)	Shift in reading frame - New Short C-terminus	No
Adult Brain	B - Lower	24	27 kDa (-E2/-E7/+E7a)	Shift in reading frame - New C-terminus	No
Adult Quad	A - Upper	25	37 kDa		Yes
Adult Quad	A - Upper	26	37 kDa		Yes
Adult Quad	B - Lower	27	27 kDa (-E2)		Yes
Adult Quad	B - Lower	28	27 kDa (-E2)		Yes
Adult Heart	A - Upper	29	37 kDa (-E7)	Shift in reading frame - New C-terminus	No
Adult Heart	A - Upper	30	37 kDa		Yes
Adult Heart	B - Lower	31	27 kDa (-E2)		Yes
Adult Heart	B - Lower	32	27 kDa (-E2)		Yes
WT MEFs	A - Upper	33	37 kDa		Yes
WT MEFs	A - Upper	34	37 kDa		Yes
WT MEFs	B - Lower	35	27 kDa (-E2)		Yes
WT MEFs	B - Lower	36	27 kDa (-E2)		Yes

Figure 3-5. The *Mbnl3* gene produces multiple isoforms through alternative splicing. *Mbnl3* isoforms were amplified from cDNA with primers flanking the coding sequence, subcloned, and sequenced. Tissues used, band excised for cloning, and annotation are included.

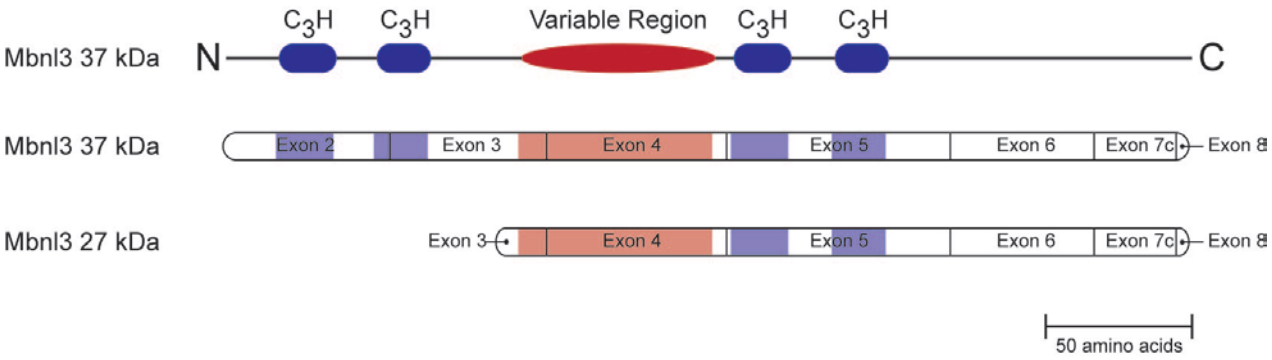

Figure 3-6. *Mbnl3* is predicted to encode several different isoforms. Mbnl3 contains two pairs of RNA binding motifs (C_3H) encoded by exons 2/3 and 5, as well as a linker region which is highly variable between family members. Exclusion of exon 2 in mature transcripts is predicted to encode a Mbnl3 N-terminal trucated isoform with one pair of C_3H binding motifs which utilizes an alternative initiation codon in exon 3.

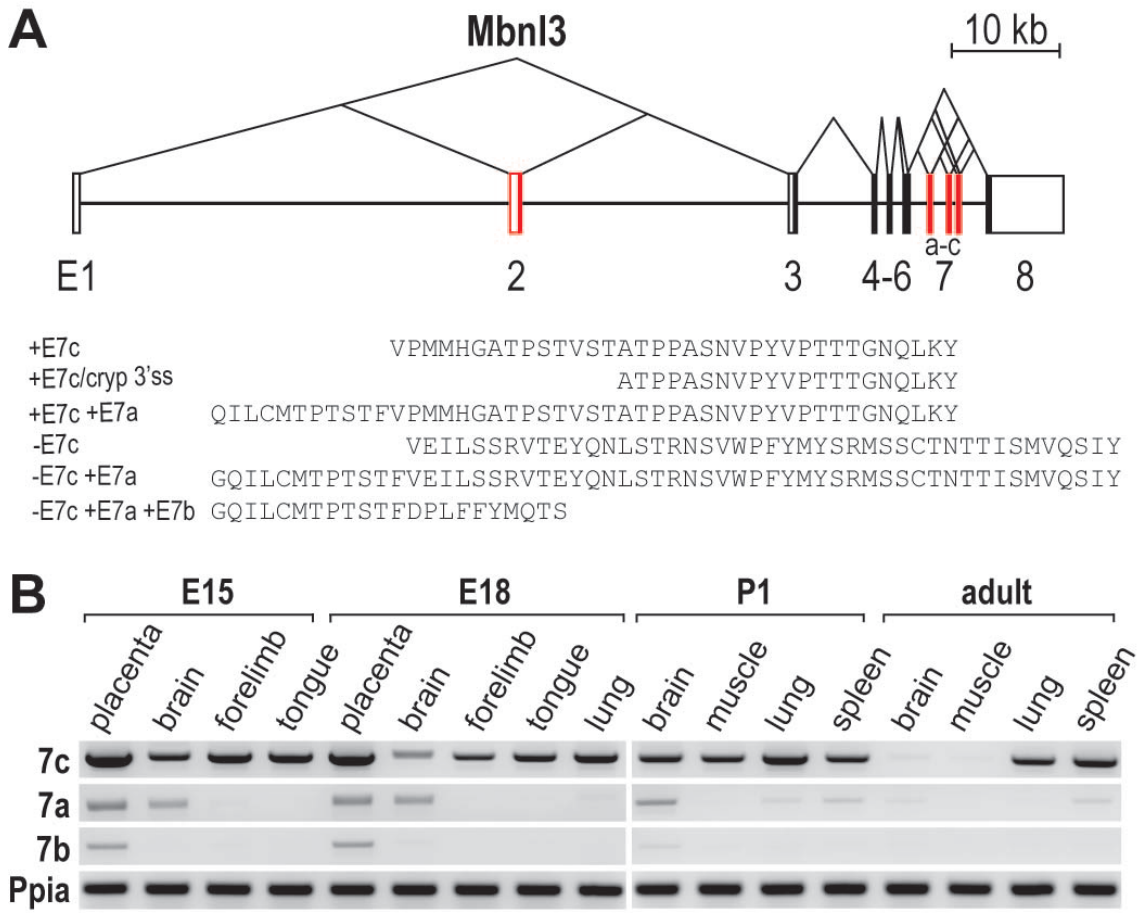

Figure 3-7. Alternative splicing of *Mbnl3* produces multiple isoforms throughout development. (A) The *Mbnl3* locus. The *Mbnl3* gene consists of 10 exons (boxes) and 9 introns (horizontal line), constitutive (black) and alternatively spliced exons (red), untranslated regions (open boxes) and coding regions (closed boxes). Alternative splicing of *Mbnl3* exons 7a, 7b, and 7c produces a variable C-terminus. (B) Exon 7c containing transcripts are predominant in all tissues examined. RT-PCR of *Mbnl3* at different murine developmental time points with a forward primer in constitutively spliced exon 3 and reverse primers in alternatively spliced exons 7a, 7b, and 7c.

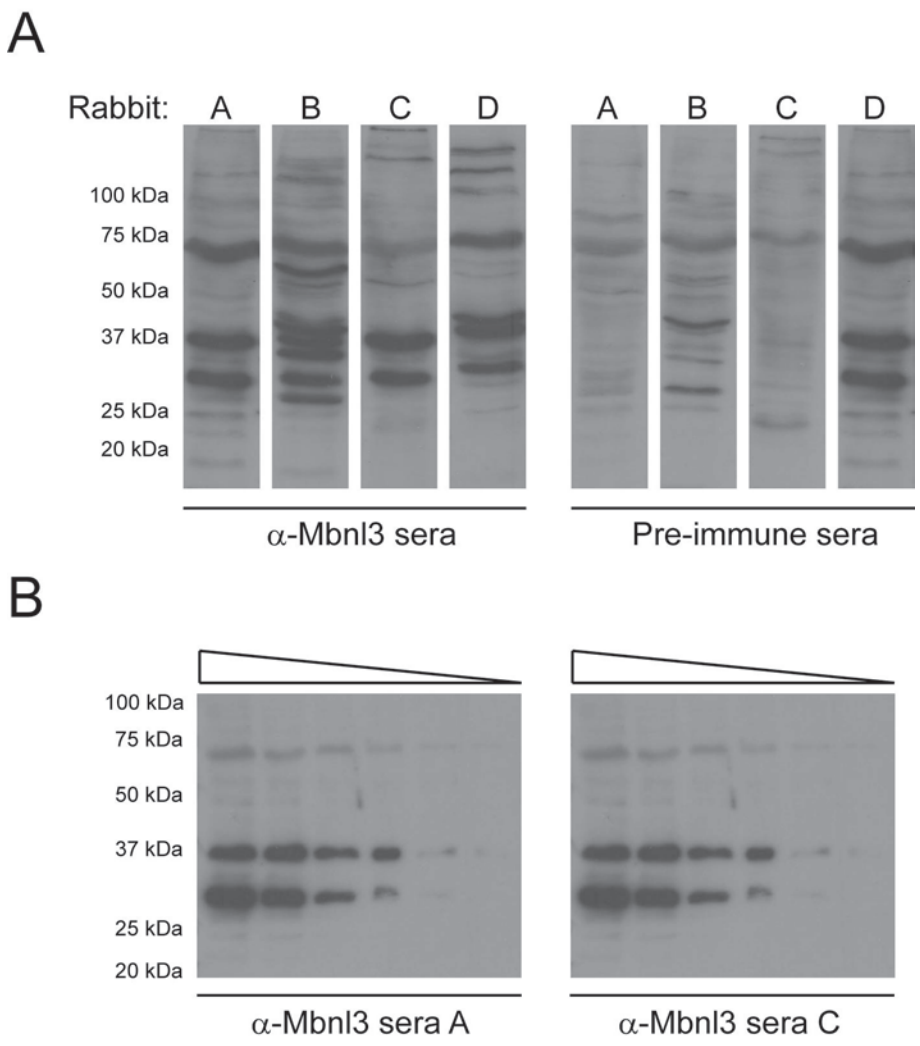

Figure 3-8. Mbnl3 antisera recognizes two distinct bands in E15 placenta whole cell lysate. (A) Immunoblot analysis using sera from four rabbits (A-D) immunized against Mbnl3 was used at a 1:500 dilution to screen for immunoreactivity in E15 placenta whole cell lysate. Pre-immune sera is used as a control. Antisera from rabbits A and C were selected for further analysis. (B) Titration of the input E15 placenta lysate is used to compare affinity between the reactive anti-sera A and C at a dilution of 1:1000.

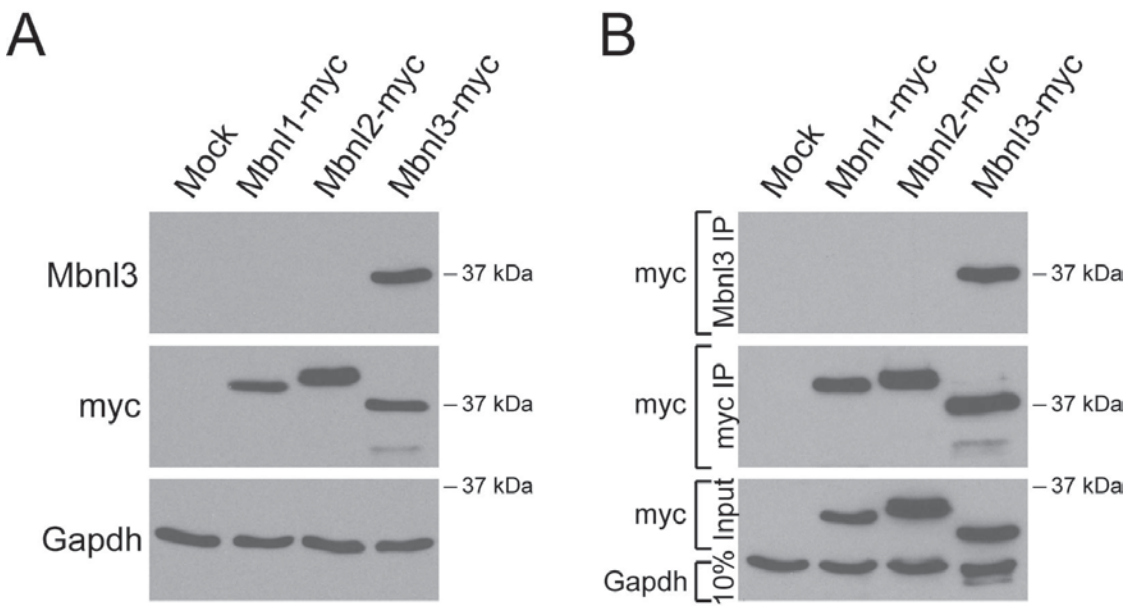

Figure 3-9. α-Mbnl3 antisera does not cross-react with other Mbnl proteins. (A) α-Mbnl3 recognizes Mbnl3, but not Mbnl1 or Mbnl2, by immunoblotting. Myc-tagged Mbnl proteins were exogenously expressed in COSM6 cells and immunoblotted with α-Mbnl3 (experimental), α-myc (expression control) or α-Gapdh (loading control). (B) α-Mbnl3 antibody immunoprecipitates (IP) Mbnl3, but not Mbnl1 or Mbnl2. Myc-tagged Mbnl proteins were exogenously expressed in COSM6 cells, immunoprecipitated with either α-Mbnl3 (experimental IP) or α-myc (control IP) antibodies, and immunoblotted with α-myc. As controls, 10% input was immunoblotted with α-myc (expression control) or α-Gapdh (loading control).

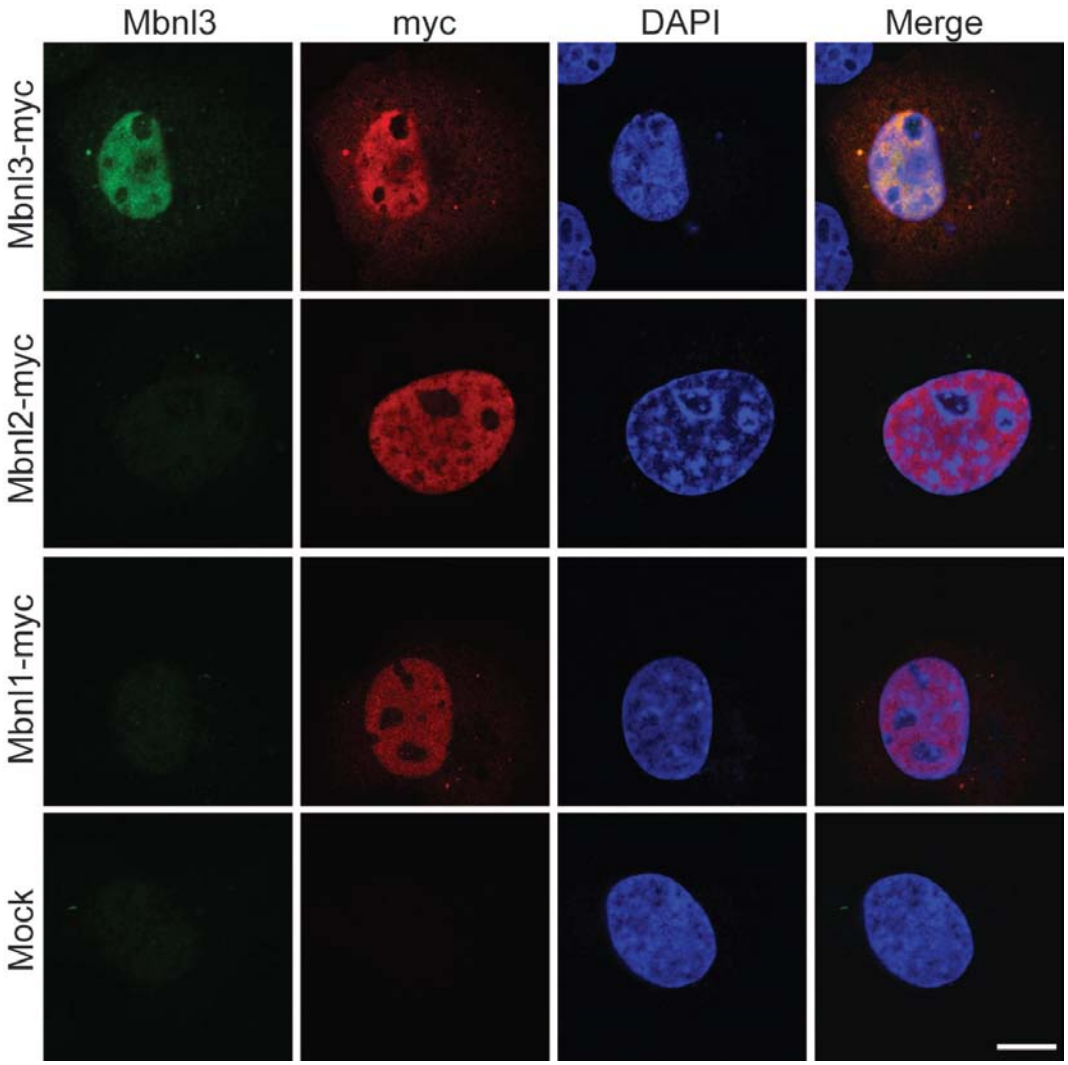

Figure 3-10. α-Mbnl3 distnguishes between Mbnl family members by immunocytochemistry. α-Mbnl3 specifically recognizes Mbnl3 by immunocytochemistry, but not Mbnl1 or Mbnl2. Myc-tagged Mbnl proteins were exogenously expressed in COSM6 cells and visualized with α-Mbnl3 (green) or α-myc (red). DAPI (blue) stains DNA and indicates nuclear location. Scale bar = 5 μm.

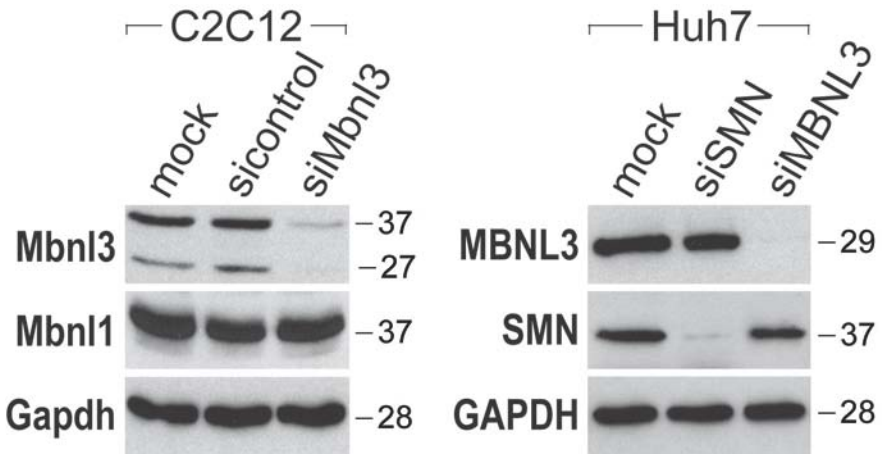

Figure 3-11. Two distinct Mbnl3 isoforms are expressed in C2C12 and Huh7 cells. Immunoblot analysis of C2C12 and Huh7 cells with α-Mbnl3 showing that siRNA against Mbnl3 (C2C12) or human MBNL3 (Huh7) specifically knocksdown expression of MBNL3, while non-specific siRNA has no effect. α-Mbnl1 and α-SMN are used to assay for efficiency of MBNL3 siRNA and mediated knockdowns. α-GAPDH is used as a loading control.

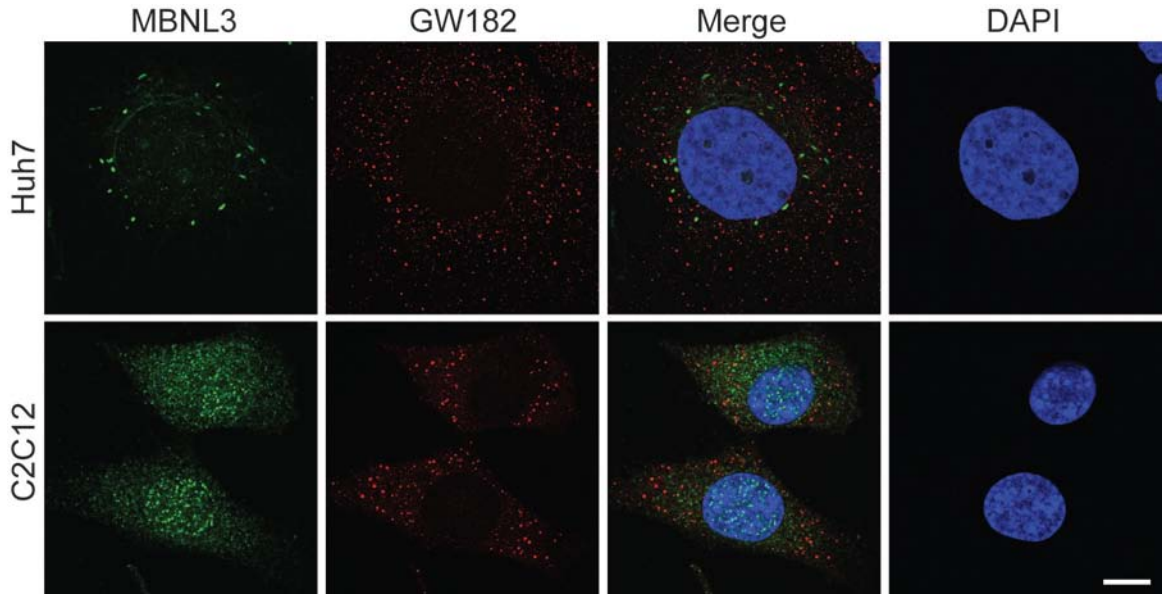

Figure 3-12. MBNL3 localizes to nuclear and cytoplasmic foci. Human MBNL3 (green) is localized to cytoplasmic foci in Huh7 cells and these foci do not colocalize with known cytoplasmic structures such as P-bodies, stained with α-GW182 (red). Mbnl3 (green) is distributed in both the nucleus and cytoplasm in C2C12 cells, but does not colocalize with P-bodies (red). DAPI (blue) stains DNA and indicates nuclear location. Scale bar = 5 μm.

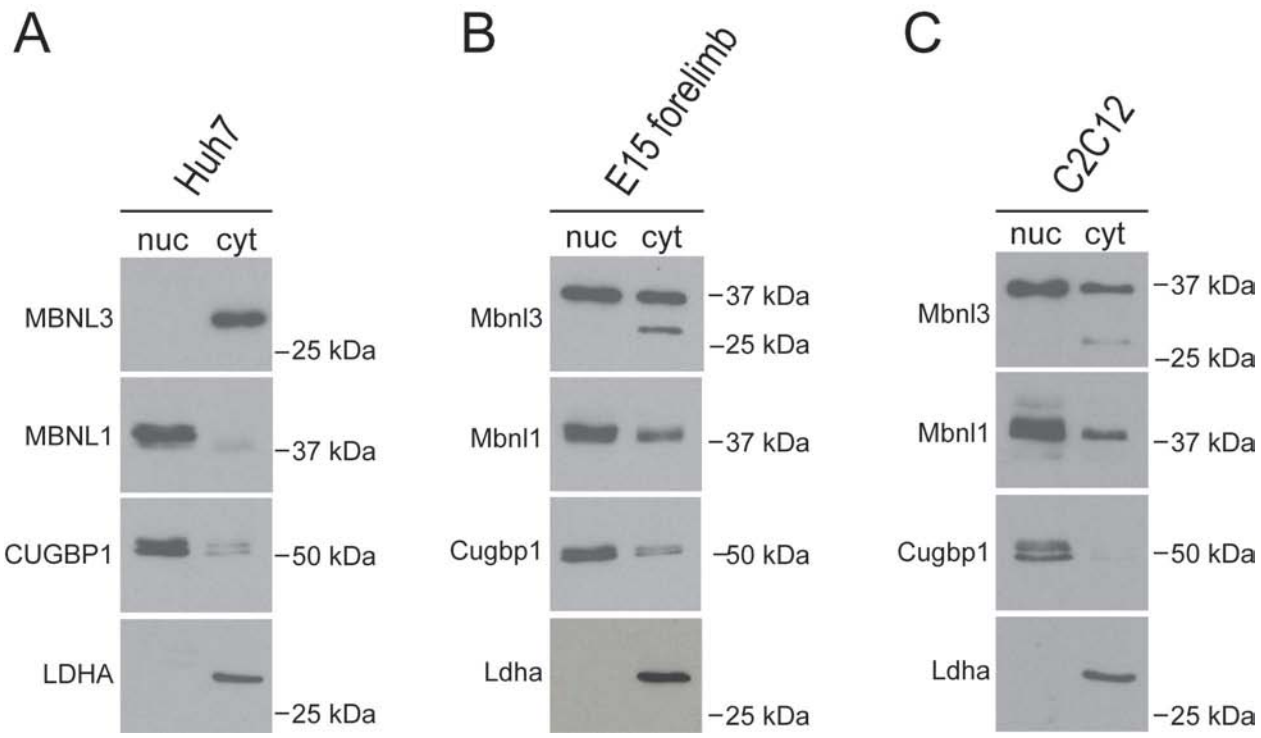

Figure 3-13. MBNL3 isoforms are found in different subcellular compartments. Immunoblot analysis of nuclear/cytoplasmic fractionations for MBNL3 (experimental), MBNL1 (family control), CUGBP1 (nuclear fractionation control), and LDHA (cytoplasmic fractionation control). The MBNL3 27 kDa/29 kDa isoform (exon 2 exclusion) is found exclusively in the cytoplasm in Huh7 (A), C2C12 (C), and murine E15 forelimb (B). MBNL3 37 kDa/39 kDa isoforms (exon 2 inclusion) are located in both the nucleus and cytoplasm in C2C12 (C) and E15 forelimb (B).

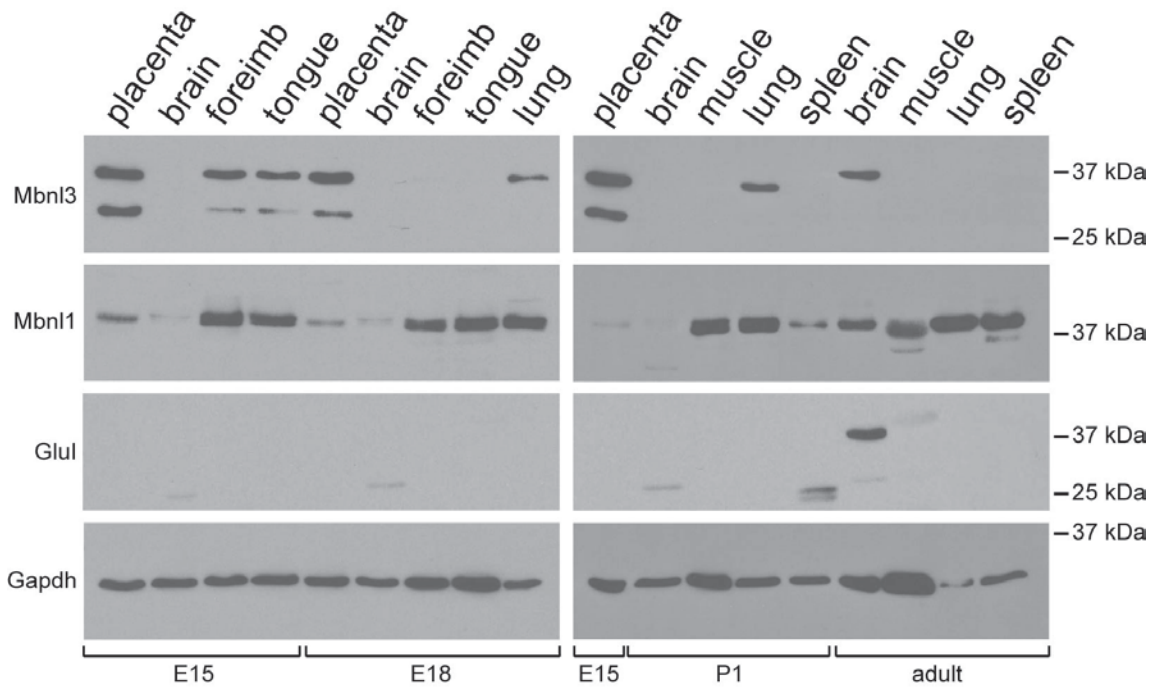

Figure 3-14. *Mbnl3* expression is restricted temporally and spatially during embryogenesis and postnatally. Murine tissues were taken at varying developmental time points and immunoblotted for Mbnl3, Mbnl1, and Gapdh. The band in adult brain is a cross reaction with glutamine synthetase (GluI).

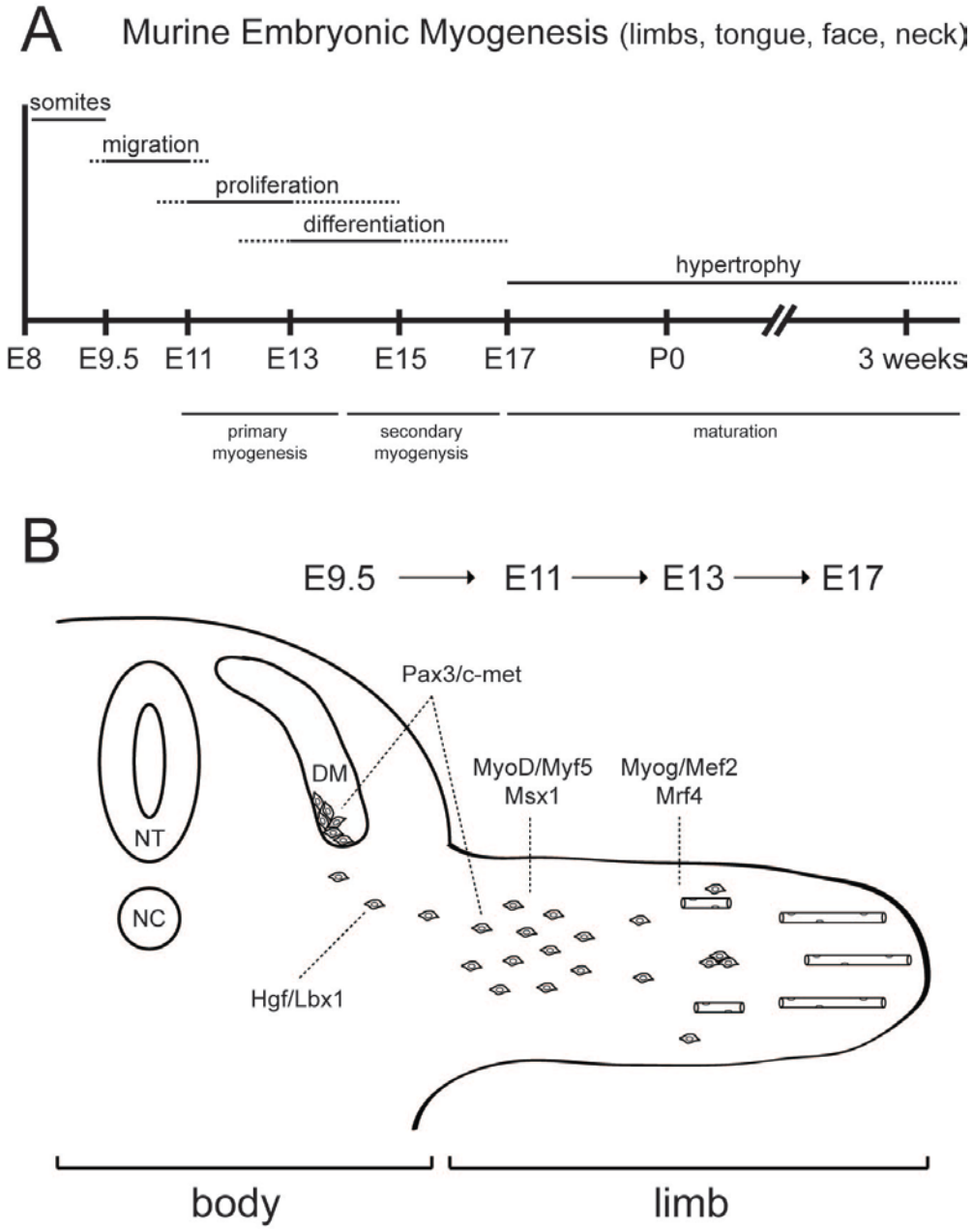

Figure 3-15. Murine embryonic myogenesis. (A) Time line of myogenic developmental windows during gestation (E8 = embryonic day 8). (B) Schematic of embryonic myogenesis. Myoblast precursors originate in the dermomyotome and travel to the developing limb bud, proliferate, and fuse to generate myotubes. Single cells represent myoblast precursors/myoblast and elongated tubes represent myotubes. Expression of critical regulatory genes during these time points are indicated by the dashed line. DM = dermomyotome; NT = neural tube; NC = notochord.

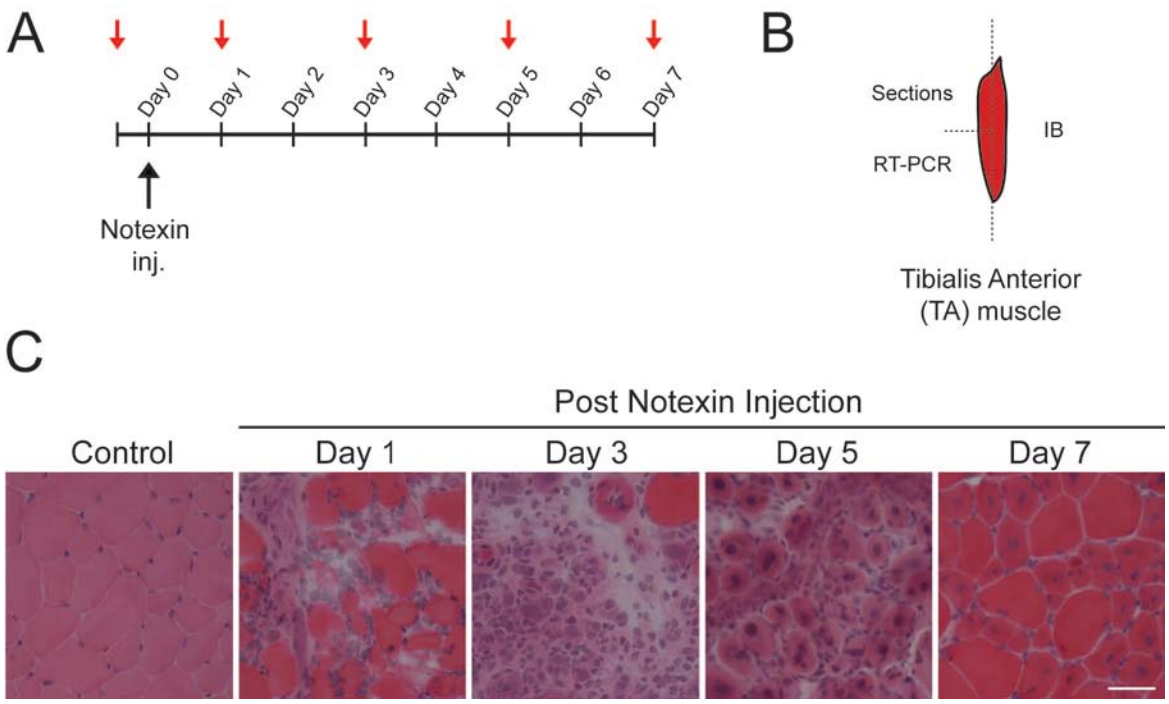

Figure 3-16. Notexin promotes murine skeletal muscle necrosis followed by regeneration *in vivo*. Timeline of Notexin injection and harvesting of the tibialis anterior (TA) muscle. (A) TA muscle was injected with Notexin and harvested at days 1, 3, 5, 7 post injection (red arrows). A pre-injection time point was used as control. (B) TA muscles were harvested at indicated time points, divided for immunoblot (IB), cryosection, and RT-PCR analysis. (C) H&E staining of 10 μm transverse TA cryosections from control and post-injection skeletal muscle. Scale bar = 50 μm.

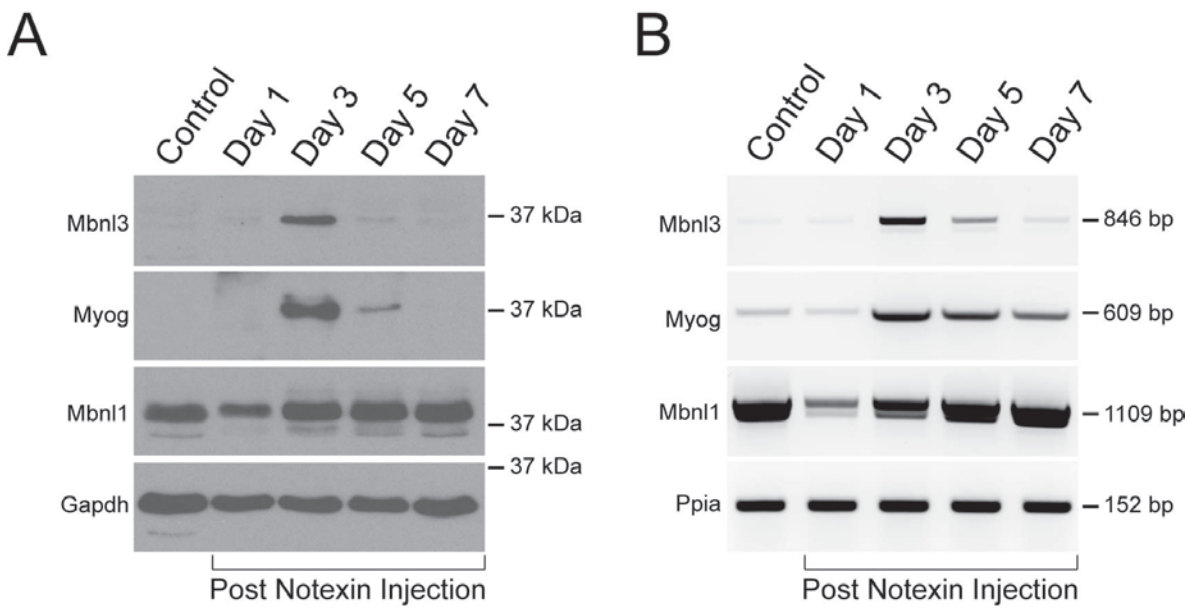

Figure 3-17. *Mbnl3* is expressed during adult skeletal muscle regeneration. (A) Immunoblot analysis of time points during Notexin mediated regeneration with antibodies specific for Mbnl3, myogenin (Myog), Mbnl1, and Gapdh. (B) RT-PCR of regenerating skeletal muscle with primers positioned in constitutive exons. *Ppia* is used as a loading control. Sizes of amplicons are indicated.

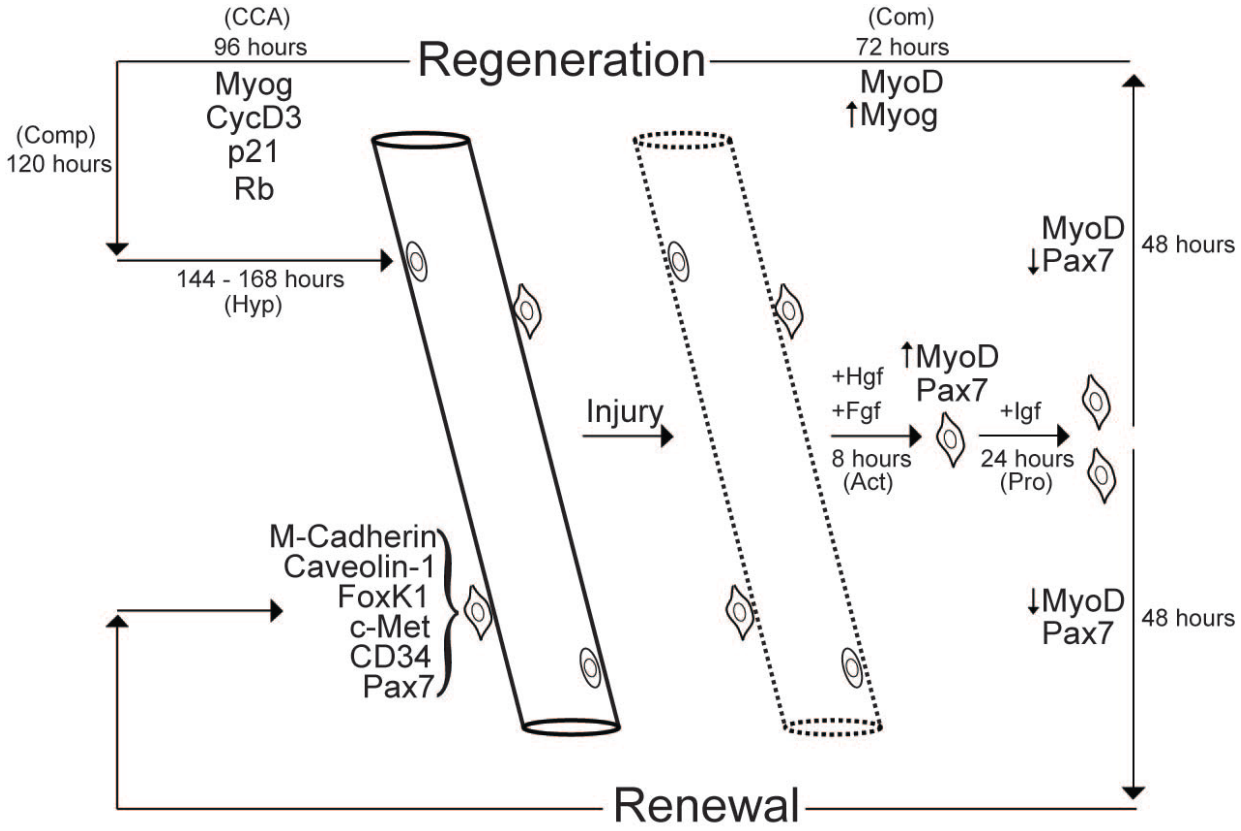

Figure 3-18. Adult skeletal muscle regeneration. Quiescent satellite cells (located adjacent to the sarcolemma) express specific markers including M-Cadherin. When the sarcolemma is injured or compromised (dashed line), satellite cells are activated. Activated satellite cells proliferate and contribute to both skeletal muscle regeneration as well as satellite cell renewal. Shortly after injury, Hepatocyte growth factor (HGF) and Fibroblast growth factor (FGF) activate quiescent satellite cells. Fgf and Insulin-like growth factor (IGF) promote proliferation of myoblasts expressing MyoD/Pax7. Forty-eight hours after activation, myoblasts contributing to regeneration downregulate Pax7 and commit to terminal differentiation, while myoblasts repopulating the satellite cell pool downregulate MyoD and enter quiescence. Abbreviations include: Act = activation, Pro = proliferation, Com = commitment, CCA = cell cycle arrest, Comp = completion, Hyp = hypertrophy.

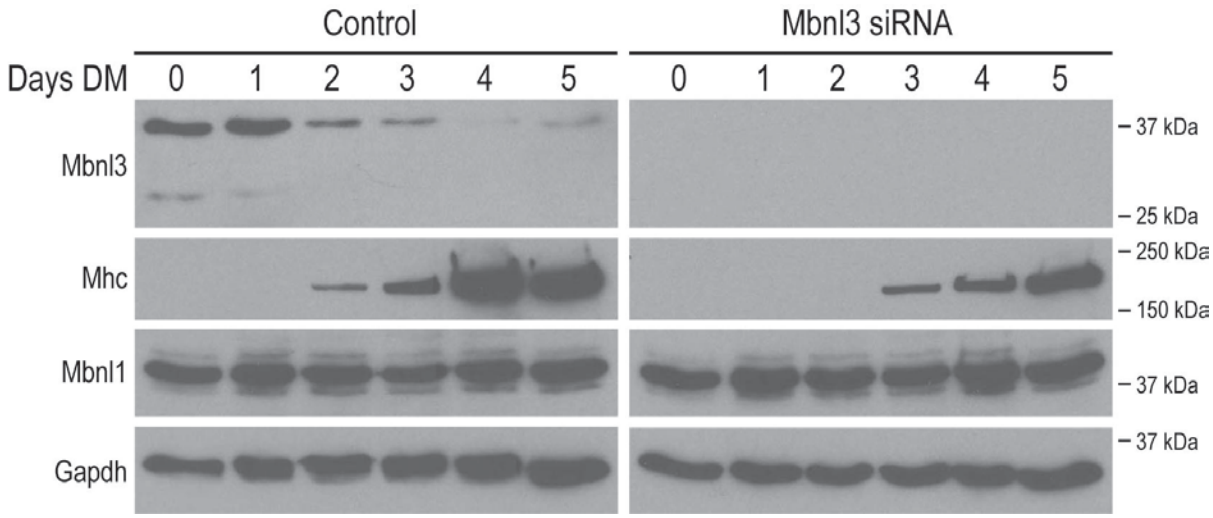

Figure 3-19. Loss of Mbnl3 inhibits myogenic differentiation in a C2C12 *in vitro* model. C2C12 cells were either mock treated (control) or transfected with *Mbnl3* siRNA (directed against the *Mbnl3* coding sequence and 3' UTR). 24 hours post siRNA, C2C12 cells were induced to differentiate by switching to differentiation media (DM). Differentiation status was monitored by immunoblot using terminal marker myosin heavy chain (Mhc). Antibodies against Mbnl3 were used to monitor knockdown, while Mbnl1 and Gapdh served as family and loading controls, respectively.

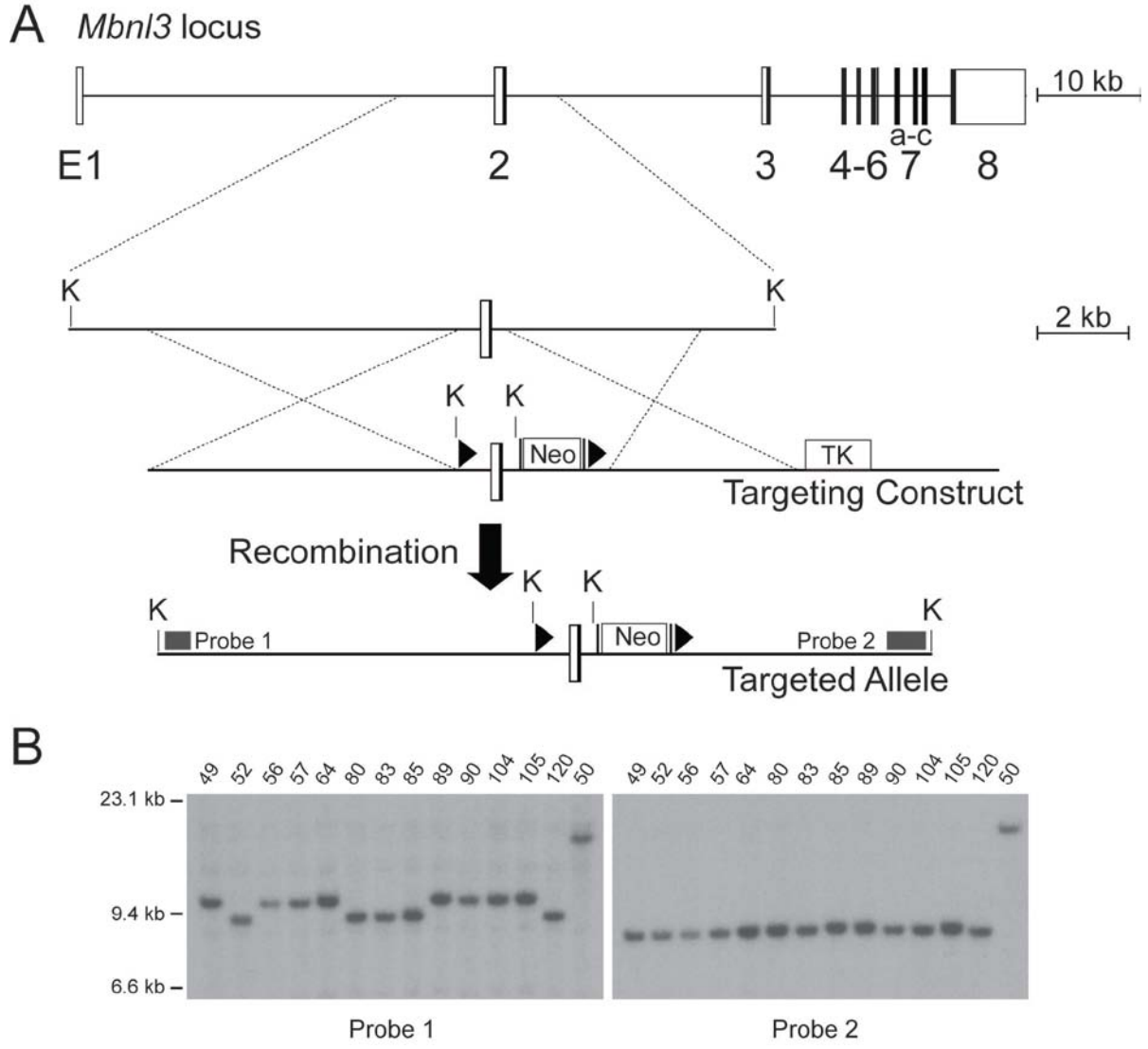

Figure 3-20. Generation of a conditional *Mbnl3* allele (*Mbnl3*$^{cond/Y}$) in ESCs. (A) *Mbnl3* exon 2 targeting strategy. *Mbnl3* consists of 10 exons (boxes) and 9 introns (horizontal line); untranslated regions (open boxes) and coding sequence (closed boxes) as well as exon numbers are indicated. Arms of homology from a *Kpn*I (K) fragment (15.5 kb) were used to target Mbnl3 exon 2 with a conditional exon 2 flanked by loxP sites. A Neomycin resistance cassette (*Neo*) and HSV-thymidine kinase cassette (*TK*) were used for positive and negative selection, respectively. Upon successful recombination, *Mbnl3* exon 2 (flanked by loxP sites) and two novel *Kpn*I restriction sites are introduced in the *Mbnl3* locus. (B) Genomic *Kpn*I digestion followed by Sourthern blotting with probes 1 and 2 (probe positions inticated by dark grey rectangle) reveal successful recombination events (correct 5' recombination = 8.7 kb, incorrect 5' recombination = 9.6 kb; 3' recombination = 7.5 kb; wild types allele = 15.5 kb). Clone identification is indicated above each blot.

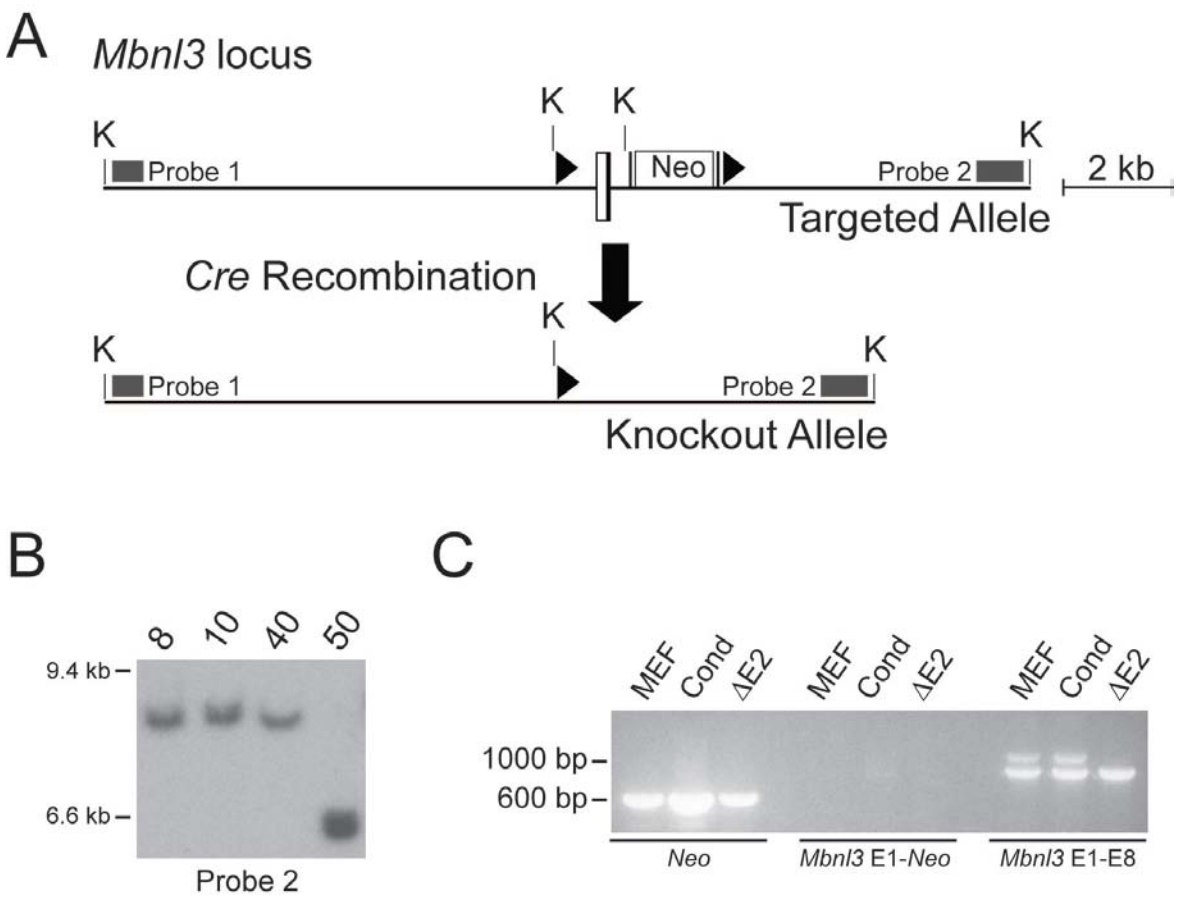

Figure 3-21. *Cre* mediated recombination of *Mbnl3*$^{cond/Y}$ allele removes exon 2 containing transcripts from ESCs. (A) *Mbnl3*$^{cond/Y}$ allele; *Kpn*I restriction sites (K), probes 1 and 2 for Southern blotting (dark grey rectangles), *Mbnl3* exon 2 (open box), neomycin resistance cassette (*Neo*). *Mbnl3* exon 2 and *Neo* are removed during *Cre* recombination. (B) Southern blotting verifies removal of *Mbnl3* exon 2 from the genome (*Mbnl3* conditional allele = 7.5 kb, *Cre* recombined allele = 6.5 kb). (C) Loss of *Mbnl3* exon 2 transcripts from ESCs. RT-PCR from mouse embryonic fibroblast (MEF) feeder cells, *Mbnl3*$^{cond/Y}$ ESCs, and *Mbnl3*$^{\Delta E2/Y}$ ESCs using primers positioned within *Neo*, forward in *Mbnl3* exon1-reverse in *Neo*, or forward in *Mbnl3* exon 1-reverse in *Mbnl3* exon 8. Expected sizes are ~1000 bp for internal *Neo*; 647 bp for *Neo* forward – *Neo* reverse; 1235 bp (+ exon 2)/ 988 bp (- exon 2) for *Mbnl3* forward exon 1-*Mbnl3* reverse exon 8.

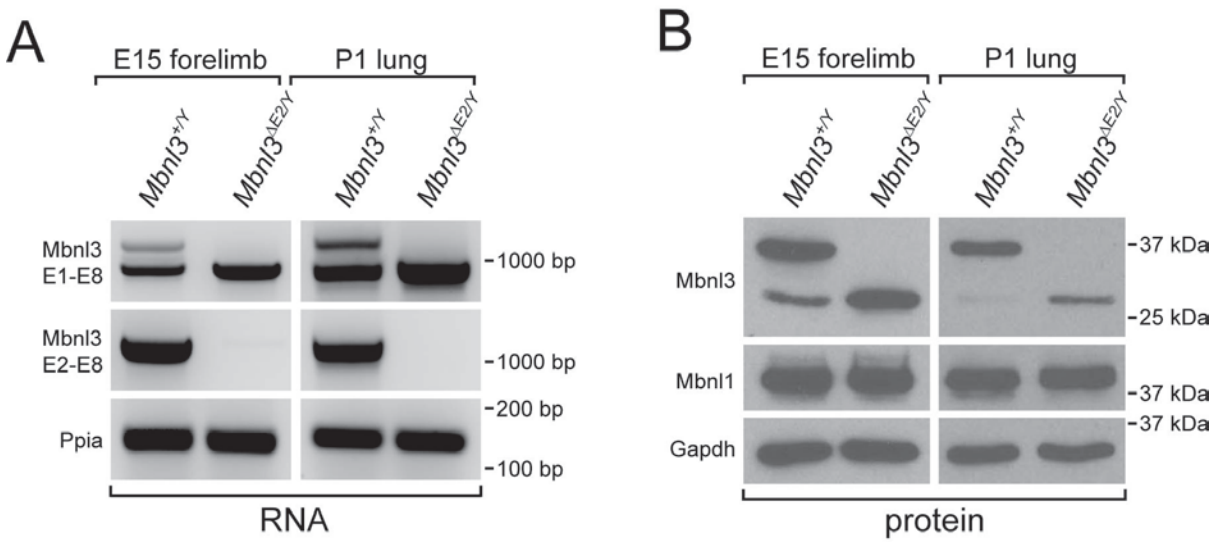

Figure 3-22. *Mbnl3*$^{\Delta E2/Y}$ mice lack full length *Mbnl3* isoforms. (A) Exon 2 is absent from *Mbnl3*$^{\Delta E2/Y}$ mRNA. RT-PCR of E15 forelimb and P1 lung with forward and reverse primers positioned in *Mbnl3* exon 1 and exon 8; *Mbnl3* exon 2 and exon 8; *Ppia* exon 3 and exon 4/5 (RT-PCR control). Expected sizes are 1235 bp (+ exon 2)/ 988 bp (- exon 2) for *Mbnl3* E1-E2; 1151 bp for *Mbnl3* E2-E8; 152 bp for *Ppia*. (B) Mbnl3 37 kDa isoforms are not present in *Mbnl3*$^{\Delta E2/Y}$ mice, while Mbnl3 27 kDa is upregulated. Immunoblot analysis of E15 forelimb and P1 lung with antibodies against Mbnl3, Mbnl1, and Gapdh.

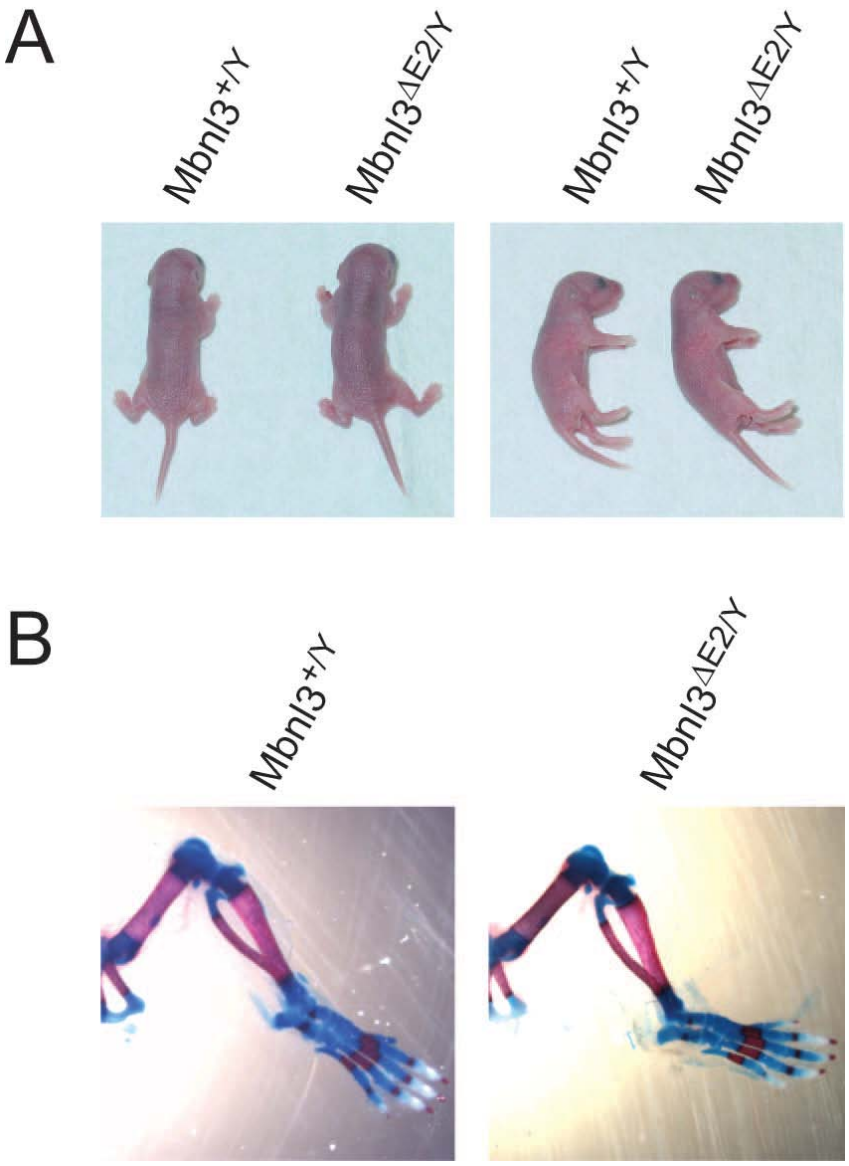

Figure 3-23. P1 *Mbnl3*$^{\Delta E2/Y}$ mice do not display defects associated with CDM. (A) *Mbnl3*$^{\Delta E2/Y}$ P1 pups are normal size at birth when compared to wild type and fail to demonstrate hypotonia, a cardinal phenotype in CDM. (B) Skeletal preps of hindlimbs from P1 wild type and *Mbnl3*$^{\Delta E2/Y}$ mice (bone is stained with Alizarin Red; cartilage is stained with Alcian Blue). Bone deformaties are not present in *Mbnl3*$^{\Delta E2/Y}$ mice.

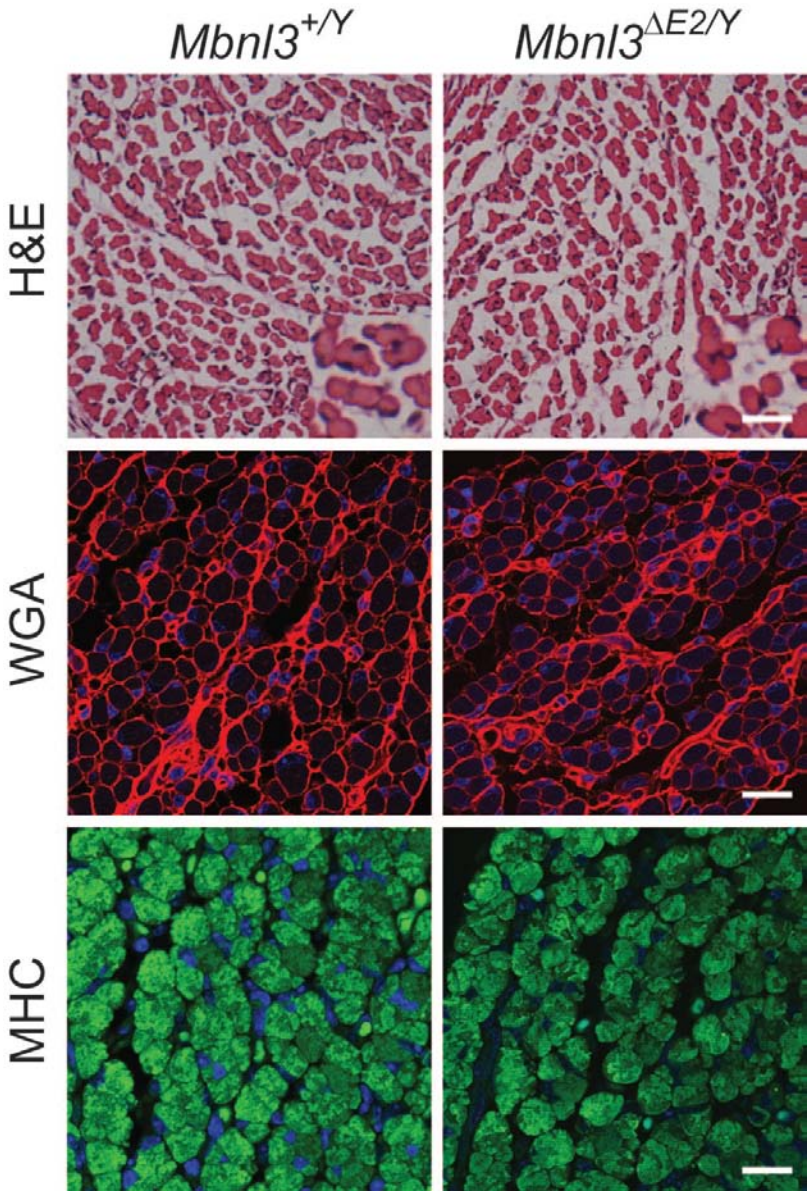

Figure 3-24. *Mbnl3*$^{\Delta E2/Y}$ neonates do not present with delayed myofibers at P1. Transverse sections from extensor carpi radialis brevis muscle. *Mbnl3*$^{\Delta E2/Y}$ P1 skeletal muscle develops normally. H&E staining (muscle morphology), Wheat Germ Agglutinin (WGA, extracellular matrix), Myosin Heavy Chain (MHC, developmental marker). Scale bar = 50 μm.

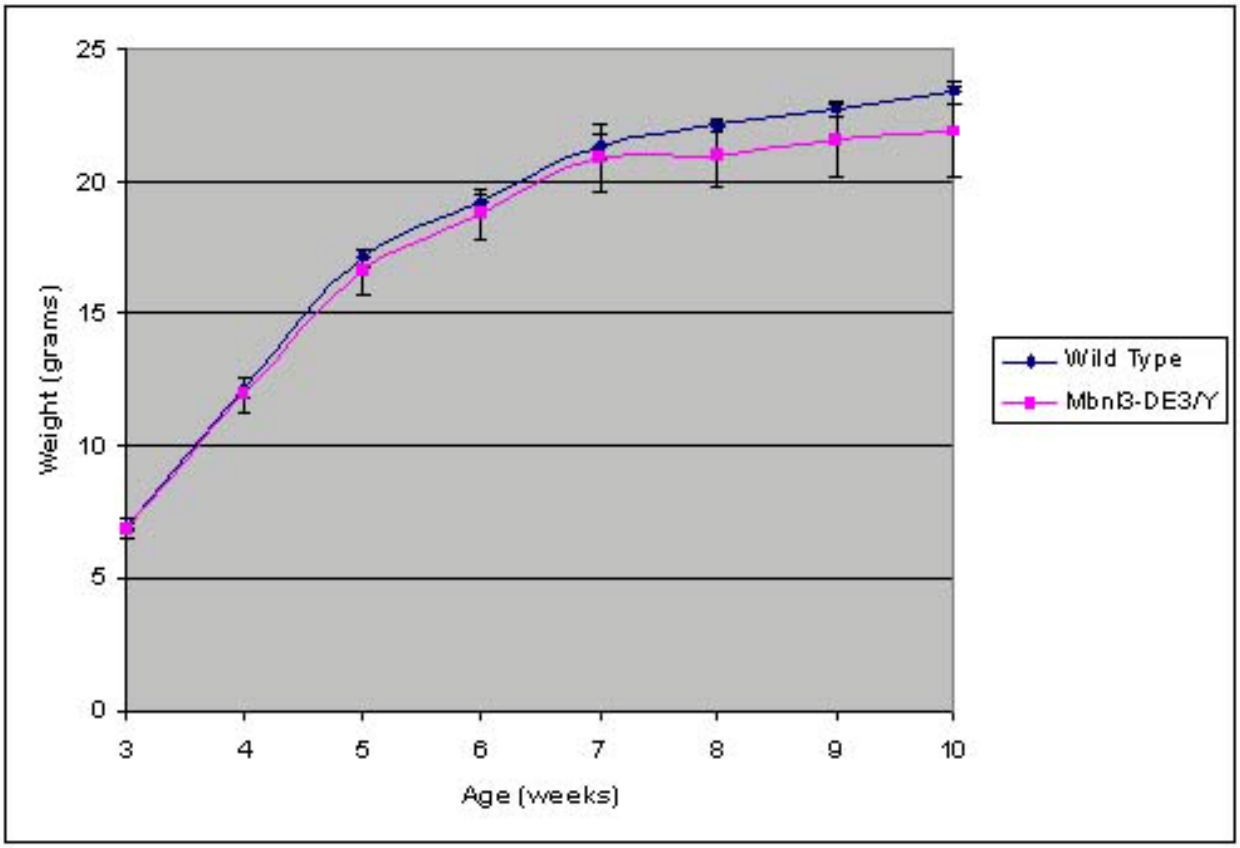

Figure 3-25. *Mbnl3*$^{\Delta E2/Y}$ mice gain weight normally up to 10 weeks of age. Weights from wild type (n=26) and *Mbnl3*$^{\Delta E2/Y}$ mice (n=17) were determined once per week from 3 weeks to 10 weeks of age (weaning to adults).

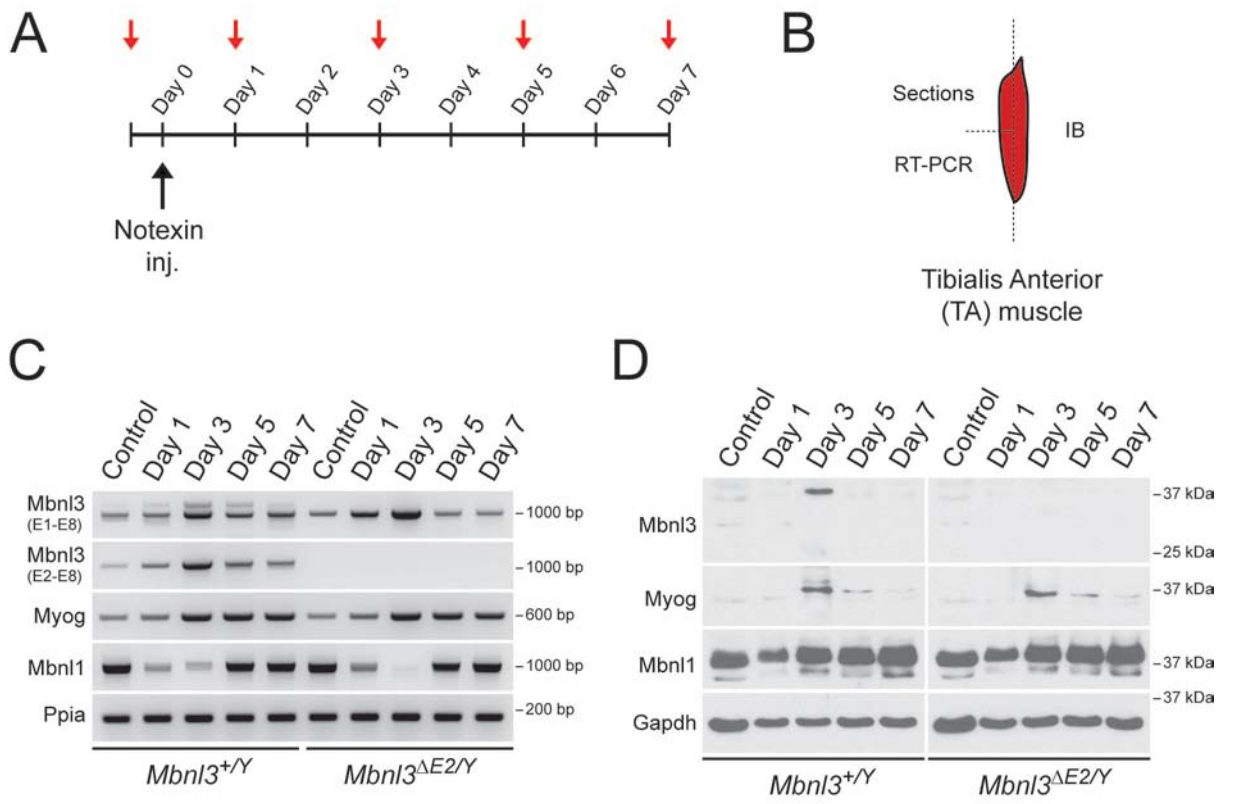

Figure 3-26. *Mbnl3*$^{\Delta E2/Y}$ mice do not express full length Mbnl3 isoforms during adult skeletal muscle regeneration. (A) Tibialis anterior (TA) muscle was injected with Notexin and harvested at Days 1, 3, 5, 7 post injection (red arrows). A pre-injection time point was used as control. (B) TA muscles were harvested at indicated time points, divided for immunoblot (IB), cryosection, and RT-PCR analysis. (C) RT-PCR of regenerating TA skeletal muscle with primers positioned in *Mbnl3* exon 1-exon 8 and exon 2-exon 8 demonstrate loss of *Mbnl3* exon 2 containing transcripts (*Mbnl3* E1-E8: exon 2 inclusion = 1235 bp; exon 2 exclusion = 988 bp. Mbnl3 E2-E8 = 1151 bp). Primers are positioned in constitutive exons of *Myogenin* (regeneration marker), *Mbnl1* (family control), and *Ppia* (loading control). (D) Immunoblot of regenerating TA muscle with antibodies against Mbnl3, Myogenin, Mbnl1, and Gapdh (loading control).

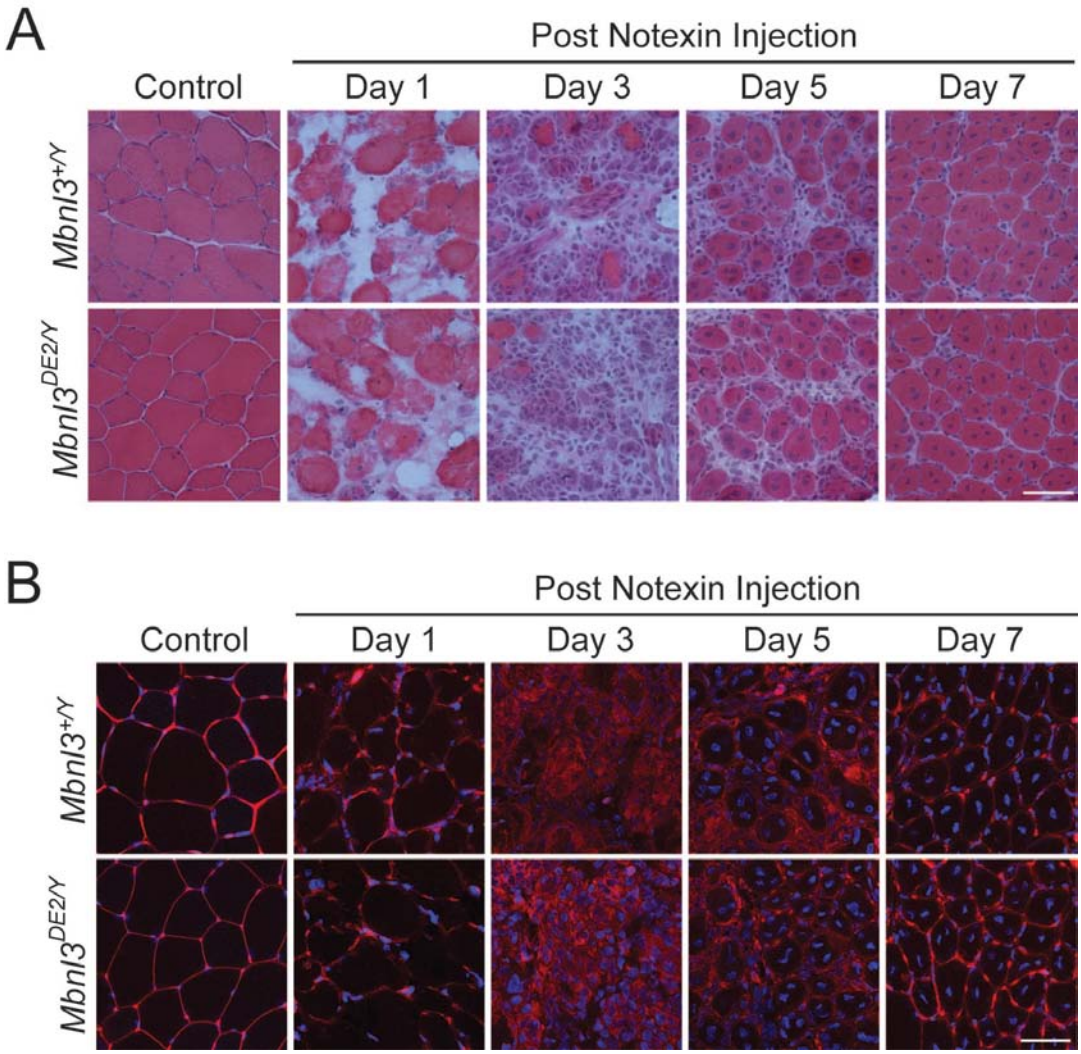

Figure 3-27. Normal skeletal muscle regeneration in *Mbnl3*$^{\Delta E2/Y}$ mice. (A) H&E staining of transverse TA cryosections (10 μm) from control and post Notexin injected skeletal muscle. (B) Wheat germ agglutinin (red) highlights the extracellular matrix during regeneration. DAPI (blue) indicates nuclear position. Scale bar = 50 μm.

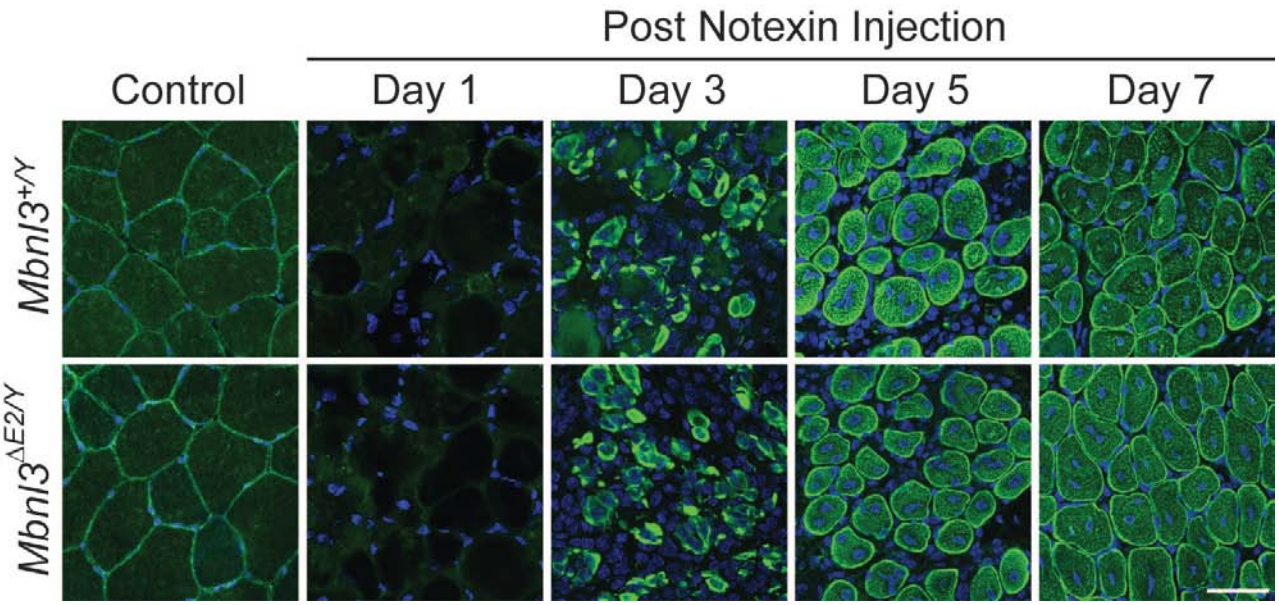

Figure 3-28. *Mbnl3*$^{\Delta E2/Y}$ normal skeletal muscle regeneration. Immunohistochemistry of transverse TA cryosections (10 μm) stained with an antibody against desmin, an intermediate filament which is expressed in both myoblasts and differentiated skeletal muscle. Scale bar = 50 μm.

CHAPTER 4
CONCLUDING REMARKS AND FUTURE DIRECTIONS

In this study, we explored the role of MBNL, a family of dsRNA binding proteins that are compromised by microsatellite expansions in the neuromuscular disorder, myotonic dystrophy. Loss of Mbnl1, in a mouse model, is sufficient to reproduce the majority of characteristic phenotypes of adult-onset DM, including muscle pathology, cataracts, and a defect in alternative splicing (Kanadia et al., 2003a). We first examined the possibility that MBNL1 regulates the alternative splicing of exons affected in disease by directly interacting with specific cis-elements in pre-mRNAs and modulating the inclusion or exclusion of that exon. Our results demonstrated that MBNL1 interacts with a motif directly upstream of *cTNT* exon 5, a fetal exon retained in DM, that is required to promote exclusion. Additionally, MBNL1 interactions with repeat RNAs show variable stabilities and toxicities, which may play a role in determining the severity of disease. Further understanding of these MBNL1 interactions with its RNA targets will be important for unraveling the complex alterations in alternative splicing and phenotypes associated with disease, as well as the contribution of other factors that potentially interact with repeat RNAs to promote DM.

We also investigated the involvement of Mbnl3 in the onset of CDM and progressive muscle wasting. Our results show that Mbnl3 protein isoforms are restricted to tissues that are affected in CDM during embryogenesis and regenerating skeletal muscle. Importantly, knockdown of *Mbnl3* expression in C2C12 cells demonstrated a delay in differentiation, consistent with the phenotype of immature skeletal muscle in CDM. This is in contrast to a previous report that a reduction in Mbnl3 expression promotes terminal differentiation. However, loss of Mbnl3 $4XC_3H$

isoforms does not inhibit embryonic muscle development or regeneration of the TA muscle in adult mice. One key difference between these experiments involves the newly identified 2XC$_3$H Mbnl3 isoform. The C2C12 delay in differentiation was dependent on the knockdown of both 4XC$_3$H and 2XC$_3$H isoforms by siRNA, while the *Mbnl3*$^{\Delta E2/Y}$ mouse still expresses 2XC$_3$H isoforms. One possibility to explain this disease is that the 2XC$_3$H isoforms can compensate for loss of the 4XC$_3$H isoforms. This observation highlights the potential importance of the Mbnl3 2XC$_3$H isoforms in development and predicts that loss of all Mbnl3 isoforms by sequestration are necessary to cause the CDM phenotype. As for the normal function of Mbnl3 in development, it will be interesting to establish the interacting RNAs and pathways that are governed by this gene. Unlike DM, very little is known concerning the molecular deficits in CDM, whether they be alternative splicing or other regulatory roles in RNA metabolism. While, the *Mbnl3*$^{\Delta E2/Y}$ mouse does not present with overt CDM or skeletal muscle wasting phenotypes, this does not mean that MBNL3 is not involved in molecular changes that contribute to disease. The unique localization pattern of Mbnl3 isoforms may provide insights into the normal function of these proteins and their dysfunction in disease. The ability to identify molecular phenotypes in DM/CDM mouse models are critical to identifying these changes in patients and developing necessary intervention or therapies.

DM and CDM represent an interesting class of disease in which a mutation displays a gain-of-function at multiple levels, potentially altering the expression and function of multiple genes in both cis and trans in a dominant manner (Fig. 3-3). This polygenic involvement likely contributes to the constellation of symptoms and

unpredictability that is characteristic of this neuromuscular disorder. However, the involvement of many genes allows us to separately model specific contributions of each factor involved to determine their specific contribution to disease in vivo. While transgenic mice expressing toxic $(CUG)_{exp}$ RNAs or knockouts modeling loss-of-function of have provided valuable insights into disease and the pathways disrupted, a more combinatorial genetic approach may be necessary to unravel additional molecular mechanisms involved in the onset of disease.

CHAPTER 5
MATERIALS AND METHODS

MBNL3

RNA Analysis

RNA was isolated from cell culture, staged C57BL6/J murine embryos, postnatal day 1 pups, and adult mouse tissues with Tri Reagent (Sigma) according to the manufacturer's protocol. For Northern blot analysis, approximately 15 μg of total RNA was denatured in Glyoxal/DMSO buffer (8% deionized glyoxal, 60% DMSO, 12 mM PIPES, 36 mM Bis-Tris, 1.2 mM EDTA, pH ~6.5) at 55 °C for 1 hr, immediately transferred to ice and resolved on a 1.2% agarose gel in 1X BPTE buffer (10 mM PIPES, 30 mM Bis-Tris, 1 mM EDTA, pH ~6.5) for 2 hr at 100 V. Following electrophoresis, the gel was rinsed in d_2H_2O, treated with 50 mM NaOH for 20 min, and neutralized in 20X SSC (pH 7) for 40 min. RNAs were transferred to Hybond-N+ nylon membranes (GE Healthcare) in 10X SSC using a neutral transfer by capillary action. Blots were crosslinked with a UV Stratalinker 1800 (Stratagene) with 120 mJoules and prehybridized for 2 hr in ExpressHyb Hybridization Solution (Clontech) at 68 °C. Northern probes were generated by [α-^{32}P] dCTP body labeling 50 ng of DNA template (*Mbnl1*, *Mbnl3*, *Ppia*) with Ready-to-Go DNA Labeling Beads – dCTP (GE Healthcare) and purified with Illustra ProbeQuant G-50 Micro Columns (GE Healthcare) according to the manufactures protocol. DNA probe in 2 μg/μL (1 mL total volume) Sheared Salmon Sperm DNA (Invitrogen) are denatured at 100 °C for 10 min and added to blot/hybridization solution and hybridized overnight at 68 °C. After hybridization, blots were washed once in 1X SSC, 0.1% SDS at room temperature on a shaker for 10 min, followed by three washes in 0.5X SSC, 0.1% SDS at 65 °C in a heated/shaking water

bath for 10 min each. Blots were exposed to Biomax Film and X-Omatic Intensifying Screens (Kodak) at -80 °C.

For cDNA preparation, approximately 25 μg of total RNA + 2.5 μg oligo $d(T)_{12-18}$ primers (Invitrogen) were denatured at 65 °C for 5 min/ice for 1 min and added to a final volume of 50 μL in 0.2 mM dNTPs, 1X 1st Strand Buffer (Invitrogen) + 8 mM DTT + 100 U RNasin (Promega) + 500 U Superscript III reverse transcriptase (Invitrogen). Reverse transcription reaction conditions were 25 °C for 5 min – 50 °C for 60 min – 70 °C for 15 min in a thermocycler. After reverse transcription, 2 U RNase H (Invitrogen) was added to the RT reaction and incubated at 37 °C for 20 min.

RT-PCR analysis was carried out using 2 μL template cDNA (i.e. 1 μg RNA input) in a 50 μL reaction containing 1X High Fidelity buffer (5 Prime), 0.4 mM dNTPs, 30 pmol forward primer, 30 pmol reverse primer, and 2.5 U Triplemaster Taq (5 Prime). PCR reaction conditions were 96 °C for 2 min (denaturation step), followed by 96 °C for 30 sec – 58 °C for 30 sec – 72 °C for 1 min (amplification step – see below for cycle numbers per primer set), followed by 72 °C for 5 min (final elongation step). Cycle numbers for amplification step (per primer pair) were as follows: 30X for MSS3655-MSS3759, MSS3655-MSS3760, MSS3655-MSS3763, MSS3586-MSS3587, MSS3225-MSS3247, MSS3648-MSS3247, MSS3580-3581, MSS3582-MSS3583, MSS3584-MSS3585; 32X for MSS3648-3652, MSS2916-3652, MSS4173-MSS4177, MSS2725-2726 (see Figure 4-1 for primer identification, position, and sequence). PCR products were resolved on a 1%-1.5% agarose gel, stained with ethidium bromide, and pictures taken on the ImageQuant 400 (GE Healthcare).

Protein Lysate and Fractionation

Whole cell lysate was prepared from cell culture, staged C57BL6/J murine embryos, postnatal day 1 pups, and adult mouse tissues by disruption with a sterile pestle in lysis buffer (20 mM HEPES-KOH, pH 8.0, 100 mM KCl, 0.1% Igepal, 0.5 μM phenylmethylsulfonyl fluoride, 5 μg/mL pepstatin A, 1 μg/mL chymostatin, 1 μM e-aminocaproic acid, 1 μM p-aminobenzamidine, 1 μg/mL leupeptin, 1 μg/mL aprotinin) on ice. Protein samples were sonicated three times at 40% amplitude and spun at 16.1 RCF for 20 min at 4 °C. Supernatant was collected, glycerol was added to a final concentration of 20%, and samples were aliquoted and stored at -80 °C. Protein concentrations were determined using the DC Protein Assay (Bio-Rad). Cellular fractionations were prepared using NE-PER (Thermo Scientific – Pierce) nuclear and cytoplasmic protein fractionation kit according to the manufacturers protocol.

Immunobloting

50 μg total protein or 35 μL immunoprecipitation eluate was resolved on a 12.5% SDS-acrylamide gel and transferred to 0.22 μm nitrocellulose (GE Water & Process Technologies) by electroblotting (using a Trans-Blot® SD semi-dry transfer cell; Bio-Rad). Blots were blocked for 45 min in 5% non-fat dry milk in PBS (pH 7.4) at room temperature and immunoblotted in 5% non-fat dry milk in PBS (pH 7.4) with 0.05% Igepal for 1 hr rotating at room temperature with the following primary antibodys: Mbnl3 purified antisera C (1:2500 dilution), Mbnl1 (A2764) antisera (1:1000 dilution), CUGBP1 (3B1), c-myc (9E10), Myosin Heavy Chain (Sigma – MY32 1:1000 dilution), Myogenin (Santa Cruz – F5D 1:1000 dilution), Lactate Dehydrogenase A (Cell Signaling – 1:1000), Glutamine Synthetase (Abcam – 1:1000), SMN (Santa Cruz – 2B1 1:1000

dilution), or Gapdh (6C5 – Abcam 1:25,000 dilution). Blots were washed three times in PBS (pH 7.4) with 0.05% Igepal for 10 min and incubated for 1 hr rotating at room temperature in blotting milk with α-rabbit IgG horseradish peroxidase or α-mouse IgG horseradish peroxidase (1:5000; GE Healthcare) secondary antibody. The membranes were washed as described above, followed by a PBS (pH 7.4) wash, developed in Amersham ECL® detection reagents (GE Healthcare) or Supersignal West Femto reagents (Thermo Scientific – Pierce) and exposed to Biomax Light Film (Kodak). To develop immunoprecipitation/immunoblots, Genscript one-step complete IP-Western kit (Genscript) was used according to the manufacturers protocol and exposed to Biomax Light Film (as described above).

Generation of an Mbnl3 Polyclonal Antibody

Mbnl3 polyclonal anti-sera was generated and purified by Genscript (Piscataway, NJ). In short, Mbnl3 C-terminal peptide (NVPYVPTTTGNQLKY) was synthesized and KLH conjugated to the N-terminus using an additional N-terminal cysteine (final immunizing peptide: KLH-CNVPYVPTTTGNQLKY). Four rabbits were injected with Mbnl3 immunizing peptide followed by two additional injections to boost antibody production/affininty. Rabbits were bleed prior to the first immunization for a pre-immune sera control for screening. Test bleeds were screened by immunoblot using 50 μg E14 placenta whole cell lysate and anti-sera at a dilution of 1:500 (immunoblot and whole cell lysate protocols were performed as described in the immunoblot section of materials and methods). Anti-seras A and C, with high specificity and low background, were further screened (at a dilution of 1:1000) by immunoblot using decreasing amount of input E14 placenta whole cell lysate (i.e. 50 μg, 25 μg, 12.5 μg, 6.25 μg, 3.13 μg, and

1.56 μg) to compare relative affinities. Anti-sera A and C were chosen for further affinity purification using the original immunizing peptide. Purified peptides were concentrated using centricon 10 concentrators (Micon) according to the manufacturer's protocol. Glycerol was added to a final concentration of 20% to α-Mbnl3 A and α-Mbnl3 C (final antibody concentrations: α-Mbnl3 A = 7.8 mg/mL, α-Mbnl3 C = 4.8 mg/mL) polyclonal antibodies and stored in aliquots at -80 °C.

Expression Vectors

Mbnl1, Mbnl2, Mbnl3 coding sequences were amplified using primers MSS3580-MSS3581, MSS3582-MSS3583, and MSS3584-MSS3585 from MEF cDNA in a 50 μL reaction containing 2 μL cDNA, 1X High Fidelity buffer (5 Prime), 0.4 mM dNTPs, 30 pmol forward primer, 30 pmol reverse primer, and 2.5 U Triplemaster Taq (5 Prime). PCR reaction conditions were 96 °C for 2 min (denaturation step), 96 °C for 30 sec – 58 °C for 30 sec – 72 °C for 1 min 30 sec (30X cycles – amplification), followed by 72 °C for 5 min (final elongation step) (see Figure 4-1 for primer identification, position, and sequence). PCR products were phenol:chloroform extracted, precipitated, and digested in a 40 μL reaction containing 1X NEB buffer 3, BSA (100 μg/mL), BamHI (40 units - NEB), and XbaI (40 units - NEB) for 2 hr at 37 °C. Following digestion, amplicons were resolved on a 1.0% agarose gel, gel extracted (Qiagen) following the manufacturers protocols, and resuspended in 50 μL water. 5 μg pcDNA 3.1/myc-His MCS A (Invitrogen) was digested in a 50 μL reaction containing 1X NEB buffer 3, BSA (100 μg/mL), BamHI (50 units - NEB), and XbaI (50 units - NEB) for 2 hr at 37 °C, resolved on a 1.0 % agarose gel, and gel extracted (Qiagen) following the manufacturers protocols, and resuspended in 50 μL water. pcDNA 3.1/myc-His and Mbnl inserts were

ligated in a 20 μL reaction containing 1X T4 DNA ligase buffer (NEB), T4 DNA ligase (5X105 units - NEB), 12 ng DNA backbone (pcDNA), and 36 ng DNA insert (Mbnl) for 1 hr, followed by transformation into Top10 cells (Invitrogen) and plating according to the manufacturers protocol. Individual colonies were picked, grown in 150 mLs LBamp overnight at 37 °C shaking, and maxiprepped (Qiagen) according to the manufacturers protocol. Expression vectors were sequenced (University of Florida ICBR sequencing core) for confirmation.

Immunoprecipitation and Mass spectrometry

Protein lysates for immunoprecipitation were prepared from cells or E15 placenta in IPP-150 with 0.1% Igepal (50 mM Tris-Cl, pH7.4, 150 mM NaCl, 0.1% Igepal, 1X PicD, PicW) by disruption with a sterile pestle in a microfuge tube on ice, followed by sonicating three times at 40% amplitude, and spinning down the lysate at 16.1 RCF for 20 min (4°C). Supernatant was transferred to a new microcentrifuge tube on ice and protein concentration was determined using DC protein assay (Bio-Rad). Lysate was left on ice until the preparation of beads was completed (see below).

For bead preparation, Protein-A Dynabeads were vortexed for 30 sec and 100 μLs were added to a microcentrifuge tube. The microcentrifuge tube was placed on the Dynal magnet (Invitrogen) for 1 min to remove the Protein-A Dynabeads from solution, followed by aspiration of the supernatant. The microcentrifuge tube was removed from the magnet and Protein-A Dynabeads were washed with 500 μL 0.1 M sodium phosphate buffer (pH 8) three times by repeating the above steps. Next, the beads were resuspended in 100 μL 0.1 M sodium phosphate buffer (pH 8.0) followed by the addition of 20 μg purified α-Mbnl3 C polyclonal antibody and incubation for 30 min at RT

with rotational mixing to promote IgG capture. Antibody-bound beads were washed two times with 0.2 M triethanolamine (pH 8.2) as described above, resuspended, and incubated in 1 mL of 20 mM dimethyl pimelimidate/0.2 M triethanolamine (pH 8.2) (DMP - Sigma) for 30 min at RT to covalently crosslink α-Mbnl3 C antibody to protein-A. The crosslinking reaction was quenched by replacing DMP/triethanolamine with 1 mL 50 mM Tris (pH 7.5) and incubating for 15 min at RT with rotational mixing. Beads were further washed with IPP-150 + 0.1% Igepal three times as described above.

For immunoprecipitation, Protein-A:α-Mbnl3 C IgG beads were resuspended in the 5 mgs of previously prepared protein lysate (~ 400 μL total volume) and incubated at 4°C for 2 hr with rotational mixing. Beads were collected by placing the microfuge tube on the magnet for 1 min, followed by a flash spin (5 sec, 16.1 RCF) to collect them at the bottom of the tube. The beads containing the immunoprecipitate were resuspended in 35 μL Laemmli buffer, incubated at 99 °C for 5 min on an Eppendorf Thermomixer (1100 RPM), followed by removal of the beads from solution using the magnet as previously described. The supernatant was collected for downstream applications, immunoblotting and mass spectrometry.

To identify the proteins immunoprecipitated using the α-Mbnl3 C polyclonal antibody, proteins were resolved on a 12.5% SDS-acrylamide gel and stained with the Novex colloidal coomassie blue kit (Invitrogen) according to the manufacturers protocol. Bands of interest were excised from the gel using separate sterile scalpels (Bard-Parker) and stored in separate microcentrifuge tubes at 4°C. Samples were submitted to the UF ICBR Proteomic Core facility for liquid chromatography-mass spectrometry (LCMS) analysis to identify proteins.

siRNA and Plasmid Transfections

C2C12 and Huh7 cells were grown on 10 cm plates in growth media containing DMEM (Invitrogen), 20% FBS (Invitrogen), 1% penicillin/streptomycin (Invitrogen) and DMEM (Invitrogen), 10% FBS (Invitrogen), and 1% penicillin/streptomycin (Invitrogen), respectively, in a humidified 37 °C, 5% CO_2 incubator. For siRNA treatment, 1.5 x 10^5 cells were seeded per well of a 6-well plate (Sarstedt) in 2 mLs antibiotic-free growth media. Cells were treated with *Mbnl3* or *MBNL3* siRNA (ON-TARGETplus SMARTpool, Dharamcon) 6-hr post seeding as follows. Two separate tubes were prepared for siRNA and transfection reagent (Dharmafect 4 - Dharmacon). For transfection of one well of a 6-well plate, 5 μL of transfection reagent was mixed with 195 μL OptiMEM I (Invitrogen) in tube 1. 10 μL of 20 μM siRNA was diluted to 1 μM by addition of 90 μL of cell culture grade PBS (pH 7.4 – Invitrogen) and 100 μL of OptiMEM I in tube 2. Tubes 1 and 2 were briefly vortexed (1 sec) and incubated for 5 min at room temeperature (RT). The contents of tube 2 were added to tube 1, vortexed briefly (1 sec), and incubated for 20 min at RT. The entire 400 μL of transfection mixture was added drop-wise into one well of 6-well plate and swirled to mix. Media was replaced by fresh C2C12 antibiotic-free growth media 24 hr post transfection.

COSM6 cells were grown in growth media containing DMEM (Invitrogen), 10% FBS (Invitrogen), 1% L-glutamine (Invitrogen), and 1% penicillin/streptomycin (Invitrogen) in a humidified 37 °C, 5% CO_2 incubator.. Six hr prior to transfection, 2.0 x 10^5 cells were seeded per well of a 6-well plate in antibiotic-free growth media. In a sterile microfuge tube, 6 μL Fugene 6 (Roche) was diluted in 180 μL OptiMEM I (Invitrogen), vortexed (1 sec) and incubated for 5 min at RT. After incubation, 2 μg of

plasmid (either myc-tagged Mbnl1, Mbnl2, Mbnl3, or empty vector pSP72) was added to the tube containing Fugene 6:OptiMEM I, vortexed (1 sec), and incubated for 15 min at RT. Following incubation, the total contents of each transfection were added to individual wells of the 6-well plate, swirled to mix, and returned to the humidified 37 °C, 5% CO_2 incubator. The following day, cells were washed with PBS (pH 7.4) and antibiotic-free growth media was replaced. Cell lysates were processed 48 hr post transfection for analysis.

C2C12 differentiation

For C2C12 cell differentiation, proliferating cells were seeded at 1.5 x 10^5 per well of a 6-well plate in C2C12 growth media. At ~ 90-95% confluency, growth media was aspirated, cells were washed twice with 2 mLs of PBS (pH 7.4), and fresh C2C12 differentiation media containing DMEM, 2% horse serum (Invitrogen), and1% Penicillin/Streptomycin was added to each well. Cells were differentiated for 5 days (media was replaced with fresh C2C12 differentiation media every 24 hr). For *Mbnl3* siRNA experiments, C2C12 cells were differentiated 24 hr post siRNA treatment.

Skeletal Muscle Regeneration

Lyophilized Notexin (Latoxan - Valence, France) was reconstituted in sterile 154 mM NaCl to a final concentration of 100 μg/mL. 10 μL aliquots (1 μg) were stored at -80 °C until use. For injections, 1X 10 μL aliquot was thawed on ice and sterile 154 mM NaCl was added to a final volume of 1 mL (working concentration is 1 μg/mL). For injections, 10-12 week old C57BL6 mice were anesthetized with isoflourene (following the University of Florida IACUC approved protocol). Using a 29-gauge needle, tibialis anterior (TA) muscles were injected with 100 μL (100 ng Notexin) of working

concentration Notexin by lateral injection and dragging the needle along the muscle lengthwise (proximal to distal) while injecting. Mice were allowed to recover in their cages while monitoring their behavior. TA muscles were taken at days 1, 3, 5, and 7 post-injection for analysis. TA muscle was divided in two lengthwise using surgical scissors, half was used for protein preparation and analysis (following whole cell lysate and immunobloting protocols) and half was snap frozen on a wooden dowel positioned in 10% gum tragacanth (in PBS, pH 7.4 - Sigma) in pre-chilled isopentane (-40 °C) and stored at -80 °C in isopentane (Fisher). TA muscle remaining after cryosectioning was used for RT-PCR analysis (see RNA Analysis).

Mbnl3$^{\Delta E2/Y}$ Mouse Generation

The *Mbnl3* conditional targeting vector was constructed using standard recombineering bacterial strains and techniques (reagents and protocols can be found at http://recombineering.ncifcrf.gov/). In short, an approximately 11 kb fragment (containing 6.8 kb sequence from directly upstream and 4.2 kb sequence from directly downstream of *Mbnl3* exon 2) surrounding *Mbnl3* exon 2 was retrieved from a C57BL6 bacteria artificial chromosome (BAC) into a high copy plasmid backbone containing the negative-selection marker Herpes Simplex Virus Thymidine Kinase gene (*HSV-TK*) downstream of the *Mbnl3* fragment. The *Mbnl3* exon 2 was flanked by an upstream loxP site and a downstream *Neo*^{R} cassestte/loxP site (see Fig. 3-16). The final targeting construct was sequenced and restriction mapped to confirm its identity. To prepare the targeting construct for ESC targeting, approximately 100 μg of the *Mbnl3* conditional targeting vector was linearized with NotI (NEB), phenol:chloroform extracted,

ethanol precipitated, and resuspended in PBS (pH 7.2) at a final concentration of ~2.2 μg/μL.

All C57BL/6 ESCs (Millipore, catalog number CMTI-2) in this targeting protocol were thawed, grown, and passaged on mitotically inactivated MEFs using standard ESC techniques in a 37 °C, 5% CO_2 humidified incubator. ESC growth media contained: 409 mLs DMEM (Invitrogen), 75 mLs Defined FBS (HyClone), 5 mLs 100X nucleoside stock (standard ESC formulation, see Millipore recommendations), 100X Pen/Strep (Invitrogen), 0.76 mLs β-mercaptoethanol (Invitrogen – 55 mM stock concentration), and 50 μL LIF (Millipore – 10^7 units/mL). For targeting, 10^8 ESCs were electroporated in 0.8 mLs PBS (pH 7.2) containing 100 μg *Mbnl3* conditional targeting vector using a BIORAD Genepulser (parameters = 0.8 volts, 3 μF, 200 Ohms, and 0.2 sec time constant in a 0.4 cm cuvette), allowed to recover for 10 min in the cuvette, and plated on 15X 10 cm gelatin coated tissue culture plates in ESC growth media. The following day, ESCs were positively selected for neomycin resistance in ESC growth media supplemented with G418 (Invitrogen) at 275 μg/mL. Two days post electroporation, ESCs were positively and negatively selected for successful recombination in ESC growth media supplemented with G418 at 275 μg/mL and FIAU (Moravek Biochemicals) at 0.2 μM for 72 hr. Following selection, individual colonies were picked with a P200 pipet tip, trypsinized, and plated in 24-well plate for colony expansion. Following expansion, individual wells of the 24-well plate were trypsinized and split, with half of the ESCs transferred to 0.5 mL freezing media in cryotubes, containing 60% growth media, 30% FBS, and 10% DMSO (Fisher Scientific), frozen in an isopropanol bath at -80 °C overnight, and transferred to $N_{(l)}$ the following day. The other half of ESCs from

individual wells of 24-well plates were transferred to wells of gelatin coated 12-well plates, and grown to confluency for genomic DNA preparation and Southern blot analysis for successful recombinants (see genomic DNA isolation and Southern blotting).

Once positive ESC clones were identified, the targeted ESCs were thawed, expanded, and additional stocks were cryopreserved. *Mbnl3*^{cond} ESCs chosen for the generation of mice were screened for murine pathogens using the IMPACT I profile at the Research Animal Diagnostic Laboratory (RADIL) at the University of Missouri and karyotyped at the University of Florida Cytogenetics Lab. *Mbnl3*^{cond} ESCs that tested negative for pathogens and displayed normal chromosome counts were sent to the University of Michigan Transgenic Core for blastocyst injection. In short, 45 B6(Cg) -*Tyr*$^{c-2J/J}$ blastocysts (i.e. albino C57BL/6) were injected with 8-16 ESCs and transferred in utero to a multiple pseudo-pregnant C57BL/6 females. The resulting chimeric mice were shipped to the University of Florida through, quarantined in isolators for pathogen testing, and released into specific pathogen free (SPF) housing.

To obtain germline transmission of the *Mbnl3*^{cond} allele, *Mbnl3*^{cond} male chimeric mice were mated with female B6(Cg) -*Tyr*$^{c-2J/J}$ females. Since *Mbnl3* is an X-linked gene, transmission through the chimeric male germline only appears in F1 females. Therefore, F1 female pups (with a black coat, indicating the gamete originated from BL6 ESCs) from the *Mbnl3*^{cond} male chimera and female B6(Cg) -*Tyr*$^{c-2J/J}$ cross possessed a *Mbnl3*$^{cond/+}$ genotype. To eliminate *Mbnl3* exon 2 and the 4XC3H Mbnl3 isoforms, *Mbnl3*$^{cond/+}$ females were crossed to B6.C-Tg(*CMV-cre*)1Cgn/J males (Jackson Labs, catalog number 006054) to obtain F2 *Mbnl3*$^{\Delta E2/+}$; *CMV-cre* females (*CMV-Cre* is also X-

linked). To obtain knockout mice, F2 *Mbnl3*$^{\Delta E2/+}$; *CMV-cre* females were crossed to wild type C57BL6/J males, F3 *Mbnl3*$^{\Delta E2/Y}$ males were analyzed for phenotypes.

Genomic DNA Isolation, Southern Blotting, Probe Generation and Genotyping

Targeted ESC genomic DNA was isolated from individual confluent wells of a 12-well plate for *Mbnl3*$^{cond/Y}$ homologous recombination screening. In short, individual wells were washed in PBS (pH 7.2 - Invitrogen) to remove debris and lysed in 0.5 mLs DNA extraction buffer (200 mM NaCl, 100 mM Tris (pH 8.5), 5 mM EDTA, 0.2% SDS, 200 μg/mL Proteinase K) overnight in a humidified 37 °C, 5% CO_2 incubator. The following day, lysates were collected and precipitated by adding an equal volume of 100% ethanol, inverting several times until the genomic DNA was visible, and spinning samples down at 16.1 RCF for 10 min in a table top centrifuge (room temperature). DNA pellets were washed in 1 mL 70% ethanol and centrifuged as before. The supernatant was removed and the genomic DNA pellet was allowed to air dry for one min, followed by resuspension in 100 μL TE buffer (pellets were not vortexed to avoid shearing) by pipeting with a wide bore tip.

ESC *Mbnl3* conditional targeting 5' (MSS3256-MSS3272) and 3' (MSS3258-MSS3259) southern probes were generated by PCR in a 50 μL reaction containing 100 ng 129 genomic DNA, 1X High Fidelity buffer (5 Prime), 0.4 mM dNTPs, 30 pmol forward primer, 30 pmol reverse primer, and 2.5 U Triplemaster Taq (5 Prime). PCR reaction conditions were 96 °C for 2 min (denaturation step), 96 °C for 30 sec – 68 °C for 1 min 15 sec (30X cycles – amplification), followed by 68 °C for 5 min (final elongation step) (see Figure 4-1 for primer identification, position, and sequence). PCR products were subcloned using the TOPO cloning kit (Invitrogen), grown in 150 mLs

LB_{amp} at 37 °C shaking overnight, and maxiprepped (Qiagen) according to the manufacturers protocols. Cloned products were sequenced (University of Florida ICBR sequencing core) for confirmation (see Figure 4-1 for primer identification, position, and sequence).

For Southern blotting, 20 μL genomic DNA (~10-20 μg) was digested in a 30 μL reaction containing 150 units KpnI (New England Biolabs - NEB), 100 μg/mL BSA, and 1X NEB buffer 1 overnight in a 37 °C water bath. The following day, the digestion was spiked with 10 μL digestion buffer containing 50 units KpnI, 100 μg/mL BSA, and 1X NEB buffer 1 for 6 hr in a 37 °C water bath. Samples were then resolved on a 1X TAE 0.8% agarose gel overnight at 30V, stained in 0.5 μg/mL ethidium bromide for 30 min (post staining, pictures were taken on the ImageQuant 400 (GE Healthcare) UV box with a ruler to accurately determine sizes), denatured for 1 hr in 0.5M NaOH/1.5M NaCl (pH ~12), neutralized for 1 hr in 0.5M Tris/1.5M NaCl (pH 7.4). DNA was transferred to Hybond-N+ nylon membranes (GE Healthcare) in 10X SSC using a neutral transfer by capillary action. Blots were crosslinked with a UV Stratalinker 1800 (Stratagene) with 120 mJoules and prehybridized for 2 hr in ExpressHyb Hybridization Solution (Clontech) at 68 °C. Southern probes were generated by [α-^{32}P] dCTP body labeling 50 ng of DNA template (5' probe 1 and 3' probe 2) with Ready-to-Go DNA Labeling Beads – dCTP (GE Healthcare) and purified with illustra ProbeQuant G-50 Micro Columns (GE Healthcare) according to the manufactures protocol. Body labeled probes are denatured at 100 °C for 10 min and added to fresh blot/hybridization solution and hybridized overnight at 68 °C. After hybridization, blots were washed once in 2X SSC, 0.1% SDS at 50 °C, once in 2X SSC, 0.1% SDS, twice in 0.5X SSC, 0.1% SDS, and

once in 0.05X SSC, 0.1% SDS (for 30 min at each wash). Blots were exposed to Biomax Film and X-Omatic Intensifying Screens (Kodak) at -80 °C.

Mbnl3$^{\Delta E2/Y}$ mice were genotyped using 100 ng genomic DNA template in a 50 μL reaction containing 1X High Fidelity buffer (5 Prime), 0.4 mM dNTPs, 30 pmol forward primer (MSS3377), 30 pmol forward primer (MSS4153), 30 pmol reverse primer (MSS3366), and 2.5 U Triplemaster Taq (5 Prime). PCR reaction conditions were 96 °C for 2 min (denaturation step), 96 °C for 30 sec – 68 °C for 1 min 30 sec (40X cycles – amplification), followed by 68 °C for 5 min (final elongation step) (see Figure 4-1 for primer identification, position, and sequence). PCR products were resolved on a 1.5% agarose gel, stained with ethidium bromide, and pictures taken on the ImageQuant 400 (GE Healthcare). Expected sizes of bands: wild type *Mbnl3* allele = 350 bp, *Mbnl3*$^{\Delta E2}$ allele = 597 bp.

Skeletal Preparations

Postnatal day 1 mouse pups were asphyxiated with CO_2, skinned and eviscerated. Cartilage was stained by incubating embryos in 0.2 g/L Alcian Blue (in 70% EtOH/30% acetic acid) overnight at room temperature with gentle agitation. The following day, embryos underwent serial rehydration in 100%, 95%, 70% and 40% EtOH and water for 30 min each with gentle agitation. Bone was stained by in 0.1% w/v Alizarin Red (in 1% KOH) overnight at room temperature. Stained embryos incubated for 1 hr each in 25% glycerol/75% KOH, 50% glycerol/50% KOH, 75% glycerol/25% KOH, and 80% glycerol to clear the embryo. Stained and cleared embryos were stored in 80% glycerol and photographed using a Leica stereoscope.

Wild Type and *Mbnl3*$^{\Delta E2/Y}$ Growth Curve

Wild Type and *Mbnl3*$^{\Delta E2/Y}$ male mice were weighed at one week intervals from 3 weeks (weaning) to 10 weeks to assay for postnatal growth (wild type n = 26, *Mbnl3*$^{\Delta E2/Y}$ n = 17). Results were expressed as mean weight for each group ± SEM at intervals of one week. Statistical significance was established using a one tailed Student's T-test with unequal variance.

Sectioning, Immunoflourescence, and H&E Staining

Forelimbs from postnatal day 1 pups were fixed overnight in 4% formaldehyde in PBS (pH 7.4) at 4 °C with agitation. The following day, forelimbs were washed three times (10 min each) in PBS (pH 7.4) and serially dehydrated in 70% ethanol (3X 20 min), 95% ethanol (3X 20 min), 100% ethanol (3X 20 min), CitriSolv (3X 20 min - Fisher Scientific). Tissues were then placed in pre-heated (65 °C) paraffin wax (Fisher Scientific) for 3X 1 hr in a 65 °C vacuum (10-15 mm Hg). Forelimbs were then placed perpendicular (with the foot oriented towards the bottom of the mold) in paraffin molds and allowed to solidify overnight at room temperature. For sectioning, 7 μm sections were take on a rotary microtome (Leica) and mounted on Superfrost slides (Fisher) and incubated at 65 °C for 1 hr in a hybridization oven to remove the paraffin and attach the sections to the slide. Sections were then rehydrated by serial incubations in CitriSolv, 100% ethonaol, 95% ethanol, 70% ethanol and d_2H_2O for 10 min each. At this point, sections were further processed for H&E or immunohistochemical analysis.

TA muscle from skeletal muscles from Notexin regeneration experiments were cut (10 μm) on a cryostat (Leica) and attached to Superfrost slides (Fisher) and stored at -20 °C. For immunohistochemical analysis, sections were thawed at RT for 30 min, fixed

in 2% formaldehyde in TBS (pH 7.5) at RT for 10 min, rinsed in TBS (pH 7.5) and permeablized in 0.1% TritonX-100 in TBS (pH 7.5) for 10 min. At this point, sections were further processed for H&E or immunohistochemical analysis.

For H&E analysis, sections were stained in Harris Hematoxylin (diluted 1:2 in water - Fisher), rinsed in running dH_2O, and dehydrated in 95% ethanol for 1 min each. Sections were then stained in Eosin-Y (Thermo Scientific) for 30 sec, followed by 3X 1 min 95% ethanol , 3X 2 min 100% ethanol, 3X 5 min CitriSolv dehydrations steps and mounted in Permount (Fisher). Pictures were taken on a Leica DM 2000 light microscope. For immunoflourescence analysis, sections were processed using the M.O.M (mouse on mouse) kit according to the manufacturers protocol. C2C12 cells and Huh7 cells were grown on 2-well Falcon slides (BD Biosciences), washed in TBS (pH 7.5), fixed in 2% formaldehyde in TBS (pH 7.5) for 10 min at RT, washed 3X 5 min in TBS (pH 7.5), and blocked/permeablized in 3% heat inactivated goat serum (Invitrogen), 0.1% TritonX-100 in TBS (pH 7.5) for 30 min at RT. Primary antibodies (in blocking/permeablization buffer) were incubated overnight at 4 °C in a humidified chamber. The following day, section were washed 3X 5 min in TBS (pH 7.5), incubated in secondary antibody (in blocking/permeablization buffer) for 2 hr at RT in a humidified chamber, and washed 3X 5 min in TBS (pH 7.5). All sections were counterstained with DAPI mounting media and pictures were taken on a Leica confocal microscope. Antibodies used for staining include, Mbnl3 purified C (1:1000), myc (1:1000 - 9E10), MHC (1:1000 - Sigma), Desmin (1:100 - Abcam), GW182 (1:6000 - Gift of Edward Chan). Secondary antibodies (used at 1:200) include α-rabbit IgG Alexaflour 488 (Mbnl3), α-mouse IgG Alexaflour 488 (Desmin, MHC), α-mouse IgG Alexaflour 568

(myc), α-human IgG Alexaflour 568 (GW182). For staining of the extracellular matrix, TRITC conjugated WGA (Sigma) was incubated at RT for 1 hr (50 μg/mL) in TBS (pH 7.5).

MBNL1

Transfections and RNA Analysis

For *α-actinin* minigene analysis, COSM6 cells were grown to 40%–60% confluency in DMEM (Invitrogen), 10% FBS (Invitrogen), 1% penicillin/streptomycin (Invitrogen) at 37 °C, 5% CO_2 and were co-transfected with 2 μg of *Drosophila α-actinin* minigene and 100 ng of GFP-Mbl isoforms in the pMV vector using Lipofectamine 2000 (Invitrogen) and Optimem (Invitrogen) according to the manufacturers protocol. Four hr after transfection, the media was changed to antibiotic-free DMEM media supplemented with 10% FBS (Invitrogen). To analyze the *Drosophila α-actinin* splicing pattern, total RNA was extracted from transfected COSM6 cells 48 hr after transfection with Tri-Reagent (Sigma) according to the manufacturers protocol. For RT, 5 μg of total RNA was treated with DNAse I (Invitrogen) and RT was performed with Superscript II (Invitrogen), random hexamers and oligo $dT_{12\text{-}18}$ (Invitrogen) following instructions from the manufacturer (Invitrogen). PCR analysis was done with 4 μL of the RT reaction as template in a standard PCR reaction. To detect spliced isoforms arising from the *Drosophila α-actinin* minigene we used primers MSS1938 and MSS1956 (see Figure 4-1 for primer identification, position, and sequence). PCR products were purified by NH_4Ac precipitation and resuspended in d_2H_2O. Half of the resuspended volume was digested with SacI (NEB). The remaining PCR products, and entire digestions, were resolved on a 2% agarose gel, stained with ethidium bromide, and pictures were taken

on the ImageQuant 400 (GE Healthcare). Figure 2-4 shows representative blots from three independent experiments. For GFP-Mbl localization experiments, COSM6 cells were transfected with either 1 μg of plasmids expressing GFP-tagged Mbl proteins and 1 μg carrier DNA (pSP72) or 1 μg plasmid expressing *DMPK*-$(CUG)_{300}$ and 1 μg plasmid expressing GFP-tagged Mbl proteins as previously described.

For *cTNT*, *IR*, and *ClaLC* splicing analysis, COSM6 or HEK293 cells were plated at $5X10^5$ cells per well in a six-well plate in DMEM (Invitrogen) supplemented with 10% FBS (Invitrogen) and 1X penicillin/streptomycin (Invitrogen). At 24 hrs after plating, the cells were transfected with either 1 μg of minigene and 2 μg of protein expression plasmid or 100 ng of minigene, and 100 ng, 300 ng and 1 μg of $(CUG)_{960}$ or $(CAG)_{960}$ repeat expressing plasmid using Fugene 6 (Roche) according to the manufacturer's directions. The total amount of plasmid transfected per well was 2 μg (pSP72 plasmid was used as a carrier to bring the total DNA to 2 μg). Protein and RNA were harvested 36–48 hr after transfection. Chicken primary muscle cultures were prepared, maintained and transfected as previously described, using 0.5 μg minigene reporter and 1 μg expression plasmid (Xu et al., 1993). RNA isolation and RT–PCR analysis for the *cTNT*, *IR*, and *ClaLC* minigenes were performed as described previously (Philips et al., 1998; Savkur et al., 2001; Stamm et al., 1999).

In Vitro Transcription and UV Crosslinking

Uniformly ^{32}P-labeled RNAs were *in vitro* transcribed using [α-^{32}P]GTP and [α-^{32}P]UTP (Perkin-Elmer) from PCR products or cloned regions of the human *cTNT* introns 4 and 5, as represented in Figure 2-8. UV-crosslinking assays were performed using radiolabeled transcripts standardized for picomoles of G and U. UV-crosslinking

assays included 1 μg of purified GST–MBNL1 in the presence of 1 μg BSA, 1 μg tRNA, 0.3 μg heparin, 0.3 mM magnesium acetate, in 2 mM magnesium acetate, 2 mM ATP, 16 mM HEPES (pH 7.9), 65 mM potassium glutamate, 0.16 mM EDTA, 0.4 mM DTT and 16% glycerol. Binding was for 10 min at 30 °C. Recombinant GST–MBNL1 protein was produced as described (Miller et al., 2000).

Plasmids

The *cTNT*, *IR* and *ClaLC* minigenes were previously described (Kosaki et al., 1998; Philips et al., 1998; Stamm et al., 1999). GFP fusions with MBNL1, 2 and 3 were provided by Dr. JD Brook (Fardaei et al., 2002). GFP–MBNL1 was found to have a novel MBNL1 isoform lacking exons 7, 9 and 10 and containing a frameshift in exon 12. Plasmids expressing DMPK exons 11–15 containing 960 interrupted CUG or CAG repeats in exon 15 were cloned using techniques as previously described (Philips et al., 1998). The MBNL mutant human cTNT minigene was generated by inverse PCR. *α-actinin* minigene and *GFP-Mbl* expression vector were constructed as previously described (Vicente et al., 2007).

Transfection of siRNA

Two custom siRNA duplexes were designed for RNAi against human MBNL1 using the Dharmacon siDESIGN program (www.dharmacon.com), and were synthesized by Dharmacon. The siRNA sequences (5' to 3') are as follows: THH31 (AACAGACAGACUUGAGGUAUG), THH2 (AACACGGAAUGUAAAUUUGCA), GFP siRNA duplex (Dharmacon). Prior to transfections, $3X10^5$ HeLa cells were plated in 2 mL of antibiotic-free growth media (DMEM supplemented with 10% FBS) per well in a six-well plate. At 12 hr after plating, the media was exchanged with 800 μL serum-free

media (DMEM) per well, and siRNA duplex (2.66 μg) was transfected using Oligofectamine (Invitrogen) according to the manufacturers protocol. After 4 hr 1 mL of 3X serum-containing media (DMEM supplemented with 30% FBS) was added to each well and returned to the incubator. After 12 hr, the 3X serum-containing media was replaced with antibiotic-free growth media and the cells were transfected with 1 μg of minigene and 2.66 μg of siRNA duplex using Lipofectamine 2000 (Invitrogen). The media was exchanged with antibiotic-free growth media 6 hr later. RNA and protein were harvested 48 hr after transfection of the minigene.

Immunoblot Analysis

Cells were harvested in protein loading buffer (62.5 mM Tris–HCl (pH 6.8), 2% SDS, 10% glycerol and 5% 2-β-mercaptoethanol) and the protein concentration was quantitated with the Non-Interfering Protein Assay (Genotech). Total protein lysates from HEK293 (20 μg) and primary chicken skeletal (30 μg) cultures were loaded on a 12.5% acrylamide gel and transferred to Immobilon-P membranes (Millipore). GFP was detected using JL-8 monoclonal antibody (BD Biosciences) at a dilution of 1:2000. The secondary antibody was a goat anti-mouse HRP conjugate (Jackson Immunoresearch) at a dilution of 1:5000. To detect endogenous MBNL1, HeLa (50 μg) protein lysates were loaded on a 12.5% acrylamide gel. Blots were probed with the monoclonal 3A4 (16 mg/mL) at a dilution of 1:500. The secondary antibody was a sheep anti-mouse HRP conjugate (Amersham Biosciences) at a dilution of 1:5000. For GAPDH in HeLa cells, 15 μg of total protein lysates was run on a 12.5% acrylamide gel, transferred to membranes and detected using the 6G5 monoclonal (Biogenesis) at a dilution of 1:100, 000. The secondary antibody was a goat anti-mouse HRP conjugate (Jackson

Immunoresearch) at a dilution of 1:5000. Blots were developed in Amersham ECL® detection reagents (GE Healthcare) or Supersignal West Femto reagents (Thermo Scientific – Pierce) and exposed to Biomax Light Film (Kodak).

Flourescent *In Situ* Hybridization and Immunocytochemistry

COSM6 cells ($1.2X10^5$) were seeded on 2-well Falcon culture slides (BD Biosciences) and transfected 24 hr after plating. Cells were washed with 1X Hanks Balanced Salt solution and then with CSK buffer (300 mM sucrose, 100 mM NaCl, 3 mM MgCl2, 1.2 mM PMSF, 10 mM PIPES, pH 6.8). The slides were incubated on ice in CSK buffer with 0.5% Triton X-100 and 10 mM vanadyl sulfate for 1 min. They were then fixed in 4% paraformaldehyde/PBS (pH 7.4) for 10 min at room temperature and washed with 70% ethanol. Cells were dehydrated with 70% ethanol overnight at 4 °C and then rehydrated with 40% formamide/2XSSC for 10 min at room temperature the following day. Cy5-labeled oligonucleotide probes (Qiagen) $(CAG)_{10}$ and $(CTG)_{10}$ were used to detect CUG and CAG repeats, respectively. The cells were incubated in probe/hybridization buffer (40% formamide, 2XSSC, 0.2% BSA, 10% dextran sulfate, 2 mM vanadyl sulfate, 1 mg/mL yeast tRNA, 50 ng/mL probe) in a humidified chamber at 37 °C for 2 hr. After hybridization, the slides were washed three times with 40% formamide/2XSSC for 30 min at 37 °C, followed by three PBS (pH 7.4) washes for 5 min each. They were then pre-blocked in 3% BSA/PBS for 15 min in a humidified chamber at room temperature followed by a wash with PBS (pH 7.4). Afterwards, the slides were incubated with primary anti-MBNL1 antibody 3A4 (10 mg/mL; 1:1000 dilution) in 3% BSA/PBS at room temperature for 1 hr in a humidified chamber, washed three times with PBS (pH 7.4) and incubated with the secondary antibody Alexa Fluor488-labeled goat anti-mouse IgG (2 mg/mL, Molecular Probes) at a dilution of

1:100 in 3% BSA/PBS (pH 7.4) at room temperature for 1 hr. Following the incubation, cells were washed three times with PBS (pH 7.4) and nuclei were stained with DAPI using Vectashield mounting media (Vector Laboratories, Inc). GFP-Mbl proteins were analyzed directly. Cells were analyzed under epifluorescence microscopy using an Axioskop2 mot plus microscope (Carl Zeiss, Inc).

Competition Assay

Whole cell lysate was generated using myc-tagged MBNL1 (41 kDa isoform) transiently transfected in COSM6 (cells were transfected using 2 μg expression vector as previously described). To prepare whole cell lysate for competition assays, 5X 10 cm tissues culture plates containing COSM6 cells ectopically expressing myc-MBNL1 were resuspended in 250 mL/plate in cold crosslinking buffer (20 mM HEPES-KOH, pH 8.0, 100 mM KCl, 0.1% Igepal, 1X PicD/PicW), sonicated 3X 40% amplitude for 5 sec, spun down at 16.1 RCF for 20 min at 4 °C. Following the spin, supernatant was transferred to a new microfuge on ice and glycerol was added to a final concentration of 20%. Whole cell lysate was aliquoted and stored at -80 °C until use.

To generate uniformly body-labeled repeat RNAs, ~1 μg linearized plasmid was transcribed with T7 RNA polymerase (Promega) in a 50 μL with [α-^{32}P]GTP according to the manufacturers protocol. After *in vitro* transcription, DNA template was digested by RQ1 RNase-free DNase (1 unit/μL - Promega) in the presence of ~1 unit/μL RNasin (Promega) and ~0.5 μg/μL yeast tRNA (Invitrogen) for 15 min at 37 °C. Following digestion, RNA is phenol:chloroform extracted, ethanol precipitated, resuspended in 10 μL RNase-free water, and resolved on a 5% acrylamide-urea gel in 1X TBE (run at 200V for 2 hr in a vertical gel). RNA bands were excised and extracted by disrupting

the acrylamide gel slices with a sterile pestle and incubation in extraction buffer (0.5 M NH_4OAc, 10 mM $Mg(OAc)_2$, 0.1 mM EDTA, 0.1% SDS) at 65 °C for 20 min (in a thermo mixer - Eppendorf). Extraction was spun down for 1 min at 16.1 RCF at RT, supernatant was passed through a mini-spin column (USA Scientific) and 0.45 μm filter (Ultra Free MC Filter - Millipore), ethanol precipitated, and stored at -80 °C until use. To generate non-labeled competitor, ~2.5 μg linearized plasmid was transcribed using the MEGAshortscript kit (Ambion) and cleaned-up using the MEGAclear kit (Ambion) according to the manufacturers protocol and stored at -80 °C until use.

Competition assays were done with 0.1 pmol uniformly body-labeled RNA in a 35 μL reaction composed of 40 mM HEPES-KOH (pH 8.0), 2 mM $Mg(OAc)_2$, 20 mM creatine phosphate, 2 mM ATP, 65 mM K-glutamate, 0.4 mM DTT, 0.16 mM EDTA, 100 mM KCl, and 15 μL myc-tagged MBNL1 whole cell lysate at 30 °C. Incubation times and introduction of 200 pmol cold competitor RNA (2000X fold excess) were done as indicated in Figure 2-13. Following competition, reactions were UV crosslinked on ice 3X 2.5 min in a UV Stratalinker 1800 (Stratagene). Excess RNA was digested by incubating crosslinked reactions in ~0.5 μg/μL RNase A (Sigma) for 20 min at 37 °C. Protein-A Sepharose (50 μL - Invitrogen) beads were washed two times by adding 500 μL IPP-150 (50 mM Tris, pH 7.4, 150 mM NaCl, 0.1% Igepal),inverting several times in a microfuge tube, flash spinning (5 sec) in a table top centrifuge, pipeting off the supernatant, and repeating. Protein-A Sepharose was resuspended in 300 μL IPP-150 with 1 μL α-myc (9E10 - 10 mg/mL) and rotated for 1 hr at 4 °C. Following antibody capture, protein-A Sepharose:α-myc was washed three times as described above, resuspended in 300 μL IPP-150, and incubated with crosslinked competition reactions

for 1 hr rotating at 4 °C to immunoprecipitate myc-MBNL1:^{32}P-RNA. Following immunoprecipitation, complexes were washed five times as described above (four washes with IPP-150 and one wash with IPP-150 without Igepal), resuspended in 35 μL 2X Laemmli buffer, resolved by 12.5% SDS-PAGE. The gel was fixed for 20 min in destain buffer (10% acetic acid, 20% methanol), dried down using the Hoefer slab gel drier (Amersham), and exposed overnight to Biomax Film and X-Omatic Intensifying Screens (Kodak) at -80 °C. Band intensity was measured using a phosphoimaging screen (Amersham) and the Typhoon 9200 (Amersham).

Primer I.D.	Gene	Location	Primer Sequence (5' to 3')
MSS 2916	Mbnl3	Exon 2 Fwd	gctctcagtatgacacctgtcaatgtagc
MSS 3648	Mbnl3	Exon 1 Fwd	gctggagtcgctcactcggagag
MSS 3655	Mbnl3	Exon 3 Fwd	ccagcacatgcaactcatgctgc
MSS 3759	Mbnl3	Exon 7a Rev	caaaagttgaagtaggagtcatgcacag
MSS 3760	Mbnl3	Exon 7b Rev	ctggtctgcatataaaagaacagtgg
MSS 3763	Mbnl3	Exon 7c Rev	gctgtagacacagtggaaggtgtagcac
MSS 3652	Mbnl3	Exon 8 Rev	tgactgcaccatgcttattgtcgtg
MSS 2725	Mbnl1	Exon 4 Fwd	gggcggaataacttgattcagcag
MSS 2726	Mbnl1	Exon 13 Rev	aaagtgtgtaggcaacagtgtgcagc
MSS 4173	Myogenin	Exon 1 Fwd	accaggagccccacttctatgatgg
MSS 4177	Myogenin	Exon 3 Rev	acatatcctccaccgtgatgctgtc
MSS 3225	Neomycin	CDS Fwd	ggagaggctattcggctatgactgg
MSS 3247	Neomycin	CDS Rev	gctcttcagcaatatcacgggtagc
MSS 3586	Ppia	Exon 4 Fwd	gcggcaggtccatctacg
MSS 3587	Ppia	Exon 4/5 Rev	gccatccagccattcagtct
MSS 3580	Mbnl1	Exon 3 Fwd	cgcggatccaacatggctgttagtgtcacaccaattcg
MSS 3581	Mbnl1	Exon 13 Rev	gctctagacatctgggtaacatacttgtggctagtcag
MSS 3582	Mbnl2	Exon 2 Fwd	cgcggatccatcatggccttgaacgttgcccccgtgag
MSS 3583	Mbnl2	Exon 9 Rev	gctctagatttcagaattatctgattggctgtggctg
MSS 3584	Mbnl3	Exon 2 Fwd	cgcggatccagtatgacacctgtcaatgtagctctaatccgtg
MSS 3585	Mbnl3	Exon 8 Rev	gctctagaatatttcaactggttgcctgtagttgttgg
MSS1938	pSG vector	backbone Fwd	gctgcaataaacaagttctgcttt
MSS1956	pSG vector	backbone Rev	agaattgtaatacgactcactatagggc
MSS3377	Mbnl3	Intron 1 Fwd	ccgccttgcttgtagttcactgtcagcaactgg
MSS4153	Mbnl3	Intron 2 Fwd	gccttgagtttgtagtaacccttatgcctttgccttctgag
MSS3366	Mbnl3	Intron 2 Rev	ggagagatggctcagtgatgaagaacacttgttgctc
MSS3256	Mbnl3	Intron 1 Fwd	cctccaaactacaagcactcagaacccgctgg
MSS3272	Mbnl3	Intron 1 Rev	ctgacactatcggtaggtctatgtcttccactaatgtcagtgc
MSS3258	Mbnl3	Intron 2 Fwd	cctaaggaaaggacttgacttcagtttctccaactggtctcc
MSS3259	Mbnl3	Intron 2 Rev	accaggtgcccaattacaatttggctgcaaggtg

Figure 4-1. Primers used for RT-PCR, *Mbnl3*$^{\Delta E2/+}$ genotyping, probe generation, and subcloning.

LIST OF REFERENCES

Accardi, A., and Pusch, M. (2000). Fast and slow gating relaxations in the muscle chloride channel CLC-1. J Gen Physiol *116*, 433-444.

Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K., and Edstrom, L. (1993). Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. Hum Mol Genet *2*, 1397-1400.

Artero, R., Prokop, A., Paricio, N., Begemann, G., Pueyo, I., Mlodzik, M., Perez-Alonso, M., and Baylies, M. K. (1998). The muscleblind gene participates in the organization of Z-bands and epidermal attachments of Drosophila muscles and is regulated by Dmef2. Dev Biol *195*, 131-143.

Ashizawa, T., Dubel, J. R., and Harati, Y. (1993). Somatic instability of CTG repeat in myotonic dystrophy. Neurology *43*, 2674-2678.

Avaria, M., and Patterson, V. (1994). Myotonic dystrophy: relative sensitivity of symptoms signs and abnormal investigations. Ulster Med J *63*, 151-154.

Barchi, R. L. (1995). Molecular pathology of the skeletal muscle sodium channel. Annu Rev Physiol *57*, 355-385.

Begemann, G., Michon, A. M., vd Voorn, L., Wepf, R., and Mlodzik, M. (1995). The Drosophila orphan nuclear receptor seven-up requires the Ras pathway for its function in photoreceptor determination. Development *121*, 225-235.

Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M., and Mlodzik, M. (1997). muscleblind, a gene required for photoreceptor differentiation in Drosophila, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. Development *124*, 4321-4331.

Belfiore, A., Frasca, F., Pandini, G., Sciacca, L., and Vigneri, R. (2009). Insulin receptor isoforms and insulin receptor/insulin-like growth factor receptor hybrids in physiology and disease. Endocr Rev *30*, 586-623.

Benders, A. A., Timmermans, J. A., Oosterhof, A., Ter Laak, H. J., van Kuppevelt, T. H., Wevers, R. A., and Veerkamp, J. H. (1993). Deficiency of Na+/K(+)-ATPase and sarcoplasmic reticulum Ca(2+)-ATPase in skeletal muscle and cultured muscle cells of myotonic dystrophy patients. Biochem J *293 (Pt 1)*, 269-274.

Berg, J., Jiang, H., Thornton, C. A., and Cannon, S. C. (2004). Truncated ClC-1 mRNA in myotonic dystrophy exerts a dominant-negative effect on the Cl current. Neurology *63*, 2371-2375.

Bergen, B. J. (1985). Evaluation of the hypotonic or floppy infant. Minn Med *68*, 341-347.

Berul, C. I., Maguire, C. T., Aronovitz, M. J., Greenwood, J., Miller, C., Gehrmann, J., Housman, D., Mendelsohn, M. E., and Reddy, S. (1999). DMPK dosage alterations result in atrioventricular conduction abnormalities in a mouse myotonic dystrophy model. J Clin Invest *103*, R1-7.

Berul, C. I., Maguire, C. T., Gehrmann, J., and Reddy, S. (2000). Progressive atrioventricular conduction block in a mouse myotonic dystrophy model. J Interv Card Electrophysiol *4*, 351-358.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem *72*, 291-336.

Bodensteiner, J. B. (2008). The evaluation of the hypotonic infant. Semin Pediatr Neurol *15*, 10-20.

Botta, A., Caldarola, S., Vallo, L., Bonifazi, E., Fruci, D., Gullotta, F., Massa, R., Novelli, G., and Loreni, F. (2006). Effect of the [CCTG]n repeat expansion on ZNF9 expression in myotonic dystrophy type II (DM2). Biochim Biophys Acta *1762*, 329-334.

Botta, A., Rinaldi, F., Catalli, C., Vergani, L., Bonifazi, E., Romeo, V., Loro, E., Viola, A., Angelini, C., and Novelli, G. (2008). The CTG repeat expansion size correlates with the splicing defects observed in muscles from myotonic dystrophy type 1 patients. J Med Genet *45*, 639-646.

Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., Hunter, K., Stanton, V. P., Thirion, J. P., Hudson, T., and 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*, 799-808.

Buckingham, M., Bajard, L., Chang, T., Daubas, P., Hadchouel, J., Meilhac, S., Montarras, D., Rocancourt, D., and Relaix, F. (2003). The formation of skeletal muscle: from somite to limb. J Anat *202*, 59-68.

Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., de Jong, P., Jansen, G., Anvret, M., Riley, B., and et al. (1992). Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature *355*, 547-548.

Campbell, C., Sherlock, R., Jacob, P., and Blayney, M. (2004). Congenital myotonic dystrophy: assisted ventilation duration and outcome. Pediatrics *113*, 811-816.

Chari, A., Paknia, E., and Fischer, U. (2009). The role of RNP biogenesis in spinal muscular atrophy. Curr Opin Cell Biol *21*, 387-393.

Charlet, B. N., Savkur, R. S., Singh, G., Philips, A. V., Grice, E. A., and Cooper, T. A. (2002). Loss of the muscle-specific chloride channel in type 1 myotonic dystrophy due to misregulated alternative splicing. Mol Cell *10*, 45-53.

Chen, M., and Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol *10*, 741-754.

Chen, W., Wang, Y., Abe, Y., Cheney, L., Udd, B., and Li, Y. P. (2007). Haploinsuffciency for Znf9 in Znf9+/- mice is associated with multiorgan abnormalities resembling myotonic dystrophy. J Mol Biol *368*, 8-17.

Cho, D. H., Thienes, C. P., Mahoney, S. E., Analau, E., Filippova, G. N., and Tapscott, S. J. (2005). Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol Cell *20*, 483-489.

Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H., and Housman, D. E. (1997). Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. Proc Natl Acad Sci U S A *94*, 7388-7393.

Dixon, R. W., and Harris, J. B. (1996). Myotoxic activity of the toxic phospholipase, notexin, from the venom of the Australian tiger snake. J Neuropathol Exp Neurol *55*, 1230-1237.

Du, H., Cline, M. S., Osborne, R. J., Tuttle, D. L., Clark, T. A., Donohue, J. P., Hall, M. P., Shiue, L., Swanson, M. S., Thornton, C. A., and Ares, M., Jr. (2010). Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. Nat Struct Mol Biol *17*, 187-193.

Echenne, B., Rideau, A., Roubertie, A., Sebire, G., Rivier, F., and Lemieux, B. (2008). Myotonic dystrophy type I in childhood Long-term evolution in patients surviving the neonatal period. Eur J Paediatr Neurol *12*, 210-223.

Eriksson, M., Hedberg, B., Carey, N., and Ansved, T. (2001). Decreased DMPK transcript levels in myotonic dystrophy 1 type IIA muscle fibers. Biochem Biophys Res Commun *286*, 1177-1182.

Fardaei, M., Rogers, M. T., Thorpe, H. M., Larkin, K., Hamshere, M. G., Harper, P. S., and Brook, J. D. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. Hum Mol Genet *11*, 805-814.

Farkas-Bargeton, E., Barbet, J. P., Dancea, S., Wehrle, R., Checouri, A., and Dulac, O. (1988). Immaturity of muscle fibers in the congenital form of myotonic dystrophy: its consequences and its origin. J Neurol Sci *83*, 145-159.

Filippova, G. N., Thienes, C. P., Penn, B. H., Cho, D. H., Hu, Y. J., Moore, J. M., Klesert, T. R., Lobanenkov, V. V., and Tapscott, S. J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet *28*, 335-343.

Fiset, C., and Giles, W. R. (2008). Cardiac troponin T mutations promote life-threatening arrhythmias. J Clin Invest.

Fu, Y. H., Friedman, D. L., Richards, S., Pearlman, J. A., Gibbs, R. A., Pizzuti, A., Ashizawa, T., Perryman, M. B., Scarlato, G., Fenwick, R. G., Jr., and et al. (1993). Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. Science *260*, 235-238.

Fu, Y. H., Pizzuti, A., Fenwick, R. G., Jr., King, J., Rajnarayan, S., Dunne, P. W., Dubel, J., Nasser, G. A., Ashizawa, T., de Jong, P., and et al. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science *255*, 1256-1258.

Fuerst, P. G., Bruce, F., Tian, M., Wei, W., Elstrott, J., Feller, M. B., Erskine, L., Singer, J. H., and Burgess, R. W. (2009). DSCAM and DSCAML1 function in self-avoidance in multiple cell types in the developing mouse retina. Neuron *64*, 484-497.

Fyrberg, E., Kelly, M., Ball, E., Fyrberg, C., and Reedy, M. C. (1990). Molecular genetics of Drosophila alpha-actinin: mutant alleles disrupt Z disc integrity and muscle insertions. J Cell Biol *110*, 1999-2011.

Geeves, M. A., and Holmes, K. C. (1999). Structural mechanism of muscle contraction. Annu Rev Biochem *68*, 687-728.

Goers, E. S., Purcell, J., Voelker, R. B., Gates, D. P., and Berglund, J. A. (2010). MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. Nucleic Acids Res.

Groh, W. J., Lowe, M. R., and Zipes, D. P. (2002). Severity of cardiac conduction involvement and arrhythmias in myotonic dystrophy type 1 correlates with age and CTG repeat length. J Cardiovasc Electrophysiol *13*, 444-448.

Gunning, P., O'Neill, G., and Hardeman, E. (2008). Tropomyosin-based regulation of the actin cytoskeleton in time and space. Physiol Rev *88*, 1-35.

Hamshere, M. G., Newman, E. E., Alwazzan, M., Athwal, B. S., and Brook, J. D. (1997). Transcriptional abnormality in myotonic dystrophy affects DMPK but not neighboring genes. Proc Natl Acad Sci U S A *94*, 7394-7399.

Hao, M., Akrami, K., Wei, K., De Diego, C., Che, N., Ku, J. H., Tidball, J., Graves, M. C., Shieh, P. B., and Chen, F. (2008). Muscleblind-like 2 (Mbnl2) -deficient mice as a model for myotonic dystrophy. Dev Dyn *237*, 403-410.

Harley, H. G., Brook, J. D., Rundle, S. A., Crow, S., Reardon, W., Buckler, A. J., Harper, P. S., Housman, D. E., and Shaw, D. J. (1992). Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature *355*, 545-546.

Harper, P. S. (1974). Proceedings: Congenital myotonic dystrophy in Britain. Arch Dis Child *49*, 820.

Harper, P. S. (1975). Congenital myotonic dystrophy in Britain. I. Clinical aspects. Arch Dis Child *50*, 505-513.

Harper, P. S. (2001). Myotonic Dystrophy Third Edition: Harcourt Publishers).

Harris, J. B., and MacDonell, C. A. (1981). Phospholipase A2 activity of notexin and its role in muscle damage. Toxicon *19*, 419-430.

Hattori, D., Chen, Y., Matthews, B. J., Salwinski, L., Sabatti, C., Grueber, W. B., and Zipursky, S. L. (2009). Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. Nature *461*, 644-648.

He, Y., and Smith, R. (2009). Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. Cell Mol Life Sci *66*, 1239-1256.

Hedberg, B., Anvret, M., and Ansved, T. (1999). CTG-repeat length in distal and proximal leg muscles of symptomatic and non-symptomatic patients with myotonic dystrophy: relation to muscle strength and degree of histopathological abnormalities. Eur J Neurol *6*, 341-346.

Hiromi, Y., Mlodzik, M., West, S. R., Rubin, G. M., and Goodman, C. S. (1993). Ectopic expression of seven-up causes cell fate changes during ommatidial assembly. Development *118*, 1123-1135.

Holt, I., Jacquemin, V., Fardaei, M., Sewry, C. A., Butler-Browne, G. S., Furling, D., Brook, J. D., and Morris, G. E. (2009). Muscleblind-like proteins: similarities and differences in normal and myotonic dystrophy muscle. Am J Pathol *174*, 216-227.

Hudson, A. J., Ebers, G. C., and Bulman, D. E. (1995). The skeletal muscle sodium and chloride channel diseases. Brain *118 (Pt 2)*, 547-563.

Iannaccone, S. T., Bove, K. E., Vogler, C., Azzarelli, B., and Muller, J. (1986). Muscle maturation delay in infantile myotonic dystrophy. Arch Pathol Lab Med *110*, 405-411.

Imarisio, S., Carmichael, J., Korolchuk, V., Chen, C. W., Saiki, S., Rose, C., Krishna, G., Davies, J. E., Ttofi, E., Underwood, B. R., and Rubinsztein, D. C. (2008). Huntington's disease: from pathology and genetics to potential therapies. Biochem J *412*, 191-209.

Inukai, A., Doyu, M., Kato, T., Liang, Y., Kuru, S., Yamamoto, M., Kobayashi, Y., and Sobue, G. (2000). Reduced expression of DMAHP/SIX5 gene in myotonic dystrophy muscle. Muscle Nerve *23*, 1421-1426.

Jensen, C. J., Oldfield, B. J., and Rubio, J. P. (2009). Splicing, cis genetic variation and disease. Biochem Soc Trans *37*, 1311-1315.

Jin, J. P., Zhang, Z., and Bautista, J. A. (2008). Isoform diversity, regulation, and functional adaptation of troponin and calponin. Crit Rev Eukaryot Gene Expr *18*, 93-124.

Jurkat-Rott, K., Lerche, H., and Lehmann-Horn, F. (2002). Skeletal muscle channelopathies. J Neurol *249*, 1493-1502.

Kahn, C. R., and White, M. F. (1988). The insulin receptor and the molecular mechanism of insulin action. J Clin Invest *82*, 1151-1156.

Kanadia, R. N., Johnstone, K. A., Mankodi, A., Lungu, C., Thornton, C. A., Esson, D., Timmers, A. M., Hauswirth, W. W., and Swanson, M. S. (2003a). A muscleblind knockout model for myotonic dystrophy. Science *302*, 1978-1980.

Kanadia, R. N., Shin, J., Yuan, Y., Beattie, S. G., Wheeler, T. M., Thornton, C. A., and Swanson, M. S. (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. Proc Natl Acad Sci U S A *103*, 11748-11753.

Kanadia, R. N., Urbinati, C. R., Crusselle, V. J., Luo, D., Lee, Y. J., Harrison, J. K., Oh, S. P., and Swanson, M. S. (2003b). Developmental expression of mouse muscleblind genes Mbnl1, Mbnl2 and Mbnl3. Gene Expr Patterns *3*, 459-462.
Kang, S., and Hong, S. (2009). Molecular pathogenesis of spinocerebellar ataxia type 1 disease. Mol Cells *27*, 621-627.

Kellerer, M., Lammers, R., Ermel, B., Tippmer, S., Vogt, B., Obermaier-Kusser, B., Ullrich, A., and Haring, H. U. (1992). Distinct alpha-subunit structures of human insulin receptor A and B variants determine differences in tyrosine kinase activities. Biochemistry *31*, 4588-4596.

Kimura, T., Takahashi, M. P., Okuda, Y., Kaido, M., Fujimura, H., Yanagihara, T., and Sakoda, S. (2000). The expression of ion channel mRNAs in skeletal muscles from patients with myotonic muscular dystrophy. Neurosci Lett *295*, 93-96.

Klesert, T. R., Cho, D. H., Clark, J. I., Maylie, J., Adelman, J., Snider, L., Yuen, E. C., Soriano, P., and Tapscott, S. J. (2000). Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. Nat Genet *25*, 105-109.

Klesert, T. R., Otten, A. D., Bird, T. D., and Tapscott, S. J. (1997). Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. Nat Genet *16*, 402-406.

Kosaki, A., Nelson, J., and Webster, N. J. (1998). Identification of intron and exon sequences involved in alternative splicing of insulin receptor pre-mRNA. J Biol Chem *273*, 10331-10337.

Koshy, B. T., and Zoghbi, H. Y. (1997). The CAG/polyglutamine tract diseases: gene products and molecular pathogenesis. Brain Pathol *7*, 927-942.

Krol, J., Fiszer, A., Mykowska, A., Sobczak, K., de Mezer, M., and Krzyzosiak, W. J. (2007). Ribonuclease dicer cleaves triplet repeat hairpins into shorter repeats that silence specific targets. Mol Cell *25*, 575-586.

Kuyumcu-Martinez, N. M., Wang, G. S., and Cooper, T. A. (2007). Increased steady-state levels of CUGBP1 in myotonic dystrophy 1 are due to PKC-mediated hyperphosphorylation. Mol Cell *28*, 68-78.

La Spada, A. R., Paulson, H. L., and Fischbeck, K. H. (1994). Trinucleotide repeat expansion in neurological disease. Ann Neurol *36*, 814-822.

Lacomis, D., Chad, D. A., and Smith, T. W. (1994). Proximal weakness as the primary manifestation of myotonic dystrophy in older adults. Muscle Nerve *17*, 687-688.

Lee, K. S., Squillace, R. M., and Wang, E. H. (2007). Expression pattern of muscleblind-like proteins differs in differentiating myoblasts. Biochem Biophys Res Commun *361*, 151-155.

Licatalosi, D. D., and Darnell, R. B. (2010). RNA processing and its regulation: global insights into biological networks. Nat Rev Genet *11*, 75-87.

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X.*, et al.* (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464-469.

Lin, X., Miller, J. W., Mankodi, A., Kanadia, R. N., Yuan, Y., Moxley, R. T., Swanson, M. S., and Thornton, C. A. (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. Hum Mol Genet *15*, 2087-2097.

Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science *293*, 864-867.

Long, J. C., and Caceres, J. F. (2009). The SR protein family of splicing factors: master regulators of gene expression. Biochem J *417*, 15-27.

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., and et al. (1992). Myotonic dystrophy

mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science *255*, 1253-1255.

Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. Science *289*, 1769-1773.

Mankodi, A., Takahashi, M. P., Jiang, H., Beck, C. L., Bowers, W. J., Moxley, R. T., Cannon, S. C., and Thornton, C. A. (2002). Expanded CUG repeats trigger aberrant splicing of ClC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. Mol Cell *10*, 35-44.

Mankodi, A., Teng-Umnuay, P., Krym, M., Henderson, D., Swanson, M., and Thornton, C. A. (2003). Ribonuclear inclusions in skeletal muscle in myotonic dystrophy types 1 and 2. Ann Neurol *54*, 760-768.

Mankodi, A., Urbinati, C. R., Yuan, Q. P., Moxley, R. T., Sansone, V., Krym, M., Henderson, D., Schalling, M., Swanson, M. S., and Thornton, C. A. (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. Hum Mol Genet *10*, 2165-2170.

Margolis, J. M., Schoser, B. G., Moseley, M. L., Day, J. W., and Ranum, L. P. (2006). DM2 intronic expansions: evidence for CCUG accumulation without flanking sequence or effects on ZNF9 mRNA processing or protein expression. Hum Mol Genet *15*, 1808-1815.

Martorell, L., Johnson, K., Boucher, C. A., and Baiget, M. (1997). Somatic instability of the myotonic dystrophy (CTG)n repeat during human fetal development. Hum Mol Genet *6*, 877-880.

Martorell, L., Monckton, D. G., Gamez, J., and Baiget, M. (2000). Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. Eur J Hum Genet *8*, 423-430.

Matthews, B. J., Kim, M. E., Flanagan, J. J., Hattori, D., Clemens, J. C., Zipursky, S. L., and Grueber, W. B. (2007). Dendrite self-avoidance is controlled by Dscam. Cell *129*, 593-604.

McClain, D. A. (1991). Different ligand affinities of the two human insulin receptor splice variants are reflected in parallel changes in sensitivity for insulin action. Mol Endocrinol *5*, 734-739.

Michalowski, S., Miller, J. W., Urbinati, C. R., Paliouras, M., Swanson, M. S., and Griffith, J. (1999). Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. Nucleic Acids Res *27*, 3534-3542.

Miller, J. W., Urbinati, C. R., Teng-Umnuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., and Swanson, M. S. (2000). Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J *19*, 4439-4448.

Milner-Brown, H. S., and Miller, R. G. (1990). Myotonic dystrophy: quantification of muscle weakness and myotonia and the effect of amitriptyline and exercise. Arch Phys Med Rehabil *71*, 983-987.

Moller, D. E., Yokota, A., Caro, J. F., and Flier, J. S. (1989). Tissue-specific expression of two alternatively spliced insulin receptor mRNAs in man. Mol Endocrinol *3*, 1263-1269.

Monferrer, L., and Artero, R. (2006). An interspecific functional complementation test in Drosophila for introductory genetics laboratory courses. J Hered *97*, 67-73.

Moolman, J. C., Corfield, V. A., Posen, B., Ngumbela, K., Seidman, C., Brink, P. A., and Watkins, H. (1997). Sudden death due to troponin T mutations. J Am Coll Cardiol 29, 549-555.

Mounsey, J. P., Mistry, D. J., Ai, C. W., Reddy, S., and Moorman, J. R. (2000). Skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. Hum Mol Genet 9, 2313-2320.

Moxley, R. T., 3rd, Griggs, R. C., Goldblatt, D., VanGelder, V., Herr, B. E., and Thiel, R. (1978). Decreased insulin sensitivity of forearm muscle in myotonic dystrophy. J Clin Invest 62, 857-867.

Moxley, R. T., Corbett, A. J., Minaker, K. L., and Rowe, J. W. (1984). Whole body insulin resistance in myotonic dystrophy. Ann Neurol 15, 157-162.

Nagy, A., Moens, C., Ivanyi, E., Pawling, J., Gertsenstein, M., Hadjantonakis, A. K., Pirity, M., and Rossant, J. (1998). Dissecting the role of N-myc in development using a single targeting vector to generate a series of alleles. Curr Biol 8, 661-664.

Novelli, G., Gennarelli, M., Zelano, G., Pizzuti, A., Fattorini, C., Caskey, C. T., and Dallapiccola, B. (1993). Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. Biochem Mol Biol Int 29, 291-297.

Ogino, S., and Wilson, R. B. (2004). Spinal muscular atrophy: molecular genetics and diagnostics. Expert Rev Mol Diagn 4, 15-29.

Ohtsuki, I., and Morimoto, S. (2008). Troponin: regulatory function and disorders. Biochem Biophys Res Commun 369, 62-73.

Orengo, J. P., Chambon, P., Metzger, D., Mosier, D. R., Snipes, G. J., and Cooper, T. A. (2008). Expanded CTG repeats within the DMPK 3' UTR causes severe skeletal

muscle wasting in an inducible mouse model for myotonic dystrophy. Proc Natl Acad Sci U S A 105, 2646-2651.

Orengo, J. P., and Cooper, T. A. (2007). Alternative splicing in disease. Adv Exp Med Biol 623, 212-223.

Osborne, R. J., Lin, X., Welle, S., Sobczak, K., O'Rourke, J. R., Swanson, M. S., and Thornton, C. A. (2009). Transcriptional and post-transcriptional impact of toxic RNA in myotonic dystrophy. Hum Mol Genet 18, 1471-1481.

Otten, A. D., and Tapscott, S. J. (1995). Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. Proc Natl Acad Sci U S A 92, 5465-5469.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40, 1413-1415.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2009). Addendum: Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 41, 762.

Pascual, M., Vicente, M., Monferrer, L., and Artero, R. (2006). The Muscleblind family of proteins: an emerging class of regulators of developmentally programmed alternative splicing. Differentiation 74, 65-80.

Pearson, C. E., Nichol Edamura, K., and Cleary, J. D. (2005). Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet 6, 729-742.

Personius, K. E., Nautiyal, J., and Reddy, S. (2005). Myotonia and muscle contractile properties in mice with SIX5 deficiency. Muscle Nerve 31, 503-505.

Philips, A. V., Timchenko, L. T., and Cooper, T. A. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. Science *280*, 737-741.

Phillips, J. E., and Corces, V. G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Planells-Cases, R., and Jentsch, T. J. (2009). Chloride channelopathies. Biochim Biophys Acta *1792*, 173-189.

Ranum, L. P., and Cooper, T. A. (2006). RNA-Mediated Neuromuscular Disorders. Annu Rev Neurosci.

Ranum, L. P., and Day, J. W. (2002). Dominantly inherited, non-coding microsatellite expansion disorders. Curr Opin Genet Dev *12*, 266-271.

Rappsilber, J., Ryder, U., Lamond, A. I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res *12*, 1231-1245.

Ray, S., Bowen, J. R., and Marks, H. G. (1984). Foot deformity in myotonic dystrophy. Foot Ankle *5*, 125-130.

Reddy, S., Mistry, D. J., Wang, Q. C., Geddis, L. M., Kutchai, H. C., Moorman, J. R., and Mounsey, J. P. (2002). Effects of age and gene dose on skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. Muscle Nerve *25*, 850-857.

Reddy, S., Smith, D. B., Rich, M. M., Leferovich, J. M., Reilly, P., Davis, B. M., Tran, K., Rayburn, H., Bronson, R., Cros, D.*, et al.* (1996). Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. Nat Genet *13*, 325-335.

Regev, R., de Vries, L. S., Heckmatt, J. Z., and Dubowitz, V. (1987). Cerebral ventricular dilation in congenital myotonic dystrophy. J Pediatr *111*, 372-376.

Ren, S. Y., Angrand, P. O., and Rijli, F. M. (2002). Targeted insertion results in a rhombomere 2-specific Hoxa2 knockdown and ectopic activation of Hoxa1 expression. Dev Dyn *225*, 305-315.

Richards, R. I., and Sutherland, G. R. (1994). Simple repeat DNA is not replicated simply. Nat Genet *6*, 114-116.

Ricker, K., Koch, M. C., Lehmann-Horn, F., Pongratz, D., Speich, N., Reiners, K., Schneider, C., and Moxley, R. T., 3rd (1995). Proximal myotonic myopathy. Clinical features of a multisystem disorder similar to myotonic dystrophy. Arch Neurol *52*, 25-31.

Roulier, E. M., Fyrberg, C., and Fyrberg, E. (1992). Perturbations of Drosophila alpha-actinin cause muscle paralysis, weakness, and atrophy but do not confer obvious nonmuscle phenotypes. J Cell Biol *116*, 911-922.

Rutherford, M. A., Heckmatt, J. Z., and Dubowitz, V. (1989). Congenital myotonic dystrophy: respiratory function at birth determines survival. Arch Dis Child *64*, 191-195.

Sabouri, L. A., Mahadevan, M. S., Narang, M., Lee, D. S., Surh, L. C., and Korneluk, R. G. (1993). Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. Nat Genet *4*, 233-238.

Sadler, A. J., and Williams, B. R. (2007). Structure and function of the protein kinase R. Curr Top Microbiol Immunol *316*, 253-292.

Sahgal, V., Bernes, S., Sahgal, S., Lischwey, C., and Subramani, V. (1983a). Skeletal muscle in preterm infants with congenital myotonic dystrophy. Morphologic and histochemical study. J Neurol Sci *59*, 47-55.

Sahgal, V., Sahgal, S., Bernes, S., and Subramani, V. (1983b). Ultrastructure of muscle spindle in congenital myotonic dystrophy. A study of preterm infant muscle spindles. Acta Neuropathol *61*, 207-213.

Sarkar, P. S., Appukuttan, B., Han, J., Ito, Y., Ai, C., Tsai, W., Chai, Y., Stout, J. T., and Reddy, S. (2000). Heterozygous loss of Six5 in mice is sufficient to cause ocular cataracts. Nat Genet *25*, 110-114.

Sarkar, P. S., Paul, S., Han, J., and Reddy, S. (2004). Six5 is required for spermatogenic cell survival and spermiogenesis. Hum Mol Genet *13*, 1421-1431.

Sarnat, H. B., and Silbert, S. W. (1976). Maturational arrest of fetal muscle in neonatal myotonic dystrophy. A pathologic study of four cases. Arch Neurol *33*, 466-474.

Savkur, R. S., Philips, A. V., and Cooper, T. A. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. Nat Genet *29*, 40-47.

Savkur, R. S., Philips, A. V., Cooper, T. A., Dalton, J. C., Moseley, M. L., Ranum, L. P., and Day, J. W. (2004). Insulin receptor splicing alteration in myotonic dystrophy type 2. Am J Hum Genet *74*, 1309-1313.

Schild, R. L., Plath, H., Hofstaetter, C., Brenner, R., Mann, E., Mundegar, R. R., Steinbach, P., and Hansmann, M. (1998). Polyhydramnios: an association with congenital myotonic dystrophy. J Obstet Gynaecol *18*, 484-485.

Schmucker, D., Clemens, J. C., Shu, H., Worby, C. A., Xiao, J., Muda, M., Dixon, J. E., and Zipursky, S. L. (2000). Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell *101*, 671-684.

Schwenk, F., Baron, U., and Rajewsky, K. (1995). A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. Nucleic Acids Res *23*, 5080-5081.

Seino, S., and Bell, G. I. (1989). Alternative splicing of human insulin receptor messenger RNA. Biochem Biophys Res Commun *159*, 312-316.

Seznec, H., Lia-Baldini, A. S., Duros, C., Fouquet, C., Lacroix, C., Hofmann-Radvanyi, H., Junien, C., and Gourdon, G. (2000). Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. Hum Mol Genet *9*, 1185-1194.

Shao, J., and Diamond, M. I. (2007). Polyglutamine diseases: emerging concepts in pathogenesis and therapy. Hum Mol Genet *16 Spec No. 2*, R115-123.

Shaper, J. H., Barker, R., and Hill, R. L. (1973). Purification of wheat germ agglutinin by affinity chromatography. Anal Biochem *53*, 564-570.

Shepard, P. J., and Hertel, K. J. (2009). The SR protein family. Genome Biol *10*, 242.

Siegel, I. M., Eisenberg, B. R., and Glantz, R. H. (1984). Contributory etiologic factor for talipes equinovarus in congenital myotonic dystrophy: comparative biopsy study of intrinsic foot musculature and vastus lateralis in two cases. J Pediatr Orthop *4*, 327-330.

Squillace, R. M., Chenault, D. M., and Wang, E. H. (2002). Inhibition of muscle differentiation by the novel muscleblind-related protein CHCR. Dev Biol *250*, 218-230.

Stamm, S., Casper, D., Hanson, V., and Helfman, D. M. (1999). Regulation of the neuron-specific exon of clathrin light chain B. Brain Res Mol Brain Res *64*, 108-118.

Stuart, C. A., Armstrong, R. M., Provow, S. A., and Plishker, G. A. (1983). Insulin resistance in patients with myotonic dystrophy. Neurology *33*, 679-685.

Tanabe, Y., and Nonaka, I. (1987). Congenital myotonic dystrophy. Changes in muscle pathology with ageing. J Neurol Sci *77*, 59-68.

Teplova, M., and Patel, D. J. (2008). Structural insights into RNA recognition by the alternative-splicing regulator muscleblind-like MBNL1. Nat Struct Mol Biol *15*, 1343-1351.

Thornton, C. A., Griggs, R. C., and Moxley, R. T., 3rd (1994). Myotonic dystrophy with no trinucleotide repeat expansion. Ann Neurol *35*, 269-272.

Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, C. A. (2000). Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. RNA *6*, 79-87.

Timchenko, N. A., Cai, Z. J., Welm, A. L., Reddy, S., Ashizawa, T., and Timchenko, L. T. (2001). RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. J Biol Chem *276*, 7820-7826.

Tominaga, K., Hayashi, Y. K., Goto, K., Minami, N., Noguchi, S., Nonaka, I., Miki, T., and Nishino, I. (2010). Congenital myotonic dystrophy can show congenital fiber type disproportion pathology. Acta Neuropathol.

Townsend, P. J., Farza, H., MacGeoch, C., Spurr, N. K., Wade, R., Gahlmann, R., Yacoub, M. H., and Barton, P. J. (1994). Human cardiac troponin T: identification of fetal isoforms and assignment of the TNNT2 locus to chromosome 1q. Genomics *21*, 311-316.

Tsilfidis, C., MacKenzie, A. E., Mettler, G., Barcelo, J., and Korneluk, R. G. (1992). Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. Nat Genet *1*, 192-195.

Utomo, A. R., Nikitin, A. Y., and Lee, W. H. (1999). Temporal, spatial, and cell type-specific control of Cre-mediated DNA recombination in transgenic mice. Nat Biotechnol *17*, 1091-1096.

Vialettes, B., Pouget, J., Viard, R., Moulin, J. P., Serratrice, G., and Vague, P. (1986). Mechanism and significance of insulin resistance in myotonic dystrophy. Horm Metab Res *18*, 395-399.

Vicente, M., Monferrer, L., Poulos, M. G., Houseley, J., Monckton, D. G., O'Dell K, M., Swanson, M. S., and Artero, R. D. (2007). Muscleblind isoforms are functionally distinct and regulate alpha-actinin splicing. Differentiation *75*, 427-440.

Vogt, B., Carrascosa, J. M., Ermel, B., Ullrich, A., and Haring, H. U. (1991). The two isotypes of the human insulin receptor (HIR-A and HIR-B) follow different internalization kinetics. Biochem Biophys Res Commun *177*, 1013-1018.

Wahl, M. C., Will, C. L., and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701-718.

Wakimoto, H., Maguire, C. T., Sherwood, M. C., Vargas, M. M., Sarkar, P. S., Han, J., Reddy, S., and Berul, C. I. (2002). Characterization of cardiac conduction system abnormalities in mice with targeted disruption of Six5 gene. J Interv Card Electrophysiol *7*, 127-135.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Ward, A. J., and Cooper, T. A. (2010). The pathobiology of splicing. J Pathol *220*, 152-163.

Warf, M. B., and Berglund, J. A. (2007). MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. RNA *13*, 2238-2251.

Wheeler, T. M., Lueck, J. D., Swanson, M. S., Dirksen, R. T., and Thornton, C. A. (2007). Correction of ClC-1 splicing eliminates chloride channelopathy and myotonia in mouse models of myotonic dystrophy. J Clin Invest *117*, 3952-3957.

Wieacker, P., Wilhelm, C., Furste, H., and Schillinger, H. (1988). [Polyhydramnios in congenital myotonic dystrophy]. Z Geburtshilfe Perinatol *192*, 36-37.

Wong, L. J., Ashizawa, T., Monckton, D. G., Caskey, C. T., and Richards, C. S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. Am J Hum Genet *56*, 114-122.

Xu, R., Teng, J., and Cooper, T. A. (1993). The cardiac troponin T alternative exon contains a novel purine-rich positive splicing element. Mol Cell Biol *13*, 3660-3674.

Yeo, G. W., Coufal, N. G., Liang, T. Y., Peng, G. E., Fu, X. D., and Gage, F. H. (2009). An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. Nat Struct Mol Biol *16*, 130-137.

Yuan, Y., Compton, S. A., Sobczak, K., Stenberg, M. G., Thornton, C. A., Griffith, J. D., and Swanson, M. S. (2007). Muscleblind-like 1 interacts with RNA hairpins in splicing target and pathogenic RNAs. Nucleic Acids Res *35*, 5474-5486.

Zhou, Z., Licklider, L. J., Gygi, S. P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature *419*, 182-185.

BIOGRAPHICAL SKETCH

Michael Poulos was born in Midland, Michigan, in 1979. He is the oldest of two sons born to David and Eileen Poulos. Michael attended Grand Valley State University from 1997-2002 where he studied biology and chemistry and earned a Bachelor of Science in biology. After graduation, Michael moved to Gainesville, Florida and joined the interdisciplinary program in biomedical sciences at the University of Florida College of Medicine. Michael did his graduate work in Dr. Maury Swanson's laboratory of the Molecular Genetics and Microbiology Department and completed his Ph.D. dissertation in May 2010. Michael plans to move to New York in the spring of 2010 to begin his postdoctoral studies in Dr. Stewart Shuman's laboratory at the Sloan-Kettering Institute .