Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy

James D. Thomas,[1] Łukasz J. Sznajder,[1] Olgert Bardhi,[1] Faaiq N. Aslam,[1] Zacharias P. Anastasiadis,[1] Marina M. Scotti,[1] Ichizo Nishino,[2] Masayuki Nakamori,[3] Eric T. Wang,[1] and Maurice S. Swanson[1]

[1]Department of Molecular Genetics and Microbiology, Center for NeuroGenetics and the Genetics Institute, College of Medicine, University of Florida, Gainesville, Florida 32610, USA; [2]Department of Neuromuscular Research, National Center of Neurology and Psychiatry, Osaka University Graduate School of Medicine, Osaka 565-0871, Japan; [3]Department of Neurology, Osaka University Graduate School of Medicine, Osaka 565-0871, Japan

Myotonic dystrophy type 1 (DM1) is a CTG microsatellite expansion (CTG^{exp}) disorder caused by expression of CUG^{exp} RNAs. These mutant RNAs alter the activities of RNA processing factors, including MBNL proteins, leading to re-expression of fetal isoforms in adult tissues and DM1 pathology. While this pathogenesis model accounts for adult-onset disease, the molecular basis of congenital DM (CDM) is unknown. Here, we test the hypothesis that disruption of developmentally regulated RNA alternative processing pathways contributes to CDM disease. We identify prominent alternative splicing and polyadenylation abnormalities in infant CDM muscle, and, although most are also misregulated in adult-onset DM1, dysregulation is significantly more severe in CDM. Furthermore, analysis of alternative splicing during human myogenesis reveals that CDM-relevant exons undergo prenatal RNA isoform transitions and are predicted to be disrupted by CUG^{exp}-associated mechanisms in utero. To test this possibility and the contribution of MBNLs to CDM pathogenesis, we generated mouse *Mbnl* double (*Mbnl1; Mbnl2*) and triple (*Mbnl1; Mbnl2; Mbnl3*) muscle-specific knockout models that recapitulate the congenital myopathy, gene expression, and spliceopathy defects characteristic of CDM. This study demonstrates that RNA misprocessing is a major pathogenic factor in CDM and provides novel mouse models to further examine roles for cotranscriptional/post-transcriptional gene regulation during development.

[*Keywords*: congenital myotonic dystrophy; MBNL; microsatellite; myoblast; myogenesis; RNA processing]

Supplemental material is available for this article.

Received April 18, 2017; revised version accepted May 26, 2017.

During tissue development, coordinated gene expression programs provide the gene products necessary to meet the functional demands of distinct cell populations. For example, skeletal muscle formation depends on a family of myogenic transcription factors that control gene expression as proliferative myoblasts differentiate into postmitotic, multinucleated myofibers (Bentzinger et al. 2012). While muscle transcriptional regulation has been the subject of numerous studies, cotranscriptional/post-transcriptional RNA processing during myogenic differentiation is less characterized, particularly in the context of human development. During myogenesis, splicing patterns and the subcellular localization of several RNA-binding proteins (RBPs) change (Bland et al. 2010), and disruption of RBP availability or altered splicing impairs myogenic differentiation (Poulos et al. 2013; Singh et al. 2014). Modulation of various RBP levels in vivo, including RBFOX, MBNL, and CELF, alters distinct stages of muscle development and/or maintenance. For example, aberrant expression of RBFOX2 inhibits myoblast fusion, while loss of RBFOX1 is implicated in later stages of muscle development, including sarcomerogenesis and functional muscle maintenance (Singh et al. 2014; Pedrotti et al. 2015). In fact, RBPs often coordinate their activities to control alternative splicing (AS) patterns in health and disease (Brinegar and Cooper 2016). To date, very little is known about the roles of RBP interactions and alternative RNA processing in neonatal myopathies such as congenital myotonic dystrophy (CDM), where failure of normal muscle development is a characteristic feature.

CDM is a prenatal-onset, multisystemic disease caused by a CTG microsatellite expansion (CTG^{exp}) in the *DMPK* gene (Ho et al. 2015). CTG copy numbers as few as 750, but typically >1000, are associated with CDM (Tsilfidis et al. 1992; Joseph et al. 1997), and mutant *DMPK* transcripts are expressed in a variety of affected tissues

Corresponding author: mswanson@ufl.edu
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.300590.117.

throughout embryogenesis (Wong and Ashizawa 1997). Prenatal symptoms of CDM include polyhydramnios and reduced fetal movement, while newborn infants display muscle immaturity, hypotonia, and life-threatening respiratory insufficiency (Ho et al. 2015; Johnson et al. 2016). Adult-onset DM (DM1) patients are asymptomatic at birth and experience symptoms in later life likely due to postnatal somatic CTG repeat expansion (Wong et al. 1995). Thus, in contrast to DM1, the prenatal CDM genome harbors pathogenic CTG^{exp} alleles, potentially subjecting these patients to CTG^{exp}-associated disease mechanisms during tissue development.

A major pathomechanism underlying adult-onset DM1 is the production of CUG^{exp} RNAs that modulate the activity of RNA processing factors, including MBNL proteins. MBNL functional inactivation in DM1 tissues results in the re-emergence of developmentally immature AS and alternative polyadenylation (APA) patterns in adult tissues as well as alterations in RNA localization and turnover (Masuda et al. 2012; Wang et al. 2012; Batra et al. 2014). Since disruption of RNA processing is a prominent feature of DM1, in this study, we tested the possibility that similar molecular mechanisms contribute to disease manifestations in CDM.

Using transcriptome profiling of CDM skeletal muscle, we provide evidence for dramatic AS and APA abnormalities in CDM infant tissues. We compared these events with a large cohort of DM1 patient muscle samples and identified CDM as a DM1-like spliceopathy that differs mostly in the degree and timing of spliceopathy onset. This model is supported by our analysis of RNA processing changes during normal human and mouse myogenesis, where we identified CDM-relevant exons undergoing RNA isoform transitions in utero. Furthermore, high *DMPK* expression was identified during prenatal myogenesis, suggesting that CUG^{exp}-associated disease mechanisms are particularly burdensome during this developmental window. We tested this hypothesis by studying one potential contributor to CDM disease; namely, MBNL loss of function. Using mouse *Mbnl* knockout models, we demonstrated that disruption of MBNL activity during prenatal myogenesis results in congenital myopathy and spliceopathy. Furthermore, a comparison of mouse *Mbnl* knockout and human CDM muscle transcriptomes reveals a cohort of conserved missplicing events and disrupted biological pathways that may be particularly important in early disease manifestations. Together, these studies provide insights into the CDM pathomechanism and, given the technical limitations of studying embryonic and fetal human tissues, establish important mouse models to explore the role of RNA processing in tissue development.

Results

Severe RNA misprocessing in CDM

To determine whether the alternative processing of RNA transcripts was affected in CDM, we performed RNA sequencing (RNA-seq) on biceps brachii muscle from disease control (spinal muscular atrophy [SMA] type 1) and CDM infants at 3–15 mo of age (Supplemental Table S1). Using MISO (Katz et al. 2010) to calculate percent spliced in (Ψ) for alternative splicing events, we identified thousands of cassette exon (CE), mutually exclusive exon (MXE), alternative 5′ splice site (A5SS), alternative 3′ splice site (A3SS), and retained intron (RI) splicing events (Fig. 1A; Supplemental Fig. S1A; Supplemental Table S2). While most splicing events showed no or subtle deviations relative to controls, hundreds displayed dramatic differential Ψ ($\Delta\Psi$; e.g., 193 MXEs) and were assigned as disease-relevant outliers. For example, *MEF2D* α1 and α2 MXEs were misspliced (Fig. 1B), an event linked to impaired myogenesis (Singh et al. 2014; Runfola et al. 2015), and misregulation of this and additional targets was confirmed by RT–PCR (Fig. 1C).

Recently, transcriptome-wide splicing analysis was performed on a large cohort of DM1 skeletal muscle samples, and a set of common missplicing events was described (Wagner et al. 2016). In our CDM data set, we found consistent missplicing of these same events and others reported previously in CDM patient-derived cells (Supplemental Fig. S1B; Fugier et al. 2011). Next, we compared CDM muscle with healthy controls, DM1 proto-mutation (proto) carriers (<100 CTG repeats as assessed from peripheral blood), and adult-onset DM1 patients (Supplemental Table S1; Wagner et al. 2016) to test for evidence of CDM-specific missplicing. In general, missplicing events were highly conserved between DM1 and CDM but were consistently more severe in the latter. For example, *MEF2D* MXE missplicing increased in CDM compared with any DM1 sample (Fig. 1D). The $\Delta\Psi$ range in this patient cohort is expected, as myopathic severity (Nakamori et al. 2013) and MBNL loss of function (Wagner et al. 2016) are also highly variable. Indeed, rank ordering of the adult muscle data sets according to inferred MBNL concentration ($[MBNL]_{inferred}$), an approximation of MBNL levels in patient muscle based on a set of 46 validated AS events (Wagner et al. 2016), revealed a correlation between functional MBNL level and abnormal *MEF2D* α_2 exon exclusion (Fig. 1E). Stronger trends were observed for other events (Supplemental Fig. S1C), and CDM muscle was among the most severely affected patient samples. Using the 695 CEs identified in Figure 1A, the mean $|\Delta\Psi|$ for CDM and adult DM1 was calculated, and greater total spliceopathy was found in CDM than in all DM1 samples compared with age-matched controls (Fig. 1F).

Since APA misregulation is another RNA processing step implicated in DM1 disease, we also performed PolyA-seq (polyadenylation [polyA] sequencing) and identified thousands of high-confidence 3′ ends, with 363 showing significant differential polyA site (PAS) selection (ΔPAS) in CDM compared with controls (Fig. 1G; Supplemental Table S2). For example, use of distal cleavage and PASs (dPASs) in *PSAP* transcripts were decreased in CDM compared with controls (Fig. 1H). Cumulatively, these results suggested that similar RNA misprocessing events observed in DM1 contribute to CDM pathogenesis, although CDM misprocessing was significantly more

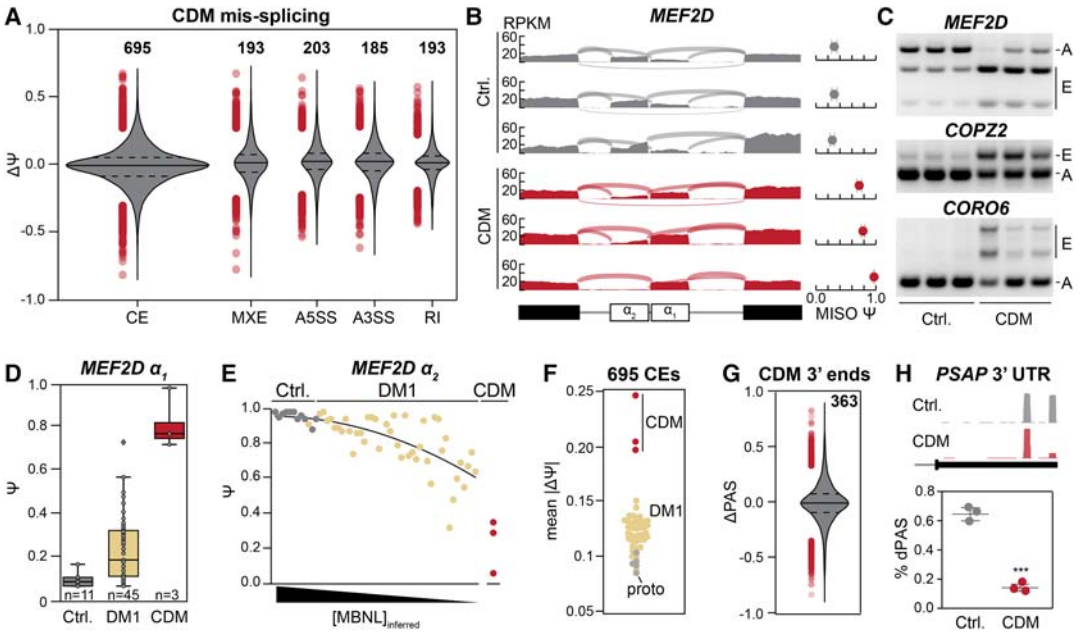

Figure 1. Prominent RNA misprocessing in CDM skeletal muscle. (*A*) Violin plots depicting the distribution of splicing events based on ΔΨ and types of AS, including CEs, MXEs, A5SSs, A3SSs, and RIs. Events with particularly dramatic ΔΨ are indicated (red circles) and quantified (red number *above* each plot). (*B*) RNA-seq read coverage across *MEF2D* MXEs α1 and α2 in disease control (SMA type 1; gray) and CDM (red) muscles. Splice junction reads are indicated with lighter gray and red lines spanning introns. Quantification of α1 Ψ is indicated at the *right*, with 95% confidence intervals shown (gray lines). (*C*) RT–PCR validation of *MEF2D*, *COPZ2*, and *CORO6* missplicing in CDM muscle. PCR products corresponding to embryonic (E) and adult (A) isoforms are indicated. (*D*) Box plots of *MEF2D* exon α1 Ψ in adult control, adult DM1, and CDM muscle. (*E*) *MEF2D* exon α2 Ψ in adult control, adult DM1, and CDM muscle based on samples ranked according to inferred MBNL concentration ($[MBNL]_{inferred}$). $[MBNL]_{inferred}$ was not calculated for CDM samples, so these points are plotted outside the fitted curve (black line). (*F*) Total spliceopathy relative to age-matched controls (mean |ΔΨ|) based on 695 CEs for proto-mutation (proto; gray circles), DM1 (yellow), and CDM (red) patients. (*G*) Violin plot depicting the distribution of misregulated APA events in CDM based on differential inclusion levels (differential polyA site [PAS] selection [ΔPAS]; *Y*-axis). Severely misregulated events are indicated as in *A*. (*H*) PolyA-seq (polyadenylation [polyA] sequencing) read coverage across the *PSAP* 3′ untranslated region in representative control and CDM muscles, with quantification based on all three samples *below*. (***) False discovery rate <0.001.

pronounced. To specifically identify events that would be susceptible to CUG^{exp} RNA toxicity in utero, we next explored AS transitions associated with prenatal muscle development.

RNA processing transitions in human muscle development

We modeled in utero transitions in RNA processing associated with human muscle development using RNA-seq data sets obtained from primary human myoblast differentiation in vitro (Trapnell et al. 2014), fetal muscle (ENCODE), and the control muscle data described above (Fig. 2A; Supplemental Table S1). All data sets were processed using the same analysis and quality control pipeline (Supplemental Fig. S2A,B) to minimize potential sample-specific biases. This strategy incorporated a permutation-based method to identify exon inclusion patterns that changed monotonically (consistently increasing or decreasing) over time and reported a *Z*-score to describe the significance of these changes (Wang et al. 2015).

Using this in silico model of myogenesis, we identified expected gene expression (Fig. 2B) and splicing (Fig. 2C) changes in differentiation-associated transcripts. Furthermore, thousands of additional CEs (Fig. 2D) and other forms of AS (Supplemental Fig. S2C; Supplemental Table S3) showed evidence of prenatal splicing transitions. Exon utilization changes were often large, as exemplified by *ITGA7* exon 25, which undergoes an ~70% shift in prenatal inclusion (Supplemental Fig. S2D). Overall, these data are consistent with the hypothesis that alternative RNA processing is an important component of normal skeletal muscle development, as has been described previously for heart morphogenesis (Kalsotra et al. 2008; Kalsotra and Cooper 2011). Importantly, the use of this in silico myogenesis data set confirmed that the infant disease control muscle data fit along the expected developmental continuum for most splicing events. Next, we tested which CDM-relevant exons are associated with muscle development in utero.

BIN1 exon 11, a CDM missplicing event, was among the myogenesis-associated exons with the greatest ΔΨ in utero (≥70%) (Fig. 2E). This in utero splicing shift was recapitulated in mice, where orthologous *Bin1* exon 11 was excluded in primary myoblasts but reached nearly 100% inclusion in both myotubes in vitro and dissected postnatal day 0 (P0) muscle (Fig. 2F). Importantly, BIN1 is essential for T-tubule maturation, and antisense

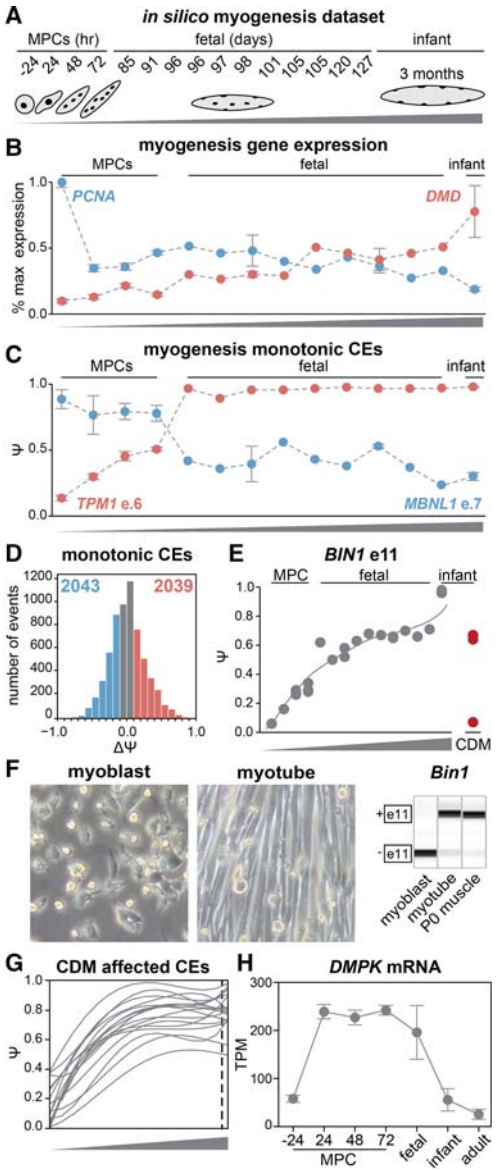

Figure 2. Failure of prenatal AS transitions in CDM. (*A*) Schematic of cellular transitions during normal human myogenesis with the associated RNA-seq data used in this study. Hours (hr) of differentiation in vitro for muscle precursor cells (MPCs), days of fetal muscle, and age (months) of infant muscle are indicated. The gray triangle indicates increasing muscle maturity and is used as the *X*-axis label throughout the figure. (*B*) Percent maximal gene expression of *PCNA* (proliferation marker) and *DMD* (differentiation marker) during muscle maturation. (*C*) *TPM1* exon 6 and *MBNL1* exon 7 (36 nucleotides [nt]) Ψ during muscle maturation. (*D*) Histogram depicting the number of monotonic CE events with consistently increasing (red) or decreasing (blue) Ψ from −24 h of myogenic differentiation in vitro to day 127 fetal muscle. (*E*) Quantification of *BIN1* exon 11 Ψ during muscle maturation. Ψ values for each CDM patient are indicated (red circle). (*F*) Phase contrast images of primary mouse myoblasts and differentiated myotubes (*left*) and RT–PCR validation of mouse prenatal *Bin1* exon 11 splicing transition (*right*). (*G*) Identification of CDM-relevant exons undergoing early isoform transitions during muscle development. (*H*) Quantification of *DMPK* RNA levels during muscle maturation.

oligonucleotide-mediated reduction of *Bin1* exon 11 inclusion is sufficient to cause myopathy in vivo (Fugier et al. 2011). Many CDM-relevant exons showed evidence of undergoing prenatal splicing transitions that largely completed prior to birth, and, in many cases, CDM Ψ resembled myoblast levels of inclusion (Supplemental Fig. S2E,F). These events displayed distinct splicing dynamics prior to birth, indicative of a temporal hierarchy in which splicing transitions are preferentially affected in CDM muscle during transcriptome maturation (Fig. 2G; Supplemental Fig. S2G). Using the same human data sets, *DMPK* RNA levels were found to be particularly high in fetal muscle as well as muscle precursor cells during myogenic differentiation in vitro (Fig. 2H). In agreement with these expression dynamics, analysis of publicly available ChIP-seq (chromatin immunoprecipitation [ChIP] combined with high-throughput sequencing) data revealed increased MYOG and MYOD occupancy over the *Dmpk* intron 1 enhancer at the initiation of C2C12 differentiation in vitro but not in fully differentiated myotubes (Supplemental Fig. S2H). Together, these data underscore the toxic burden of CUG^{exp} RNA during myogenesis.

While these CDM skeletal muscle transcriptome analyses provide important insights into disease pathogenesis, the use of human tissues for mechanistic studies of embryonic and neonatal tissue development is limited. Furthermore, the control and CDM tissues were collected from patients between 3 and 15 mo of age, so the presence of degeneration–regeneration-associated secondary effects (Orengo et al. 2011; Bachinski et al. 2014) confounds identification of disease-initiating RNA misprocessing events. To overcome these limitations, we next tested the utility of *Mbnl* knockout mice as models for CDM-associated cellular and molecular phenotypes.

Congenital myopathy and spliceopathy in Mbnl1; Mbnl2 double-knockout mice

Our analysis of the CDM muscle and human myogenesis transcriptomes motivated us to test whether the absence of MBNL activity during myogenesis would disrupt prenatal RNA processing and result in congenital phenotypes. In agreement, dramatic neonatal lethality was observed in *Mbnl1*; *Mbnl2* conditional double-knockout mice (Lee et al. 2013), accompanied by a reduction in total body weight at birth (Fig. 3A,B). Moribund P0 double-knockout mice were cyanotic (Fig. 3C) with labored breathing (Supplemental Movie S1). To test for congenital myopathy, we analyzed soleus and tibialis anterior muscles from hindlimb cross-sections (Supplemental Fig. S3A) and observed misshapen myofibers (Supplemental Fig. S3B) with an overall reduction in cross-sectional area (Supplemental Fig. S3C). Histopathological changes were also noted in P0 double-knockout quadriceps (Fig. 3D) with associated missplicing of a CDM-relevant exon, *Bin1* exon 11 (Fig. 3E). These data demonstrated that MBNL proteins regulate prenatal RNA processing essential for normal myogenesis and neonatal viability.

To characterize the congenital spliceopathy in double-knockout muscle and link a cohort of events with

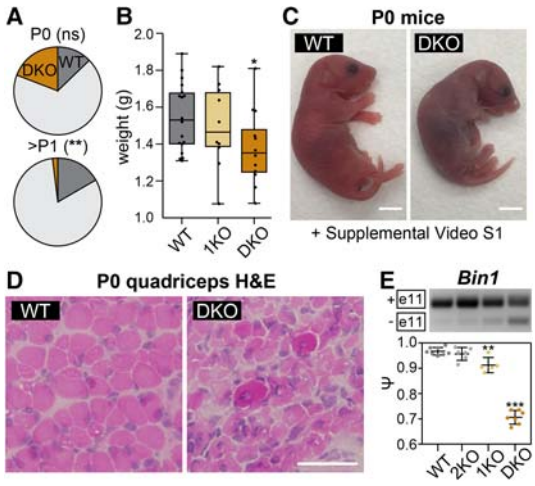

Figure 3. Congenital myopathy in *Mbnl* double-knockout mice. (*A*) Pie chart depicting the proportion of wild-type and double-knockout animals born (P0; *top*; [ns] not significant) and surviving beyond 1 d (>P1; *bottom*; [**] $P < 0.01$, χ^2 test). The light-gray section is the proportion of all other genotypes generated. (*B*) Box plot of P0 wild-type (WT), *Mbnl1*$^{-/-}$;*Mbnl2*$^{C/C}$;*Mbnl3*$^{C/Y}$;*Myog-Cre*$^{-/-}$ (1KO), and double-knockout (DKO) mouse total body weight. (*) $P < 0.05$, Student's *t*-test. (*C*) P0 wild-type and double-knockout mice. Moribund double-knockout mice were cyanotic with labored breathing (see Supplemental Movie S1). (*D*) Hematoxylin and eosin (H&E)-stained sections of wild-type and double-knockout quadriceps. (*E*) Representative RT–PCR for *Bin1* exon 11 in P0 quadriceps muscle (*top*) and quantification based on replicates (*bottom*). (**) $P < 0.01$; (***) $P < 0.001$, ANOVA.

neonatal phenotypes, we performed RNA-seq on dissected P0 quadriceps. We identified hundreds of misregulated CE events (Fig. 4A) such as the abnormal inclusion of *Bin1* exon 7 (Fig. 4B) as well as other AS modes in double-knockout muscle (Supplemental Table S4). Misregulated CEs were validated using RT–PCR (Fig. 4C; Supplemental Fig. S4A), and, while some events (e.g., *Atp2a1* exon 22) were misspliced in *Mbnl* single knockouts, total spliceopathy was significantly greater in double-knockout muscle (Fig. 4D), which indicated that MBNL paralogs provide functional compensation during myogenesis as reported in C2C12 myoblasts and adult muscle (Wang et al. 2012; Lee et al. 2013). Using mouse embryonic forelimb (ENCODE) as well as P0 and adult muscle RNA-seq data sets, we generated an in silico model of mouse myogenesis similar to that used for humans (Supplemental Table S1). In agreement with our observations in human muscle development, prominent prenatal splicing transitions were identified (Supplemental Fig. S4B), such as *Sorbs1* exon 15, which is severely misspliced in double-knockout muscle (Supplemental Fig. S4C).

Next, we tested for overlap between mouse double-knockout and CDM patient muscle missplicing and identified >2000 orthologous genes with evidence of missplicing in both data sets (Supplemental Fig. S4D). To refine this analysis, we specifically focused on orthologous CEs (see the Materials and Methods). After filtering events ($|\Delta\Psi| \geq 0.10$, CI ≤ 0.30 required in both data sets), 144 orthologous exons (Supplemental Table S5) were identified as being misspliced in P0 double-knockout and CDM muscle, with many showing concordant misregulation (Fig. 4E). For example, *CACNA2D1* exon 19 showed $\Delta\Psi$ of 0.57 and 0.59 in CDM and double-knockout muscle, respectively (Fig. 4F). The developmental transition for this exon initiates between embryonic day 15.5 (E15.5) and P0 and is recapitulated during C2C12 differentiation in vitro, suggesting that inclusion of this exon is important for myogenesis (Supplemental Fig. S4E,F). Although double-knockout mice present with CDM phenotypes and spliceopathy, combined loss of MBNL1 and MBNL2 was insufficient to model the extensive gene expression changes occurring in CDM (Fig. 4G). While many variables could cause larger gene expression changes in human tissue versus mouse models, we hypothesized that concomitant loss of the third MBNL paralog (MBNL3) might be responsible for these additional features.

Precocious expression of muscle-specific transcripts in Mbnl3 knockout myoblasts

MBNL3 is expressed in C2C12 myoblasts but declines during in vitro differentiation and is not detectable in adult skeletal muscle (Poulos et al. 2013). To extend these prior observations, MBNL3 expression was measured throughout early myogenesis using single myofibers from adult mice (Pasut et al. 2013). MBNL3 was specifically expressed in activated PAX7$^+$ satellite cells associated with single myofibers grown in culture for 3 d but not in quiescent PAX7$^+$ satellite cells associated with immediately fixed fibers (Fig. 5A). In myogenin (MYOG)-positive primary myoblasts, MBNL3 was also detectable, primarily in the cytoplasm (Fig. 5B). These observations suggested that MBNL3 functions in myoblasts but is not required for the maintenance of satellite cell quiescence.

To study the roles of MBNL3, a *Mbnl3* conditional null allele (*Mbnl3*^{condWL}) was generated with loxP sites flanking *Mbnl3* exons 2 and 7c (Supplemental Fig. 5A). Targeted embryonic stem cells (Supplemental Fig. S5B) were used to generate mouse conditional lines that were subsequently crossed to a *CMV-Cre*-expressing line, and both Mbnl3 RNA (Supplemental Fig. S5C) and MBNL3 protein (Supplemental Fig. S5D) isoforms were undetectable in tissues from these *Mbnl3* knockout mice. Furthermore, RNA-seq analysis of *Mbnl3* knockout myoblasts confirmed the expected *Mbnl3* exon 2–7c deletion (Supplemental Fig. S5E). In agreement with our previous studies on *Mbnl3* isoform knockouts that express a short MBNL3 isoform (Poulos et al. 2013), the *Mbnl3* knockout primary myoblasts displayed defective in vitro differentiation phenotypes (Fig. 5C). In addition to this in vitro differentiation defect, *Mbnl3* knockout myoblasts showed aberrant cell spreading when grown at high density or in reduced serum medium (Supplemental Fig. S5F).

To test for transcriptome abnormalities associated with early myogenic defects in *Mbnl3* knockout myoblasts, RNA-seq was performed on subconfluent cultures grown

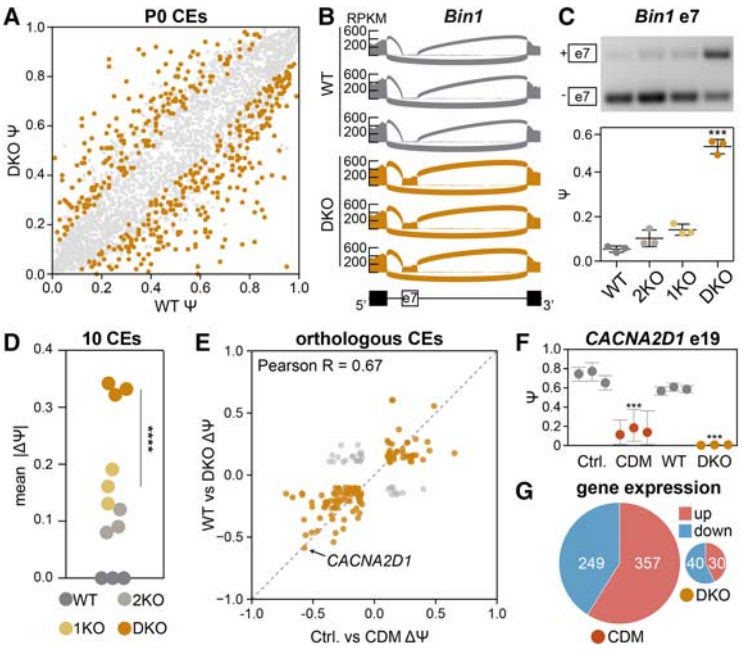

Figure 4. Congenital spliceopathy in *Mbnl* double-knockout mice. (*A*) Scatter plot comparing CE Ψ between P0 wild-type (WT) and double-knockout (DKO) muscle. Significant events are highlighted in orange. (*B*) RNA-seq read coverage across *Bin1* exon 7 in wild-type and double-knockout muscle. (*C*) Representative RT–PCR for *Bin1* exon 7 in P0 quadriceps muscle (*top*) and quantification based on replicates (*bottom*). (***) $P < 0.001$, ANOVA. (*D*) Total spliceopathy (mean $|\Delta\Psi|$) based on 10 CEs in *Mbnl1*$^{+/+}$;*Mbnl2*$^{C/C}$;*Mbnl3*$^{+/Y}$;*Myog-Cre*$^{+/-}$ (2KO), *Mbnl1*$^{-/-}$;*Mbnl2*$^{C/C}$;*Mbnl3*$^{C/Y}$;*Myog-Cre*$^{-/-}$ (1KO), and double-knockout muscle compared with wild-type controls. (****) $P < 0.0001$, ANOVA. (*E*) Scatter plot of ΔΨ values for double-knockout and CDM orthologous CEs. Events misspliced in double-knockout muscle that correlate with CDM are highlighted in orange. The theoretical perfect correlation (dashed gray line) is also shown. (*F*) Quantification of *CACNA2D1* exon 19 Ψ in CDM and double-knockout muscle. (***) Monotonicity *Z*-score >2.0. (*G*) Pie chart of total gene expression changes identified in CDM and double-knockout muscle.

for 0 or 6 h in differentiation medium. Prominent splicing misregulation was observed in *Mbnl3* knockout myoblasts (Supplemental Fig. S5G), such as abnormal inclusion of previously unannotated *Fermt2* tandem CEs (Supplemental Fig. S5H). *Fermt2* encodes Kindlin-2, a scaffolding protein essential for myocyte elongation (Dowling et al. 2008). Surprisingly, for many misspliced transcripts, we observed inclusion of adult pattern exons, such as the *Neb* MXEs (Fig. 5D). To confirm the normal developmental pattern of misspliced exons, we used the P0 wild-type and double-knockout muscle splicing data set, where compound loss of MBNL1 and MBNL2 results in reversion to developmentally immature AS patterns. Transcriptome-wide quantification revealed that 260 CEs displayed opposite patterns of regulation between *Mbnl3* knockout myoblasts compared with double-knockout muscle (Fig. 5E), and the emergence of adult pattern exons was validated in additional myoblast isolates (Fig. 5F). These results suggested that MBNL3 represses adult pattern RNA isoforms, as has been proposed previously (Lee et al. 2010). In agreement, *Mbnl3* knockout myoblasts also displayed early changes in mRNA abundance compared with wild-type controls (Supplemental Fig. S5I), with increases in many transcripts associated with terminal differentiation (Supplemental Fig. S5J) such as *dystrophin* (Supplemental Fig. S5K). Given these myogenic abnormalities in *Mbnl3* knockout myoblasts, we next determined whether novel phenotypes were present in mice that were deficient in all three MBNL paralogs.

MBNL triple knockouts recapitulate CDM-associated spliceopathy and gene expression changes

To generate *Mbnl1; Mbnl2; Mbnl3* muscle-specific (*Myog-cre*) triple-knockout mice, the *Mbnl3*^{condWL} allele was crossed onto the double-knockout background, and the presence of all three targeted alleles was confirmed by tail genomic DNA PCR (Fig. 6A). As for double-knockout mice, neonatal moribund triple knockouts displayed failure to thrive (Fig. 6B), muscle histopathology, and evidence of respiratory distress; for mice that survived into adulthood, dramatic respiratory distress, skeletal muscle wasting, and a reduction in muscle strength were observed (Fig. 6C; Supplemental Fig. S6A; Supplemental Movie S2).

To test for congenital transcriptome dysfunction in P0 triple-knockout mice, RNA-seq was performed at the same sequencing depth as double-knockout mice (Supplemental Fig. S6B). Triple knockouts recapitulated nearly all missplicing events identified in double knockouts (e.g., 93 out of 100 of the most reproducible misregulated CEs in double knockouts) but also presented with additional missplicing (Supplemental Fig. S6C; Supplemental Table S4). For example, 262 additional misregulated CEs (Fig. 6D) were identified in triple knockouts and included abnormal inclusion of two small (6-nucleotide [nt]) exons in *Ap2m1* and *Clip1* (Fig. 6E; Supplemental Fig. S6D). However, most triple-knockout-specific missplicing events were relatively subtle ($|\Delta\Psi| \leq 0.20$), and, although wild-type, double-knockout, and triple-knockout replicates distinctly clustered, total spliceopathy was only modestly increased in triple-knockout compared with double-knockout muscle (Fig. 6F). These data suggested that congenital spliceopathy was primarily due to compound loss of MBNL1 and MBNL2, but further loss of MBNL3 reveals hundreds of additional missplicing events.

In addition to increased spliceopathy, hundreds of differentially expressed genes were observed in wild-type versus triple-knockout muscle (Fig. 6G). While only 70 genes reached filtering thresholds (adjusted $P < 0.01$; $|\log_2$ fold change$| > 1$) in double-knockout mice, ~400 gene

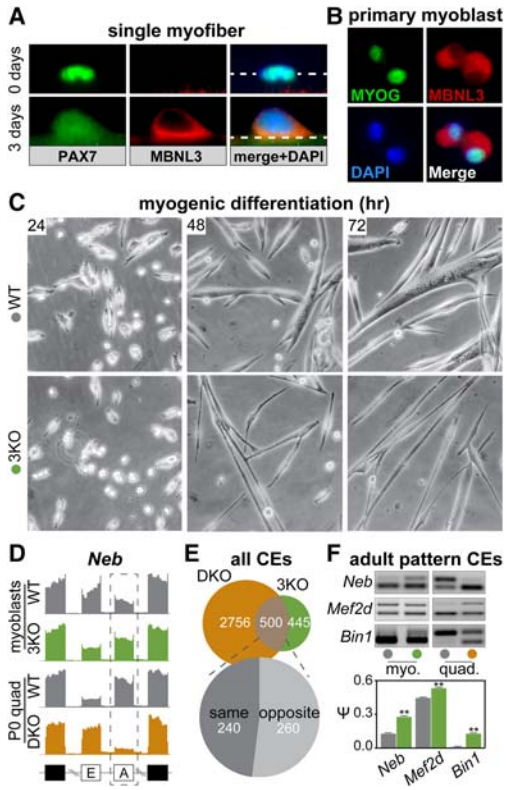

Figure 5. Emergence of adult pattern RNA isoforms in *Mbnl3* knockout myoblasts. (*A*) Immunofluorescence labeling of PAX7 (green) and MBNL3 (red) in freshly isolated (0 d) or cultured (3 d) single myofibers. (*B*) Immunofluorescent labeling of MYOG (green) and MBNL3 (red) in cultured primary myoblasts. (*C*) Representative phase contrast image of wild-type (WT) and *Mbnl3* knockout (3KO) myoblasts 24, 48, and 72 h following induction of differentiation in vitro. (*D*) RNA-seq read coverage across *Neb* MXEs depicting loss of the adult pattern exon in double-knockout (DKO) muscle and increased inclusion of the same exon in *Mbnl3* knockout compared with wild-type myoblasts. (*E*) Overlap of all CEs identified in the P0 double-knockout muscle and *Mbnl3* knockout myoblast RNA-seq data sets. Of the 500 overlapping events, 260 events show opposite directions in splicing shift, with double-knockouts and *Mbnl3* knockouts typically showing loss and gain of adult pattern exons, respectively. (*F*) RT–PCR for selected splicing events showing opposite splicing shifts from *E*. Data were generated from separate isolates of wild-type and *Mbnl3* knockout myoblasts as used for RNA-seq. Quantification of Ψ for each splicing event is shown *below*. (**) $P < 0.01$, Student's *t*-test.

expression changes were observed in triple knockouts (Fig. 6H). Strikingly, these changes were enriched in stress-related biological pathways similar to those in CDM muscle, such as hypoxia (Fig. 6H). Overall, a comparison of statistically significant gene ontology categories ($P < 0.01$; minimum enrichment ≥1.5) revealed 49 conserved pathways between triple-knockout and CDM muscle (Supplemental Fig. S6E,F; Supplemental Table S6). Because hypoxia was one of the major enriched categories and respiratory insufficiency is an important CDM-associated phenotype, we analyzed the diaphragm of P0 *Mbnl* knockout mice and found that acetylcholine receptor (AChR) organization and morphology were disrupted in P0 triple knockouts (Fig. 6I; Supplemental Fig. S6G,H). Cumulatively, our results indicate that loss of MBNL RNA processing activity in utero disrupts AS switches essential for myogenesis, resulting in impaired muscle development, respiratory system immaturity, and perinatal lethality, which are characteristic pathological features of CDM.

Discussion

RNA misprocessing in CDM

Key insights into the regulation of AS and APA during postnatal tissue development and regeneration have emerged from studies on pathomechanisms involved in DM (Kalsotra and Cooper 2011; Scotti and Swanson 2016). Here, we tested whether disruption of AS and APA developmental switching in utero are a hallmark of CDM disease. Our results confirm and extend prior work of *BIN1* pre-mRNA missplicing in CDM patient-derived cells (Fugier et al. 2011) by showing misregulation of hundreds of RNA alternative processing events in CDM biopsies. Using in silico models of human and mouse muscle development as well as MBNL loss-of-function models, we show that many CDM-relevant splicing events undergo prenatal transitions and are susceptible to CUG^{exp}-associated pathomechanisms during embryogenesis. Overall, this study demonstrates that RNA processing transitions are a prominent feature of the developing muscle transcriptome, and transcriptome-wide characterization links a cohort of RNA misprocessing outcomes with congenital myopathy.

These findings highlight a critical distinction between the congenital and adult-onset forms of DM: the developmental timing of CUG^{exp} toxicity. Most CDM cases are associated with maternal transmission of relatively large CTG^{exp} alleles (Myring et al. 1992; Tsilfidis et al. 1992), and pathogenic CUG^{exp} RNAs are expressed in a variety of prenatal/neonatal tissues (Wong and Ashizawa 1997). To date, splicing transitions for DM-relevant transcripts have been described largely as postnatal phenomena (Lin et al. 2006; Brinegar et al. 2017), and, while disruption of these events affects muscle function (Mankodi et al. 2002) and maintenance (Rau et al. 2015), misregulation of postnatal exons is not expected to prevent the morphogenesis of newborn muscle. Here, we provide evidence that major splicing transitions for exons affected in CDM occur in utero, are conserved during mouse prenatal myogenesis, and are disrupted in the absence of MBNL proteins along with coincident congenital myopathy in compound knockout mice. Furthermore, many of these splicing transitions are recapitulated during human and mouse myotube formation in vitro, suggesting that these events are associated with specific phases of cellular differentiation (e.g., myoblast fusion) rather than maturation of existing myotubes. In addition to providing insights into the CDM pathomechanism, this work provides a comprehensive survey of AS and APA patterns in CDM

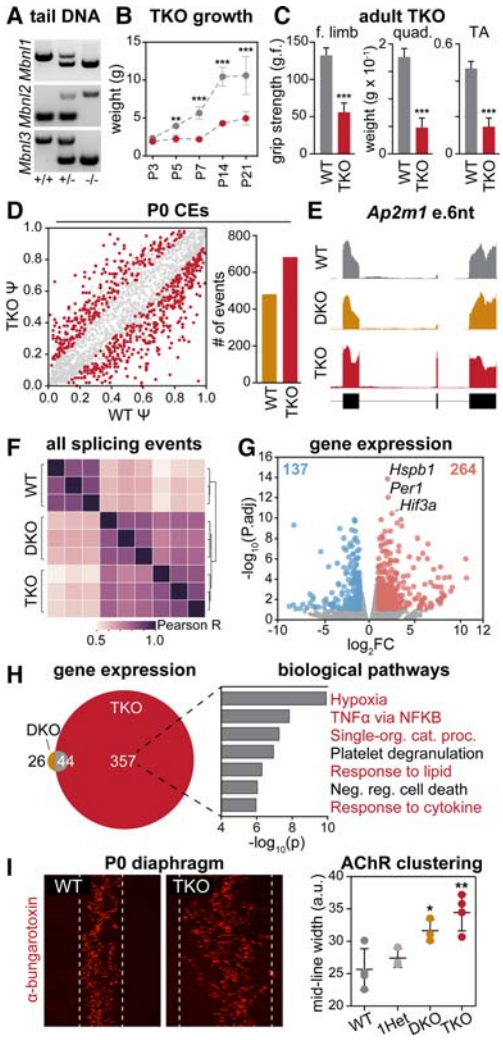

Figure 6. *Mbnl* triple-knockout mice reveal additional congenital phenotypes. (*A*) PCR of tail-derived gDNA from mice harboring wild-type (+/+) or heterozygous (+/−) or homozygous (−/−) mutant *Mbnl* alleles. (*B*) Postnatal growth of wild type (WT) and triple knockouts (TKO). (**) $P < 0.01$; (***) $P < 0.001$, Student's *t*-test. (*C*) Quantification of forelimb grip strength (grams force [g. f.]) and quadriceps and TA weight in adult triple-knockout mice. (***) $P < 0.001$, Student's *t*-test. (*D*) Scatter plot of misspliced CEs in P0 triple-knockout skeletal muscle compared with wild-type controls (*left*) and total number of misspliced CEs in double knockouts and triple knockouts (*right*). (*E*) RNA-seq read coverage across a 6-nt exon in *Ap2m1* (e.6 nt) in wild type, double knockouts (DKO), and triple knockouts. (*F*) Unsupervised clustering based on the correlation of Ψ values for all muscle AS events identified in *Mbnl* knockout muscle. (*G*) Scatter plot depicting genes up-regulated (red) and down-regulated (blue) in triple-knockout muscle. Gray circles indicate genes with nonsignificant expression changes. (*H*) Overlap in the number of differentially expressed genes in double-knockout and triple-knockout mice. For genes misregulated in triple-knockout muscle, enriched biological pathways are shown, with those common between CDM and triple-knockout muscle highlighted in red. (*I*) Immunofluorescence labeling of wild-type and triple-knockout acetylcholine receptors (AChRs) along the midline of P0 diaphragms (*left*), and quantification based on average midline width is shown (*right*). (*) $P < 0.05$; (**) $P < 0.01$, ANOVA.

skeletal muscle as a benchmark for cell and animal CDM disease models.

Mbnl loss-of-function models for CDM pathogenesis

Although prenatal *DMPK* CTG repeat length is often assessed in unaffected tissues (e.g., chorionic villus sampling), correlations between repeat copy number and CDM disease remain controversial, and pathomechanisms beyond RNA toxicity have been proposed (Barceló et al. 1994; Geifman-Holtzman and Fay 1998; Barbé et al. 2017). For example, the *DMPK* CTG^{exp} is flanked by the upstream *DMWD* and downstream *SIX5* genes, and prior studies have suggested that CDM is associated with disease-specific expression changes in these flanking genes (Klesert et al. 1997, 2000). Mechanistically, this model is supported by CDM-specific hypermethylation upstream of the *DMPK* CTG^{exp} tract (Steinbach et al. 1998; Barbé et al. 2017), which might contribute to altered CUG^{exp} RNA levels and/or influence expression of the downstream *SIX5* gene (Harley et al. 1992, 1993; Steinbach et al. 1998; Castel et al. 2011; Yanovsky-Dagan et al. 2015). However, loss of SIX5 in mice does not produce the neonatal features of CDM, and our analysis failed to detect significant expression differences of either *DMWD* ($\log_2$ fold change = 0.51; adjusted $P = 0.62$) or *SIX5* ($\log_2$ fold change = −0.05; adjusted $P = 1.0$) between control and CDM muscle. Furthermore, DMPK is dispensable for normal muscle development and function in mice (Carrell et al. 2016), supporting a minimal contribution of haploinsufficiency models to CDM pathogenesis.

Given that pre-mRNA processing transitions are a common feature of developing human muscle and that many of these events show evidence of misregulation in CDM, we tested a model in which the presence of toxic CUG^{exp} RNAs in utero disrupts MBNL activity, leading to dysregulation of pre-mRNA processing patterns essential for normal muscle formation and function. Consistent with this model, neonatal *Mbnl* compound knockout mice show characteristic manifestations of CDM, including frequent perinatal lethality, respiratory distress, reduced newborn body weight and postnatal growth, muscle histopathology, congenital spliceopathy, and gene expression abnormalities. These findings extend beyond our previous observations of muscle wasting in *Mbnl* muscle-specific double-knockout mice (Lee et al. 2013) by demonstrating that muscle development is also severely compromised by compound loss of MBNL1 and MBNL2 activity. A central question is which of the hundreds of detectable RNA misprocessing events in CDM muscle truly contribute to abnormal muscle development. It is likely that splicing events with postnatal transitions (e.g., *ATP2A1* e.22 and *INSR* e.11) can be ruled out, as can those with detectable missplicing in *Mbnl1* and *Mbnl2* single-knockout mice, since these animals do not display overt congenital myopathy. We believe our utilization of newborn mouse muscle for transcriptome analysis minimizes the contribution of postnatal muscle degeneration/regeneration to secondary missplicing (Orengo et al. 2011) and supports the identification of disease-initiating events.

An unexpected role for MBNL3 in myogenesis

Given that MBNL1 and MBNL2 loss results in developmentally immature AS and APA patterns, the emergence of adult pattern exons in *Mbnl3* knockout myoblasts was surprising, particularly since all three MBNLs share similar binding site preferences. MBNL3 is a weaker splicing factor than MBNL1 (Sznajder et al. 2016), and it is possible that MBNL3 occupancy partially blocks recruitment of MBNL1 (the stronger splicing factor) to target RNAs, resulting in overall weaker adult pattern splicing activity. In the absence of MBNL3 in *Mbnl3* knockout myoblasts or during normal muscle differentiation where MBNL3 levels decrease, MBNL1 binding availability increases, allowing more robust adult pattern splicing to proceed. Interestingly, RBFOX proteins share binding motifs similar to those of MBNLs and regulate many of the same targets as MBNLs (Singh et al. 2014), suggesting that these factors may coordinate in the precise timing of muscle differentiation events.

As splicing was our major outcome measure in CDM patient and *Mbnl* knockout muscles, potential contributions of RNA mislocalization and altered stability to myopathy are currently unclear, but, based on prior studies, it is likely that these pathways (or other cytoplasmic events) contribute to CDM pathogenesis (Masuda et al. 2012; Wang et al. 2012, 2015). In support of this possibility, MBNL3 is localized primarily in the cytoplasm, and proliferating *Mbnl3* knockout myoblasts show precocious expression of terminal differentiation gene products, including *Dmd*, *Dag*, and *Tnnt2*, but are fusion-defective following induction of myogenic differentiation. Since multiple genes encode factors involved in the signaling, actin dynamics, and vesicular trafficking events required for membrane fusion (Posey et al. 2011; Abmayr and Pavlath 2012), future studies will focus on MBNL3 RNA targets and MBNL3-based mechanisms required for myogenic differentiation. In summary, this study provides testable mammalian models and mechanistic insights into the molecular etiology of CDM disease and affords further insights into the role of RNA processing in embryonic and fetal development.

Materials and methods

Human muscle biopsies and genotyping

Biceps branchii autopsy tissue was obtained from three CDM infants (ages 3, 4, and 15 mo) and three disease controls (SMA type 1, all 3-mo of age). The presence of a *DMPK* CTG^{exp} was confirmed by Southern blot analysis using DNA isolated from skeletal muscle.

Mbnl knockout models

Mbnl1 constitutive (*Mbnl1*$^{\Delta E3/\Delta E3}$), *Mbnl2* conditional (*Mbnl2*$^{C/C}$), compound *Mbnl1; Mbnl2*, and *Mbnl3* isoform (*Mbnl3*$^{\Delta E2/Y}$) knockout mice have been described (Lee et al. 2013; Poulos et al. 2013). The *Mbnl3* conditional whole-locus knockout targeting vector was created using standard recombineering bacterial strains and techniques using protocols 1–4 (http://web.ncifcrf.gov/research/brb/protocol.aspx) and is described more thoroughly in the Supplemental Material. Model genotypes were as follows: wild-type (*Mbnl1*$^{+/+}$; *Mbnl2*$^{C/C}$; *Mbnl3*$^{C/Y}$; *Myog-Cre*$^{-/-}$), 1KO (*Mbnl1*$^{-/-}$; *Mbnl2*$^{C/C}$; *Mbnl3*$^{C/Y}$; *Myog-Cre*$^{-/-}$), 2KO (*Mbnl1*$^{+/+}$; *Mbnl2*$^{C/C}$; *Mbnl3*$^{+/Y}$; *Myog-Cre*$^{+/-}$), double-knockout (*Mbnl1*$^{-/-}$; *Mbnl2*$^{C/C}$; *Mbnl3*$^{+/Y}$; *Myog-Cre*$^{+/-}$), and triple-knockout (*Mbnl1*$^{-/-}$; *Mbnl2*$^{C/C}$; *Mbnl3*$^{C/Y}$; *Myog-Cre*$^{+/-}$). All animal procedures were reviewed and approved by the University of Florida Institutional Animal Care and Use Committee.

RNA-seq and PolyA-seq library preparation

For infant control and CDM samples, RNA-seq libraries were prepared using the SMART-Seq version 4 ultralow input RNA kit (Clontech) per the manufacturer's instructions and sequenced using an Illumina HiSeq 2000. For PolyA-seq libraries, total RNA (130 ng) was used as a starting template, and libraries were generated as described (Batra et al. 2014) with the modification of adding barcodes to the library amplification primers to accommodate multiplex sequencing. For mouse P0 muscle and primary myoblast samples, RNA was isolated using the Direct-Zol RNA miniprep kit (Zymo Research). RNA-seq libraries were prepared from total RNA (500–700 ng) using the stranded RNA-seq kit with RiboErase (Kapa Biosystems) per the manufacturer's protocol. The CDM PolyA-seq and all mouse libraries were sequenced using an Illumina NextSeq 500.

RNA-seq analysis

Reads were aligned to the human (hg19) or mouse (mm10) genomes using HISAT2 (Kim et al. 2015), and primary mapped alignments were retrieved using SAMtools. RSeQC was used to assess quality metrics, including uniform gene coverage, annotated splice junction saturation, annotated gene expression saturation, and transcript integrity values. Quantification of exon inclusion levels was performed using MISO (Katz et al. 2010) based on MISO Annotations version 2.0. Only events with read coverage $\geq$20 and Bayes factor $\geq$5 for at least one pair-wise comparison were considered. To identify reproducible changes between control and experimental groups, monotonicity *Z*-scores (Wang et al. 2015) were calculated for each splicing event and used for filtering ($|Z\text{-score}| \geq 1.5$). To identify conserved missplicing between humans and mice, we focused on CEs and used the University of California at Santa Cruz liftover tool to convert hg19 MISO exon coordinates to mm10 coordinates; 4192 CE splicing events were recovered, and 3652 of these were represented in both the double-knockout and CDM RNA-seq. Events were filtered ($|\Delta\Psi| \geq 0.10$; CI ≤ 0.30; required in both data sets) to retrieve the final event list (Supplemental Table S5). Visualization of splicing results was performed using the Sashimi Plots package.

To identify differentially expressed genes, reads were aligned and associated with genomic features using Quantas (Zhang et al. 2014). Reproducible changes were identified using edgeR (Robinson et al. 2010), and gene lists were filtered ($|\log_2$ fold change$| \geq 1$; adjusted $P \leq 0.01$) to identify significant changes. To identify enriched biological pathways, gene ontology analysis was performed using Metascape (Tripathi et al. 2015) using gene sets from Reactome gene sets, GO Biological Processes, Hallmark gene sets, and Kyoto Encyclopedia of Genes and Genomes pathways and default statistical thresholds (minimum pathway overlap $\geq$3; *P*-value <0.01; enrichment $\geq$1.5). Gene ontology was also performed using GOseq (Young et al. 2010) to correct for potential transcript length biases. For some analyses (e.g., Supplemental Fig. S5I), transcripts per million (TPM) were calculated using Kallisto (Bray et al. 2016), and individual TPM values were summed for a given gene ID. Plotting and correlation tests were performed

using custom scripts largely using the Seaborn Python package (http://seaborn.pydata.org/index.html).

PolyA-seq analysis

PolyA-seq data were processed as described previously (Batra et al. 2014). Briefly, reads were mapped to the human (hg19) genome using OLego (Wu et al. 2013) and filtered to assign high-confidence PASs, and ΔPAS was measured using a modified version of Quantas to identify reproducible changes between control and CDM samples (false discovery rate ≤ 0.05; dI ≥ |0.01|).

RT–PCR splicing validations

RNA was isolated using the Direct-Zol RNA miniprep kit (Zymo Research) along with an on-column gDNA digestion step. Total RNA (500 ng) was used for cDNA synthesis using random hexamers and SuperScript IV RT (Invitrogen) according to the manufacturer's protocol. PCR amplification was performed at an annealing temperature of 58°C using primers positioned in constitutive exons neighboring alternative exons. PCR products were resolved on ethidium bromide-stained agarose gels or a Fragment Analyzer (Advanced Analytical). When necessary, PCR products were digested with appropriate restriction enzymes to facilitate resolution of similarly sized bands.

Muscle histology and diaphragm analysis

P0 hindlimbs were harvested, embedded in OCT, and frozen in liquid nitrogen-cooled 2-methylbutane. Sections (10 µm) were cut using a −20°C cryostat and stained with hematoxylin and eosin (H&E) according to Cure CMD protocol MDC1A_M.1.2.004. Imaging was performed on a Leica DM2000 with a Leica N Plan 5×/0.12 objective and Q-Imaging MicroPublisher 5.0 RTV CCD color camera. Alternatively, sections were labeled with Alexa fluor 488 (AF488)-conjugated wheat germ agglutinin. Confocal imaging was performed on a Zeiss LSM 880 with Airyscan using a plan-apochromat 63×/1.4 oil DIC M27 objective. DAPI and AF488 were excited with a 405-nm diode laser and 488-nm argon laser, respectively. Emission was detected with beam splitters 405/488/561 nm with spectral filters 411/475 nm for DAPI and 517/606 nm for AF488. Single focal planes were acquired and processed with Zen software (intensity thresholds and γ corrections). Myofiber cross-sectional areas were measured for tibialis anterior ($n = 600$) and soleus ($n = 300$) muscles in triplicate using ImageJ. Statistical significance was determined using GraphPad Prism 7 by Kruskal-Wallis test with Dunn's post-hoc (triple asterisks indicate $P < 0.001$). For analysis of AChR organization and size, dissected P0 diaphragms were washed in PBS and fixed in 10% neutral buffered formalin (Fisher) overnight at 4°C. The next day, diaphragms were washed in PBS, and AChRs were labeled using AF488-conjugated bungarotoxin and mounted using Vecta-Shield with DAPI (Vector Laboratories). Confocal imaging was performed as described above with the subsequent modification. AF568 was excited with a 561-nm laser and detected with beam splitter 458/561 nm with spectral filter 568/712 nm. A total of 100–200 slices covered 10–20 µm. Images were processed with Zen 2 and presented as maximum intensity projections from all *z* slices. About 70 neuromuscular junctions were measured with ImageJ software. Statistical significance was determined in GraphPad Prism by Kruskal-Wallis test with Dunn's post-hoc (triple asterisks indicate $P < 0.001$).

Single myofiber isolation

Single myofibers were isolated from EDL muscle after digestion with collagenase (Sigma) for ~1 h at 37°C following previously described protocols (Pasut et al. 2013). Single isolated myofibers were either immediately fixed and stored in PBS at 4°C or cultured for 3 d prior to immunofluorescence experiments. For culturing, floating myofibers were grown at 37°C and 5% CO_2 in myofiber culture medium (DMEM, 20% FBS, 1% chicken embryo extract, 1× penicillin/streptomycin).

Immunofluorescence

Immunofluorescence was performed on single myofibers and primary myoblasts using similar procedures. Samples were fixed in 10% neutral buffered formalin (Fisher) for 10 min at room temperature, washed twice with PBS, and permeabilized using 0.2% Triton-X 100 in PBS for 10 min. Blocking was performed using 5% normal goat serum and 0.1% Triton-X in PBS for 1 h at room temperature followed by incubation with primary antibodies overnight at 4°C. The next day, samples were washed and incubated with AF488-labeled secondary antibodies in blocking solution for 1 h at room temperature, washed, and mounted with Vecta-Shield plus DAPI (Vector Laboratories). PAX7 (1:100; supernatant, Developmental Studies Hybridoma Bank), MYOG (1:500; Santa Cruz Biotechnology, F5D), and MBNL3 (1:500) (Poulos et al. 2013) were used. Image acquisition was performed using a Zeiss Axioskop 2 Mot Plus with Zeiss plan-neofluar 40×/0.75 objective and AxioCam MRC5 CCD camera. DAPI and AF488 signal was acquired using Zeiss filter set 2 and Olympus filter number 41001, respectively.

Isolation and culture of primary myoblasts

Primary myoblasts were isolated based on previously described protocols (Rando and Blau 1994). Briefly, total hindlimb musculature from mice <6 wk of age was minced using sterile razor blades and digested in a collagenase/dispase mixture for 1 h at 37°C and 5% CO_2. The resulting cell slurry was triturated using a serological pipette and filtered through a 100-µm cell strainer. After centrifugation, the cell pellet was resuspended in myoblast growth medium (Ham's F10, 20% FBS, 1× penicillin/streptomycin, 2.5 ng/µL bFGF), and cells were grown on gelatin-coated tissue culture dishes at 37°C and 5% CO_2. Wild-type and *Mbnl3* knockout myoblasts were prepared simultaneously and cultured until nearly >90% of cells stained positive for desmin. To induce in vitro differentiation, cultures at ~70% confluency were switched to differentiation medium (DMEM, 2% horse serum, 1× penicillin/streptomycin).

Accession number

Sequencing data have been deposited in Gene Expression Omnibus under accession number GSE97806.

Acknowledgments

We thank the CDM and SMA families for their important contributions to this study, University of Florida Research Computing for computational resources, and A. Berglund, J. Bubenik, and K. Esser for comments on the manuscript. This study was supported by grants from the National Institutes of Health (NS058901 and NS98819 to M.S.S., and OD017865 to E.T.W.), Grants-in-Aid for Challenging Exploratory Research and Scientific Research (B) from the Japan Society for the Promotion of

Science (KAKENHI 15K15339 and 16H05321 to M.N.), and a post-doctoral fellowship award from the Myotonic Dystrophy and Wyck Foundations to Ł.J.S.

References

Abmayr SM, Pavlath GK. 2012. Myoblast fusion: lessons from flies and mice. *Development* **139:** 641–656.

Bachinski LL, Baggerly KA, Neubauer VL, Nixon TJ, Raheem O, Sirito M, Unruh AK, Zhang J, Nagarajan L, Timchenko LT, et al. 2014. Most expression and splicing changes in myotonic dystrophy type 1 and type 2 skeletal muscle are shared with other muscular dystrophies. *Neuromuscul Disord* **24:** 227–240.

Barbé L, Lanni S, López-Castel A, Franck S, Spits C, Keymolen K, Seneca S, Tomé S, Miron I, Letourneau J, et al. 2017. CpG methylation, a parent-of-origin effect for maternal-biased transmission of congenital myotonic dystrophy. *Am J Hum Genet* **100:** 488–505.

Barceló JM, Pluscauskas M, MacKenzie AE, Tsilfidis C, Narang M, Korneluk RG. 1994. Additive influence of maternal and offspring DM-kinase gene CTG repeat lengths in the genesis of congenital myotonic dystrophy. *Am J Hum Genet* **54:** 1124–1125.

Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, et al. 2014. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell* **56:** 311–322.

Bentzinger FC, Wang Y, Rudnicki MA. 2012. Building muscle: molecular regulation of myogenesis. *Cold Spring Harb Perspect Biol* **4:** a008342.

Bland CS, Wang ET, Vu A, David MP, Castle JC, Johnson JM, Burge CB, Cooper TA. 2010. Global regulation of alternative splicing during myogenic differentiation. *Nucleic Acids Res* **38:** 7651–7664.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34:** 525–527.

Brinegar AE, Cooper TA. 2016. Roles for RNA-binding proteins in development and disease. *Brain Res* **1647:** 1–8.

Brinegar AE, Xia Z, Loehr JA, Li W, Rodney GG, Cooper TA. 2017. Extensive alternative splicing transitions during postnatal skeletal muscle development are required for calcium handling functions. bioRxiv doi.org/10.1101/124230.

Carrell ST, Carrell EM, Auerbach D, Pandey SK, Bennett CF, Dirksen RT, Thornton CA. 2016. Dmpk gene deletion or antisense knockdown does not compromise cardiac or skeletal muscle function in mice. *Hum Mol Genet* **25:** 4328–4338.

Castel A, Nakamori M, Tomé S, Chitayat D, Gourdon G, Thornton CA, Pearson CE. 2011. Expanded CTG repeat demarcates a boundary for abnormal CpG methylation in myotonic dystrophy patient tissues. *Hum Mol Genet* **20:** 1–15.

Dowling JJ, Vreede AP, Kim S, Golden J, Feldman EL. 2008. Kindlin-2 is required for myocyte elongation and is essential for myogenesis. *BMC Cell Biol* **9:** 36.

Fugier C, Klein AF, Hammer C, Vassilopoulos S, Ivarsson Y, Toussaint A, Tosch V, Vignaud A, Ferry A, Messaddeq N, et al. 2011. Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat Med* **17:** 720–725.

Geifman-Holtzman O, Fay K. 1998. Prenatal diagnosis of congenital myotonic dystrophy and counseling of the pregnant mother: case report and literature review. *Am J Med Genet* **78:** 250–253.

Harley HG, Brook JD, Rundle SA, Crow S, Reardon W, Buckler AJ, Harper PS, Housman DE, Shaw DJ. 1992. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* **355:** 545–546.

Harley HG, Rundle SA, MacMillan JC, Myring J, Brook JD, Crow S, Reardon W, Fenton I, Shaw DJ, Harper PS. 1993. Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am J Hum Genet* **52:** 1164–1174.

Ho G, Cardamone M, Farrar M. 2015. Congenital and childhood myotonic dystrophy: current aspects of disease and future directions. *World J Pediatr* **4:** 66–80.

Johnson NE, Butterfield R, Berggren K, Hung M, Chen W, DiBella D, Dixon M, Hayes H, Pucillo E, Bounsanga J, et al. 2016. Disease burden and functional outcomes in congenital myotonic dystrophy. *Neurology* **87:** 160–167.

Joseph JT, Richards CS, Anthony DC, Upton M, Perez-Atayde AR, Greenstein P. 1997. Congenital myotonic dystrophy pathology and somatic mosaicism. *Neurology* **49:** 1457–1460.

Kalsotra A, Cooper TA. 2011. Functional consequences of developmentally regulated alternative splicing. *Nature Rev Genet* **12:** 715–729.

Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, Cooper TA. 2008. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc Natl Acad Sci* **105:** 20333–20338.

Katz Y, Wang ET, Airoldi EM, Burge CB. 2010. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods* **7:** 1009–1015.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12:** 357–360.

Klesert TR, Otten AD, Bird TD, Tapscott SJ. 1997. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat Genet* **16:** 402–406.

Klesert TR, Cho DH, Clark JI, Maylie J, Adelman J, Snider L, Yuen EC, Soriano P, Tapscott SJ. 2000. Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat Genet* **25:** 105–109.

Lee K-S, Cao Y, Witwicka HE, Tom S, Tapscott SJ, Wang EH. 2010. RNA-binding protein Muscleblind-like 3 (MBNL3) disrupts myocyte enhancer factor 2 (Mef2) β-exon splicing. *J Biol Chem* **285:** 33779–33787.

Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, Warren SA, Chamberlain CM, Finn D, Hong H, et al. 2013. Compound loss of muscleblind-like function in myotonic dystrophy. *EMBO Mol Med* **5:** 1887–1900.

Lin X, Miller JW, Mankodi A, Kanadia RN, Yuan Y, Moxley RT, Swanson MS, Thornton CA. 2006. Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. *Hum Mol Genet* **15:** 2087–2097.

Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA. 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.

Masuda A, Andersen H, Doktor T, Okamoto T, Ito M, Andresen B, Ohno K. 2012. CUGBP1 and MBNL1 preferentially bind to 3′ UTRs and facilitate mRNA decay. *Sci Rep* doi: 10.1038/srep00209.

Myring J, Meredith AL, Harley HG, Kohn G, Norbury G, Harper PS, Shaw DJ. 1992. Specific molecular prenatal diagnosis for

the CTG mutation in myotonic dystrophy. *J Med Genet* **29:** 785–788.

Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, Dekdebrun J, Heatwole CR, McDermott MP, Chen T, et al. 2013. Splicing biomarkers of disease severity in myotonic dystrophy. *Ann Neurol* **74:** 862–872.

Orengo JP, Ward AJ, Cooper TA. 2011. Alternative splicing dysregulation secondary to skeletal muscle regeneration. *Ann Neurol* **69:** 681–690.

Pasut A, Jones AE, Rudnicki MA. 2013. Isolation and culture of individual myofibers and their satellite cells from adult skeletal muscle. *J Vis Exp* **22:** e50074.

Pedrotti S, Giudice J, Dagnino-Acosta A, Knoblauch M, Singh RK, Hanna A, Mo Q, Hicks J, Hamilton S, Cooper TA. 2015. The RNA-binding protein Rbfox1 regulates splicing required for skeletal muscle structure and function. *Hum Mol Genet* **24:** 2360–2374.

Posey AD, Demonbreun A, McNally EM. 2011. Ferlin proteins in myoblast fusion and muscle growth. *Curr Top Dev Biol* **96:** 203–230.

Poulos MG, Batra R, Li M, Yuan Y, Zhang C, Darnell RB, Swanson MS. 2013. Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. *Hum Mol Genet* **22:** 3547–3558.

Rando TA, Blau HM. 1994. Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy. *J Cell Biol* **125:** 1275–1287.

Rau F, Lainé J, Ramanoudjame L, Ferry A, Arandel L, Delalande O, Jollet A, Dingli F, Lee K-YY, Peccate C, et al. 2015. Abnormal splicing switch of DMD's penultimate exon compromises muscle fibre maintenance in myotonic dystrophy. *Nat Commun* **6:** 7205.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26:** 139–140.

Runfola V, Sebastian S, Dilworth JF, Gabellini D. 2015. Rbfox proteins regulate tissue-specific alternative splicing of Mef2D required for muscle differentiation. *J Cell Sci* **128:** 631–637.

Scotti MM, Swanson MS. 2016. RNA mis-splicing in disease. *Nature Rev Genet* **17:** 19–32.

Singh RK, Xia Z, Bland CS, Kalsotra A, Scavuzzo MA, Curk T, Ule J, Li W, Cooper TA. 2014. Rbfox2-coordinated alternative splicing of Mef2d and Rock2 controls myoblast fusion during myogenesis. *Mol Cell* **55:** 592–603.

Steinbach P, Gläser D, Vogel W, Wolf M, Schwemmle S. 1998. The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. *Am J Hum Genet* **62:** 278–285.

Sznajder ŁJ, Michalak M, Taylor K, Cywoniuk P, Kabza M, Wojtkowiak-Szlachcic A, Matłoka M, Konieczny P, Sobczak K. 2016. Mechanistic determinants of MBNL activity. *Nucleic Acids Res* **44:** 10326–10342.

Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32:** 381–386.

Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder L, et al. 2015. Meta- and orthogonal integration of influenza 'OMICs' data defines a role for UBR4 in virus budding. *Cell Host Microbe* **18:** 723–735.

Tsilfidis C, MacKenzie AE, Mettler G, Barceló J, Korneluk RG. 1992. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nat Genet* **1:** 192–195.

Wagner SD, Struck AJ, Gupta R, Farnsworth DR, Mahady AE, Eichinger K, Thornton CA, Wang ET, Berglund AJ. 2016. Dose-dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. *PLoS Genet* **12:** e1006316.

Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S, et al. 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150:** 710–724.

Wang ET, Ward AJ, Cherone JM, Giudice J, Wang TT, Treacy DJ, Lambert NJ, Freese P, Saxena T, Cooper TA, et al. 2015. Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. *Genome Res* **25:** 858–871.

Wong LJ, Ashizawa T. 1997. Instability of the (CTG)n repeat in congenital myotonic dystrophy. *Am J Hum Genet* **61:** 1445–1448.

Wong LJ, Ashizawa T, Monckton DG, Caskey CT, Richards CS. 1995. Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am J Hum Genet* **56:** 114–122.

Wu J, Anczuków O, Krainer AR, Zhang MQ, Zhang C. 2013. OLego: fast and sensitive mapping of spliced mRNA-Seq reads using small seeds. *Nucleic Acids Res* **41:** 5149–5163.

Yanovsky-Dagan S, Avitzour M, Altarescu G, Renbaum P, Eldar-Geva T, Schonberger O, Mitrani-Rosenbaum S, Levy-Lahad E, Birnbaum RY, Gepstein L, et al. 2015. Uncovering the role of hypermethylation by CTG expansion in myotonic dystrophy type 1 using mutant human embryonic stem cells. *Stem Cell Reports* **5:** 221–231.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* **11:** 1–12.

Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, et al. 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34:** 11929–11947.