Zebrafish *mbnl* mutants model physical and molecular phenotypes of myotonic dystrophy

Melissa N. Hinman[1], Jared I. Richardson[4,6], Rose A. Sockol[1], Eliza D Aronson[1], Sarah J. Stednitz[2], Katrina N. Murray[3], J. Andrew Berglund[4,6], and Karen Guillemin[1,5*]

[1]Institute of Molecular Biology, University of Oregon, Eugene, Oregon, United States of America

[2]Institute of Neuroscience, University of Oregon, Eugene, Oregon, United States of America

[3]Zebrafish International Resource Center, University of Oregon, Eugene, Oregon, United States of America

[4]RNA Institute, State University of New York at Albany, Albany, New York, United States of America

[5]Humans and the Microbiome Program, CIFAR, Toronto, Ontario, Canada

[6]Department of Biochemistry & Molecular Biology, Center for NeuroGenetics, University of Florida, Gainesville, Florida, United States of America

*Corresponding author

E-mail: kguillem@uoregon.edu (KG)

Short title: Zebrafish *mbnl* mutant models of myotonic dystrophy

Abstract

The muscleblind RNA binding proteins (MBNL1, MBNL2, and MBNL3) are highly conserved across vertebrate species and are important regulators of alternative splicing and other aspects of RNA metabolism. Loss of MBNL protein function through sequestration by CUG or CCUG RNA repeats is largely responsible for the phenotypes of the human genetic disorder myotonic dystrophy (DM). We generated the first stable zebrafish (*Danio rerio*) models of DM-associated *MBNL* loss of function through CRISPR-based mutation of the three zebrafish *mbnl* genes. In contrast to mouse models, zebrafish double and triple homozygous *mbnl* mutants were viable to adulthood. Zebrafish *mbnl* mutants displayed disease-relevant physical phenotypes including decreased body size and impaired movement. They also exhibited widespread alternative splicing changes, including the misregulation of many DM-relevant exons. Physical and molecular phenotypes were more severe in compound *mbnl* mutants than in single *mbnl* mutants, suggesting partially redundant functions of Mbnl proteins in zebrafish. The high fecundity and larval optical transparency of this complete series of zebrafish *mbnl* mutants will make them useful for studying DM-related phenotypes and how individual Mbnl proteins contribute to them, and for testing potential therapeutics.

Author summary

The human genetic disorder myotonic dystrophy (DM), which leads to a range of dysfunction in muscles and other parts of the body, is caused in part by decreased functioning of the Mbnl family of proteins. We generated a series of zebrafish mutants that mimicked DM-associated Mbnl loss of function through mutation of the zebrafish

mbnl genes. The mutant series recapitulated a range of DM-associated symptoms, including decreased body size and impaired motor function, with greater gene loss resulting in more severe physical defects. The mutant series also exhibited widespread and dose-dependent changes in alternative splicing, a molecular process whose dysregulation is known to underlie DM symptoms. The new *mbnl* mutant zebrafish, which can be generated in large numbers relatively inexpensively and studied as optically transparent, live animals, will be important complements to existing mouse models for studying DM. Specifically, they will be valuable for studying the onset of disease symptoms and screening for potential therapeutics in live animals.

Introduction

The muscleblind (MBNL) family of RNA binding proteins is highly conserved in structure and function across multicellular species, and vertebrates have three family members: MBNL1, MBNL2, and MBNL3 [1]. MBNL proteins use zinc finger motifs (two pairs in the case of vertebrates) to recognize and bind to consensus YGCY sequences (where Y is a pyrimidine) in target RNAs and regulate multiple aspects of RNA metabolism, the most well-characterized of which is alternative splicing [2-9]. In general, MBNL proteins promote exon skipping when bound to YGCY sequences in the intron upstream of an alternative exon, and promote exon inclusion when bound downstream of an alternative exon [3, 8, 10]. Regulation of alternative splicing by MBNL proteins is critical for the molecular function of many target genes. For example, MBNL proteins promote skipping of exon 5 in the *MBNL1* and *MBNL2* mRNAs by binding to highly conserved YGCY elements in the upstream intron, affecting the subcellular localization and alternative splicing regulatory activity of the resulting proteins [11-13].

The regulation of RNA metabolism by MBNL proteins is crucial for the proper functioning of many body systems, as shown by the multi-systemic phenotypes of the human genetic *MBNL* loss of function disorder, myotonic dystrophy (DM). In individuals affected by DM, expression of an expanded CTG repeat in the 3' UTR of the *DMPK* gene (DM1) or an expanded CCTG repeat in an intron of the *CNBP* gene (DM2) leads to the formation of RNA stem loop structures that contain dozens to thousands of YGCY motifs that bind and sequester MBNL proteins, preventing them from performing their normal functions [14-19] (Fig 1A). In DM1, disease severity increases with increasing CTG repeat length, likely due in part to increased sequestration of MBNL proteins [20]. DM is most well-known for its skeletal muscle phenotypes, including weakness, myotonia, atrophy, and pain. However, individuals with DM experience a wide variety of symptoms in many tissues that can be as debilitating as muscle phenotypes, including but not limited to cataracts, breathing difficulties, behavioral and psychological disorders, sleep disorders, insulin resistance, heart conduction abnormalities, cardiomyopathy, and alterations in the motility and microbiota of the gut [21-26]. Misregulation of specific alternative splicing events was shown to contribute to disease-related phenotypes [26]. For example, missplicing of *RYR1* and *ATP2A1* contributed to altered calcium homeostasis in DM1 muscle [27, 28].

Mouse *Mbnl* mutants greatly improved our understanding of the biological roles of individual MBNL proteins and how they may contribute to DM. Mouse *Mbnl1* is widely expressed and *Mbnl1*$^{-/-}$ mice exhibited changes in alternative splicing, myotonia, myopathy, and cataracts, but did not have other DM-related phenotypes such as muscle weakness, suggesting that loss of *Mbnl1* function alone was not sufficient to cause all

DM-related phenotypes [29, 30]. *Mbnl2* expression is enriched in the brain, and *Mbnl2*$^{-/-}$ mice exhibited DM-like CNS phenotypes including missplicing in the brain, abnormal sleep patterns, and spatial memory deficits [31]. Mouse *Mbnl3* is expressed primarily during embryonic development and injury-induced adult skeletal muscle regeneration. *Mbnl3*$^{-/-}$ mice exhibited delays in injury-induced muscle regeneration, accelerated onset of age-related pathologies, and gene expression changes, but displayed minimal changes in alternative splicing, suggesting that *Mbnl3* played a more modest role in DM-related missplicing [32, 33]. Compound *Mbnl* mutant mice had more severe phenotypes than single mutant mice, suggesting that Mbnl proteins had redundant functions in mice. *Mbnl1*$^{-/-}$*;Mbnl2*$^{-/-}$ mice were embryonic lethal, while *Mbnl1*$^{-/-}$*;Mbnl2*$^{+/-}$ mice were viable but had severe missplicing, decreased body weight, myotonia, progressive muscle weakness, and reduced lifespan [30]. Muscle-specific double (*Mbnl1*$^{-/-}$*; Mbnl2*$^{C/C}$*; Mbnl3*$^{+/Y}$*; Myog-Cre*$^{+/-}$) and triple (*Mbnl1*$^{-/-}$*; Mbnl2*$^{C/C}$*; Mbnl3*$^{C/Y}$*; Myog-Cre*$^{+/-}$) mutant mice were reduced in size, had dramatic and widespread alternative splicing changes, muscle weakness, and most died in the neonatal period due to respiratory distress [34]. Compound *Mbnl* mutant mice recapitulated many of the phenotypes associated with severe forms of DM, but their utility in future experiments is limited because they are difficult to generate in large numbers.

We sought to model DM-associated loss of MBNL function using another vertebrate organism, the zebrafish. Like mice, zebrafish have one ortholog of each of the human *MBNL* genes [35]. There are several advantages of using zebrafish over mice to model DM. Zebrafish embryos develop rapidly and externally, making studies of developmental phenotypes easier than in mice. Zebrafish embryos are optically

transparent, which allows for direct analysis of DM-related phenotypes such as altered gut motility and abnormal heart rhythm. Finally, zebrafish produce hundreds of embryos at once, and adults can be maintained much more cheaply than mice, which enables experiments to be performed with large numbers of animals and improves statistical power when studying subtle or variable phenotypes. Two zebrafish models of *MBNL* loss of function disorders were generated previously, but both had severe limitations. Zebrafish embryos that were injected with CUG repeat RNA had subtle developmental abnormalities, but did not exhibit alternative splicing changes, and the injected RNA was degraded by two days post fertilization (dpf) [36]. This model did not allow for the study of phenotypes beyond early development, and could not distinguish between the effects of loss of individual Mbnl proteins. Another model employed morpholinos, but these are also transient and can have off-target effects. Morpholino-based knockdown of *mbnl2* in zebrafish embryos resulted in profound morphological abnormalities that were inconsistent with the much milder phenotypes observed in *Mbnl2*$^{-/-}$ mice [31, 37].

In the current study, we used CRISPR-Cas9 based mutagenesis to make homozygous zebrafish *mbnl1*, *mbnl2,* and *mbnl3* loss of function mutants, which were crossed to generate double (*mbnl1/mbnl2*, *mbnl2/mbnl3,* and *mbnl1/mbnl3*) and triple (*mbnl1/mbnl2/mbnl3*) homozygous mutants, all of which were viable to adulthood. Zebrafish *mbnl* mutants exhibited DM-relevant physical phenotypes including decreased body size and impaired motor function. They also exhibited widespread alternative splicing changes, including many of the same changes that were present in individuals with DM and mouse *Mbnl* mutants. These alternative splicing changes occurred both in larval and adult fish, but were most dramatic in adult skeletal and heart muscle. As in

mice, double and triple homozygous zebrafish *mbnl* mutants exhibited more severe phenotypes than single mutants. Thus, zebrafish *mbnl* mutants had physical and molecular phenotypes consistent with those present in DM, and are powerful new vertebrate models for studies of *MBNL* function.

Results

A comprehensive set of zebrafish *mbnl* mutants was generated

To model DM-associated *MBNL* loss of function in zebrafish, we used CRISPR/Cas9 technology to mutate zebrafish *mbnl1*, *mbnl2,* and *mbnl3*. We targeted constitutively included exons in each zebrafish *mbnl* gene [35], and established two independent *mbnl1* mutant lines ($1^{A/A}$ and $1^{B/B}$), one *mbnl2* mutant line ($2^{A/A}$), and three independent *mbnl3* mutant lines ($3^{A/A}$, $3^{B/B}$, $3^{C/C}$) (Fig 1B, S1 Table, and S2 Table). We were unable to examine protein levels in the mutant lines, as several commercially available antibodies did not recognize zebrafish Mbnl proteins (see materials and methods for panel of antibodies tested). However, each zebrafish *mbnl* mutation was predicted to result in loss of protein function due to the direct insertion of an early stop codon (3^{C}), or a frameshift leading to a premature termination codon (1^{A}, 1^{B}, 2^{A}, 3^{A}, and 3^{B}) (Fig 1B and S2 Table). Each mutation was predicted to disrupt one or more of the zinc fingers that are important for RNA binding (Fig 1B) [9]. Homozygous larval mutants exhibited decreased ($2^{A/A}$, $3^{A/A}$, and $3^{B/B}$) or unchanged ($1^{A/A}$, $1^{B/B}$, and $3^{C/C}$) levels of the mRNA expressed from the mutated gene, with minimal effect on the expression of the other *mbnl* family members (S1 Fig).

All homozygous *mbnl1, mbnl2,* and *mbnl3* mutants survived to adulthood in roughly Mendelian ratios (S3 Table), were fertile, and did not exhibit the dramatic

morphological phenotypes that were previously observed in *mbnl2* morpholino-injected larvae [37]. We crossed single *mbnl* mutant fish to generate double *mbnl* mutants ($1^{B/B};2^{A/A}$, $1^{B/B};3^{C/C}$, and $2^{A/A};3^{C/C}$). All double *mbnl* mutants were also viable to adulthood and fertile (S3 Table). This contrasted with mouse models, in which $mbnl^{-/-};mbnl2^{-/-}$ mutants were embryonic lethal, and to our knowledge represented the first time that *mbnl1/mbnl3* and *mbnl2/mbnl3* double homozygous mutant animals were generated [30]. We further crossed double homozygous animals to generate triple homozygous mutants ($1^{B/B};2^{A/A};3^{C/C}$). Surprisingly, these fish were viable to adulthood, albeit in lower than expected numbers (S3 Table), but they were unable to produce embryos except through *in vitro* fertilization. In summary, we created the first complete panel of vertebrate *mbnl1, mbnl2,* and *mbnl3* single, double, and triple homozygous mutants for modeling DM and studying the functions of Mbnl proteins.

Mutation of *mbnl* genes led to decreased zebrafish size

Mouse compound *Mbnl* loss of function mutants, but not single mutants, were significantly smaller than WT mice [30, 34]. To determine whether loss of *mbnl* function had a similar effect in zebrafish, we raised WT, double homozygous *mbnl* mutant, and triple homozygous *mbnl* mutant fish under identical conditions and measured their lengths in early adulthood. WT fish were significantly larger than $1^{B/B};3^{C/C}$ fish, which in turn were significantly larger than $1^{B/B};2^{A/A}$ fish (Fig 2A). $2^{A/A};3^{C/C}$ fish were slightly smaller than WT fish, but the difference was not statistically significant (Fig 2A). Triple homozygous *mbnl* mutant fish were similar in size to $1^{B/B};2^{A/A}$ fish (Fig 2A). Interestingly, mutation of *mbnl1*, *mbnl2,* or *mbnl3* alone was not sufficient to decrease adult zebrafish size (S2A,B,C Fig). Taken together, these results suggested that

zebrafish Mbnl proteins have at least partially redundant functions, as the growth defect was more dramatic in compound mutants than in single mutants.

To assess whether zebrafish *mbnl* mutants also had a growth defect during larval development, we measured 7 dpf larval fish whose nutrition was derived from the yolk rather than feeding. All double homozygous *mbnl* mutant larvae were similar in size to WT, except for $2^{A/A};3^{C/C}$ larvae, which were slightly smaller than WT (Fig 2B). $1^{B/B};2^{A/A};3^{C/C}$ larvae were similar in size to their $1^{+/B};2^{A/A};3^{C/C}$ and $2^{A/A};3^{C/C}$ clutch mates, and all single homozygous *mbnl* mutant larvae were either similar in size or slightly larger than WT (Fig 2C and S2D Fig). These results suggest that the reduced size phenotype in zebrafish *mbnl* mutants arose later in development.

Histological analysis of adult skeletal muscle was performed by an expert zebrafish pathologist to determine whether the fish genotypes with the most profound size phenotypes, $1^{B/B;}2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$, also had changes in adult muscle fiber morphology. We did not observe obvious signs of DM-related muscle histopathology in the *mbnl* mutant fish (S3 Fig), such as myofiber atrophy or centralized nuclei as has been described in compound *mbnl* mutant mice and in individuals with DM [26, 34], but we cannot rule out the possibility that the mutant fish had subtle alterations in muscle structure or function.

Zebrafish *mbnl* mutants had altered movement

Given that individuals with DM have impaired motor function, we asked whether zebrafish *mbnl* mutants exhibited altered swimming behavior. We performed a test in which individual adult fish were introduced to a novel tank and their position within the tank was tracked for five minutes (Fig 3A,B). A similar test was previously performed in

DM patients, in which the distance walked in six minutes was used as a measure of motor function [38, 39]. While all single homozygous mutants swam equal or greater distance than WT, two of the double homozygous *mbnl* mutants ($1^{B/B};2^{A/A}$ and $1^{B/B};3^{C/C}$) and the $1^{B/B};2^{A/A};3^{C/C}$ mutants swam significantly decreased distances compared to WT (Fig 3C,D,E). One explanation for this result is impaired motor function, although this phenotype needs to be explored further. While performing these tests, we noticed a striking difference in the amount of time *mbnl* mutants and WT fish spent near the bottom of the tank. We quantified this by measuring the distance from the bottom of the tank in each frame for each fish and averaging this over the recording period. Except for $2^{A/A};3^{C/C}$ mutants, all *mbnl* mutants spent significantly more time toward the bottom of the tank than WT fish (Fig 3F,G,H). Avoiding the top of the tank is a well-known indicator of anxiety in zebrafish, and future tests will explore whether *mbnl* mutants exhibit other signs of anxiety or perhaps have a morphological or motor deficit that explains their lower swimming position [40]. Overall, these results indicated that loss of *mbnl* function resulted in altered movement in fish, but the exact mechanisms underlying these changes are unknown.

Zebrafish *mbnl* mutants displayed changes in alternative splicing across tissues

Given the DM-relevant size and movement phenotypes of *mbnl* mutant fish, we next asked whether zebrafish *mbnl* mutants modeled the alternative splicing changes observed in humans with DM and in mouse DM models. Our initial studies used the extensively characterized suppression of *mbnl1* exon 5 and *mbnl2* exon 5 inclusion by Mbnl proteins as a model system [11]. The intronic sequences immediately upstream of *mbnl1* and *mbnl2* exon 5, which contain putative YGCY Mbnl protein binding sites, are

highly conserved between human and zebrafish, suggesting functional importance (S4 and S5 Fig).

We harvested total RNA from whole larval zebrafish as well as adult skeletal muscle, heart, brain, cornea, and intestine. Quantitative RT-PCR analyses of WT fish indicated that *mbnl1, mbnl2*, and *mbnl3* mRNAs were all present in larvae and across adult tissues, and that levels of all three were highest in skeletal muscle and brain (S6A,B,C Fig). Inclusion of *mbnl1* exon 5 and *mbnl2* exon 5, as measured by semi-quantitative RT-PCR, was significantly elevated in zebrafish *mbnl* mutants across all tissues (Fig 4 and S7 Fig). The phenotype was most dramatic in skeletal muscle and heart, where *mbnl1* and *mbnl2* exon 5 were almost entirely skipped in WT and predominantly included in $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$ mutants (Fig 4B,C and S7B,C Fig). The brain and cornea exhibited *mbnl* splicing patterns quite similar to one another, but the magnitudes of splicing changes between WT and mutants were slightly lower than those observed in muscle (Fig 4D,E and S7D,E Fig). Modest but significant splicing changes were observed in whole larvae and intestine (Fig 4A,F and S7A,F Fig).

In most tissues, double and triple mutants exhibited larger *mbnl1* and *mbnl2* splicing changes than single mutants (Fig 4 and S7 Fig). This is consistent with the idea that zebrafish Mbnl proteins, like mouse Mbnl proteins, have partially redundant functions when it comes to splicing regulation [30, 34]. Strikingly, the genotypes with the strongest splicing phenotypes, $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$, also had dramatic size and movement phenotypes (Figs 2-4 and S7 Fig).

In mouse model and DM patient tissues, it was previously demonstrated that some regulated exons were much more sensitive than others to changes in overall

levels of free MBNL proteins [41]. It appears that zebrafish also followed this pattern. For example, in skeletal muscle and heart, *mbnl1* mutation alone was sufficient to increase *mbnl1* exon 5 inclusion, whereas mutation of both *mbnl1* and *mbnl2* was required to increase *mbnl2* exon 5 inclusion (Fig 4B,C and S7B,C Fig). Overall, these results suggest that, as in other systems, zebrafish Mbnl proteins work in concert to regulate alternative splicing across tissues.

Misregulation of alternative splicing was widespread in zebrafish *mbnl* mutants

We expanded our analysis to determine whether our findings related to misregulation of *mbnl1* exon5 and *mbnl2* exon 5 alternative splicing in *mbnl* mutant fish applied transcriptome wide. We performed RNA-Seq analysis using RNA isolated from the skeletal muscle of adult WT, $1^{B/B}$, $2^{A/A}$, $3^{C/C}$, $1^{B/B}$;$2^{A/A}$, $1^{B/B;}3^{C/C}$, $2^{A/A}$;$3^{C/C}$, and $1^{B/B}$;$2^{A/A}$;$3^{C/C}$ zebrafish. There were no significant changes in the levels of total *mbnl1*, *mbnl2*, or *mbnl3* mRNAs in mutants compared to WT (S8 Fig). Hundreds of significantly misregulated RNA alternative splicing events were detected in all *mbnl* mutants, with misregulated cassette exons outnumbering alternative 3' and 5' splice sites, retained introns, and mutually exclusive exons (Fig 5A and S4 Table). $1^{B/B}$;$2^{A/A}$ and $1^{B/B}$;$2^{A/A}$;$3^{C/C}$ fish had the largest number of significantly misregulated splicing events (Fig 5A). We analyzed the magnitude of the splicing change (absolute value of the change in percent spliced in (ΔPSI) between WT and mutant) for the 62 cassette exon events that were significantly misregulated in the majority of *mbnl* mutants (at least 4 of 7). The highest magnitude of splicing change across events was observed in $1^{B/B}$;$2^{A/A}$ and $1^{B/B}$;$2^{A/A}$;$3^{C/C}$ fish (Fig 5B).

We next asked which genotypes were most similar in their overall splicing phenotypes. All cassette exon splicing changes that were observed in any mutant compared to WT were compiled into a single list, and the top 800 events with the largest variability in PSI were identified. Using the gap statistic, we determined that all samples could be clustered into two groups based on the splicing phenotype for those events. A K-means cluster analysis with two centers was then performed, in which all WT and mutant samples were analyzed for their similarity in PSI for the 800 events. The analysis indicated that all $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$ fish were more similar to each other than they were to WT fish or any of the other *mbnl* mutant fish (Fig 5C). Taken together, these results indicated that compound loss of function of zebrafish Mbnl proteins led to widespread and robust changes in alternative splicing.

We analyzed the frequency of the previously described Mbnl protein binding motif, YGCY, within and surrounding the cassette exons that were significantly misregulated in $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$ mutants as well as those misregulated cassette exons that were in common between the two mutants. YGCY motifs were enriched in the introns downstream of exons whose inclusion was activated by the presence of Mbnl proteins, while YGCY motifs were enriched upstream and within the cassette exons that were repressed by Mbnl proteins (Fig 5D). This was consistent with findings in other DM model systems, and suggests that zebrafish Mbnl proteins played a direct role in regulating the inclusion of cassette exons[1, 3, 8, 10].

Zebrafish *mbnl* mutants exhibited disease-relevant alternative splicing changes

Given the widespread changes in alternative splicing in *mbnl* mutant zebrafish, we asked whether these changes were conserved with those identified in human DM

patients. Using publicly available datasets, we identified cassette exons whose inclusion was significantly misregulated in DM1 tibialis anterior muscle biopsy tissues or in DM1 patient derived myoblasts compared to healthy control tissues, and compared them with the zebrafish RNA-Seq data. We identified 25 orthologous cassette exons that were misregulated in both $1^{B/B};2^{A/A};3^{C/C}$ mutant fish and in DM1 tibialis muscle and 40 that were misregulated in both $1^{B/B};2^{A/A};3^{C/C}$ mutant fish and in the DM1 myoblasts (Fig 6A, S9A Fig, and S5 Table). The inclusion of these alternative exons tended to be misregulated in the same direction in zebrafish DM models and in DM1 patient-derived tissues (Fig 6A and S9A Fig). As anticipated, *mbnl1* exon 5 and *mbnl2* exon 5 appeared on both lists of orthologous misregulated cassette exons, and PSI values were strikingly similar as determined by RNA-Seq and RT-PCR (Figs 4B and 6B and S7B and S9B Figs). Two other randomly selected orthologous exons, *aplp2* exon 7 and *atp6v1h* exon6, also showed similar splicing phenotypes by RNA-Seq and RT-PCR, and contained potential YGCY Mbnl binding motifs in their surrounding intronic sequences (S9C,D,E,F, S10, and S11 Figs). Fig 6A and S9A Fig do not represent an exhaustive list of orthologous misregulated exons, as a manual literature search revealed several additional orthologs of disease-associated exons that were misregulated in *mbnl* mutant model fish, some of which were only annotated in an earlier zebrafish genome assembly than the one used to generate Fig 6A and S9A (GRCz10, S4 Table). For example, the decreased inclusion of alternative exons within the human *ATP2A1* and *RYR1* genes has been shown to contribute to altered calcium homeostasis in DM muscle, and we detected decreased inclusion of orthologous exons in zebrafish *mbnl* mutants (Fig 6C,D,E,F, and S12 and S13 Fig)[27, 28]. Likewise, we observed dramatic

mis-splicing of zebrafish *ank3b* exon 36, the human ortholog of which is known to be misregulated in human DM1 heart samples (Fig 6G,H and S14 Fig)[42]. Taken together, these results indicate that the coordinated splicing regulation program of the Mbnl proteins is well-conserved between zebrafish and humans.

Discussion

For the first time, we have modeled DM-associated *MBNL* loss of function by generating a complete panel of vertebrate *mbnl* single, double, and triple homozygous mutants. These mutants will be invaluable for investigating how vertebrate Mbnl proteins contribute both individually and in concert to DM-relevant phenotypes in a living organism. Existing *Mbnl* mutant mice have provided important insights into the biological functions of individual Mbnl proteins [29-34]. However, the compound *Mbnl* mutant mouse models that most closely mimic DM phenotypes are challenging to breed in large numbers and most die during the neonatal period, making studies of adult phenotypes difficult. In contrast, all double and triple homozygous *mbnl* mutant zebrafish are viable to adulthood, and fertile double homozygous mutants can be incrossed to generate hundreds of mutants for experiments. Even triple homozygous *mbnl* mutants, which do not mate on their own, can be generated in large numbers through *in vitro* fertilization or through incrosses of $1^{+/B};2^{A/A};3^{C/C}$ fish. The ability to study hundreds of zebrafish mutants at once provides the statistical power to analyze subtle or variable phenotypes.

We were initially surprised that zebrafish $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$ mutants were viable, as this contradicted previous findings in mouse models [30, 34]. We can imagine a few possible explanations for this finding. First, one or more of the zebrafish

mbnl CRISPR mutations may not have resulted in a complete loss of function. In particular, $2^{A/A}$ mutants had the potential to produce a partially functional protein containing one of the two sets of zinc finger motifs (Fig 1B), although we are confident that $2^{A/A}$ mutants represented at least a partial loss of function given their splicing phenotypes and reduced *mbnl2* mRNA levels (Figs 4 and 5 and S1 and S7 Figs). In contrast to mice, in which *Mbnl2* expression is enriched in the brain[31], all three zebrafish *mbnl* genes are expressed broadly (S6 Fig), and perhaps the presence of partially functional mbnl2 protein across tissues including skeletal muscle was sufficient for zebrafish survival. Larvae injected with *mbnl2* morpholinos were reported to have dramatic morphological phenotypes [37], although we were unable to recapitulate these results using the same morpholinos. The fact that our $2^{A/A}$ fish did not exhibit similar phenotypes could be explained by partially retained Mbnl2 function in $2^{A/A}$ fish. We suspect, however, that off target effects of the *mbnl2* morpholinos contributed to the observed morphological defects, given the mild phenotypes of $Mbnl2^{-/-}$ mice[31]. Another explanation for the discrepancy between *mbnl2* morphant and mutant phenotypes could be genetic compensation, although we did not observe a substantial increase in expression of other *mbnl* family members in *mbnl2* mutants (S1 and S8 Figs)[43]. Even if the *mbnl* mutant fish retained some protein function, they were still valid models of DM, as sequestration of MBNL proteins by CUG/CCUG repeats is likely not complete in the human disease state.

An additional possibility is that the zebrafish *mbnl* mutations did result in complete loss of function, but that Mbnl function is not required for viability in fish. Most muscle-specific double and triple *Mbnl* mutant mice died during the neonatal period

when they transitioned abruptly between receiving oxygen through the placenta and breathing on their own. The small fraction of muscle-specific compound *Mbnl* mutant mice that were able to make it through this transition survived to adulthood [26]. Perhaps compound *mbnl* mutant fish were better able to overcome this developmental hurdle because larval fish undergo a much more gradual transition between receiving oxygen through diffusion to receiving oxygen through the gills.

Zebrafish *mbnl* mutants had physical phenotypes that were relevant to human DM disease and mouse DM models. Adult $1^{B/B};2^{A/A}$, $1^{B/B};3^{C/C}$, and $1^{B/B};2^{A/A};3^{C/C}$ fish were significantly smaller than their WT tank mates (Fig 2A). Neonatal muscle-specific compound mutant mice likewise were substantially smaller than their littermates [34]. In the case of fish, the size phenotype did not arise early in development, as 7 dpf larval *mbnl* mutant fish were not smaller than WT fish (Fig 2B,C and S2D Fig). We do not know the mechanism underlying the size phenotype in *mbnl* mutant fish, but one possibility is impaired movement leading to decreased ability to access food. Consistent with this idea, we found that $1^{B/B};2^{A/A}$, $1^{B/B};3^{C/C}$, and $1^{B/B};2^{A/A};3^{C/C}$ fish swam a significantly decreased distance compared to WT fish during a five-minute swim test (Fig 3D,E). Interestingly, human DM patients exhibited decreased performance on a similar 6-minute walk test due to muscle weakness and impaired motor function, leading us to hypothesize that *mbnl* mutant fish also had muscle weakness [38, 39]. Although basic histological analysis did not indicate obvious signs of muscle fiber pathology in *mbnl* mutant fish (S3 Fig), more in-depth analyses are needed to determine whether subtle changes in muscle structure or function are present that could lead to muscle weakness. We also found that almost all genotypes of *mbnl* mutant fish stayed closer

to the bottom of the tank on average than WT fish (Fig 3F,G,H), which could also decrease access to food as fish were fed from the top of the tank. The *mbnl* mutant fish could have stayed at the bottom of the tank due to motor dysfunction or a morphological defect, but this phenotype is also a well-characterized indicator of anxiety in fish [40]. Individuals with DM often show signs of avoidant personality disorder with anxiety, and an intriguing possibility is that *mbnl* mutant fish also have increased anxiety [44, 45]. Our future studies will explore in depth the mechanisms underlying decreased size and altered swimming behavior in *mbnl* mutant fish.

Like humans with DM and mouse models, zebrafish *mbnl* mutants exhibited widespread changes in alternative splicing (Figs 4-6 and S7 and S9 Figs). Splicing changes in zebrafish *mbnl* mutants occurred in both larvae and adults, and across many DM-relevant tissues including skeletal muscle, heart, brain, and intestine (Fig 4 and S7 Fig). Mis-splicing was also observed in the cornea, suggesting that zebrafish *mbnl* mutants may also model the genetic disorder Fuchs endothelial corneal dystrophy, a subtype of which was recently shown to be caused by an expanded CUG repeat that is expressed from an intron of the *TCF4* gene and is associated with MBNL protein sequestration and mis-splicing (Fig 4 and S7 Fig)[46-49]. Hundreds of alternative splicing changes were observed when each of the zebrafish *mbnl* genes was mutated alone, suggesting that each plays a role in splicing regulation (Fig 5A). Consistent with mouse studies was the finding that splicing changes were more numerous and larger in magnitude in double and triple homozygous zebrafish mutants than in single homozygous mutants, indicating partially redundant functions of Mbnl proteins (Fig 5) [30, 34]. A notable exception were the $2^{A/A};3^{C/C}$ mutants, which exhibited fewer overall

dysregulated alternative splicing events and lacked the swimming position phenotype of *2*$^{A/A}$ mutants and *3*$^{C/C}$ mutants (Fig 3F,G and Fig 5A). In mouse models, many cassette exon events were shown to be dysregulated in opposite directions in *Mbnl3* mutants and other *Mbnl* mutants[34]. Perhaps opposing splicing regulation could be a contributing factor to the mild phenotypes of *2*$^{A/A}$*;3*$^{C/C}$ mutants.

Our data indicated that, just like other model organisms, zebrafish Mbnl-mediated alternative splicing regulation was quite complex. Mbnl proteins suppressed inclusion of some exons, while they promoted inclusion of others (Fig 6 and S9 Fig). As observed in other organisms, YGCY putative Mbnl binding motifs were enriched upstream of exons that were suppressed by zebrafish Mbnl proteins, and enriched downstream of exons whose inclusion was promoted (Fig 5D) [1, 3, 8, 10]. Also similar to observations in other systems, some zebrafish exons were more sensitive to changes in overall *mbnl* concentration than others [41]. For example, in skeletal muscle, *mbnl2* exon 5 required the loss of both *mbnl1* and *mbnl2* function for increased inclusion, whereas *mbnl1* exon 5 only required the loss of *mbnl1* function (Fig 4B and S7B Fig).

Many of the alternative splicing changes observed in *mbnl* mutant DM model fish were conserved with changes observed in DM1 patient tissues (Fig 6 and S9 Fig). These included several that are known to contribute to disease phenotypes, such as changes in *ATP2A1* and *RYR1* splicing that contribute to dysregulation of calcium homeostasis[27, 28]. However, not all missplicing events were conserved between species. For example, *mbnl* mutants did not exhibit changes in alternative splicing of the zebrafish orthologs of *CLCN1*, whose splicing misregulation contributes to myotonia in DM, so the zebrafish is unlikely to be an appropriate model for studying myotonia,

although it can be used to study many other important disease phenotypes [50, 51]. The complete panel of zebrafish *mbnl* mutants and accompanying RNA-Seq data that we generated will be valuable in future studies to determine the specific alternative splicing changes that underlie phenotypes. For example, exons that were misregulated in all zebrafish *mbnl* mutants except for $2^{A/A};3^{C/C}$ mutants are strong candidates for contributing to the anxiety-like phenotype (Fig 3F,G,H). Likewise, exons whose mis-regulation was not conserved between humans with DM and zebrafish are candidates for contributing to disease-related changes in muscle histology that were not observed in the zebrafish *mbnl* mutants (S3 Fig). In summary, we successfully generated a complete panel of zebrafish single, double, and triple homozygous *mbnl* mutants for modeling myotonic dystrophy. These fish exhibited molecular and physical phenotypes similar to those observed in humans with DM. We propose that zebrafish $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{F/F}$ mutants, which exhibited dramatic alternative splicing, body size, and movement phenotypes, represent severe forms of DM1 with long CTG repeats, while other double and single *mbnl* mutants model more moderate forms of the disease. The new zebrafish models will be highly useful in future studies of the functions of individual Mbnl proteins and how they contribute to disease relevant phenotypes. Strengths of using zebrafish over other vertebrate model systems are that large numbers of fish can be generated to study subtle or variable phenotypes, and that transparent larvae can be used to directly study disease-related phenotypes, such as altered gut motility and heart abnormalities, in live animals.

Materials and methods

Generation of mutant zebrafish

All zebrafish experiments were performed with the guidance and approval of the University of Oregon Institutional Animal Care and Use Committee (PHS assurance number D16-00004, protocol AUP-15-98). Guide RNAs (gRNAs) targeting zebrafish (*Danio rerio*) *mbnl1*, *mbnl2*, and *mbnl3* were designed using the Chop Chop website (http://chopchop.cbu.uib.no). DNA templates for the gRNAs were generated by a template-free Phusion polymerase (New England Biolabs) PCR reaction using a common scaffold primer (gRNA scaffold, S1 Table) and a gene-specific primer (*mbnl1* gRNA1, *mbnl1* gRNA2, *mbnl2* gRNA, *mbnl3* gRNA1, or *mbnl3* gRNA2, S1 Table), then cleaned using the QIAquick PCR Purification Kit (Qiagen). gRNAs were transcribed from DNA templates using a MEGAscript kit (Ambion) and purified by phenol-chloroform extraction and isopropanol precipitation. Cas9 RNA was generated by linearizing the pT3TS-nls-zCas9-nls plasmid [52] with XbaI, purifying it using the QIAquick Gel Extraction Kit (Qiagen), performing an in vitro transcription reaction using the T3 mMESSAGE kit (Invitrogen), and purifying the RNA using the RNeasy Mini kit (Qiagen). AB zebrafish embryos were microinjected at the one cell stage with 1-2 nL of a mixture containing 100ng/µL Cas9 mRNA, 50 ng/µL gRNA, and phenol red, and raised to adulthood. Mosaic mutants were identified by PCR amplification and Sanger sequencing of fin clip DNA using primers specific to the targeted region (S1 and S2 Tables), and outcrossed to wildtype AB zebrafish to generate heterozygotes. Fish with predicted loss of function mutations were identified by Sanger sequencing (S2 Table), and further crossed to generate single, double, and triple *mbnl1*, *mbnl2*, and *mbnl3* homozygotes. Fish were genotyped by restriction fragment length polymorphism analysis using the primers and restriction enzymes indicated in Table S2. Sperm from

all zebrafish mutant lines were cryopreserved and are available upon request from the corresponding author.

Measurement of zebrafish size

Young adult fish (2-4 months old) (Fig 2A and S2A,B,C Fig) were anesthetized in 168 mg/L tricaine methane sulfonate, photographed against a white background with a ruler using a tablet computer, finclipped, and genotyped. To ensure identical density and feeding conditions, fish were compared with others from the same clutch (Fig S2A,B,C) or different clutches that had been raised together in the same tank (Fig 2A). For measurement of larval fish, unfed embryos of different genotypes were raised to 7 days post fertilization (dpf) in separate dishes at a density of one fish per mL of embryo medium, with the exception of the larvae in Fig 2C, which were grown in the same dish and genotyped after measurement. Fish were anesthetized in 168 mg/mL tricaine methane sulfonate, laid out on a microscope slide in 3% methylcellulose, and imaged using a Leica M165FC microscope. An investigator that was blinded to genotype measured the distance from snout to caudal peduncle or posterior tip of the notochord (standard length, SL) of each fish [53]. Data were analyzed by ordinary one-way ANOVA with multiple comparison correction (S5 Table).

Alternative splicing analysis by RT-PCR

Fish were euthanized by tricaine methane sulfonate overdose (larvae) or hypothermic shock (adults). Tissues of interest were dissected and flash frozen in 1 mL of Trizol (Ambion), thawed, and homogenized with a mortar and pestle (larvae) or Bullet Blender Storm 24 (adult tissues). Chloroform (200 µL) was added to each tube followed by mixing, centrifugation at 12,000 g for 10 minutes at 4°C, transfer of the aqueous

phase to a separate tube, addition of 200 µL ethanol, and binding of sample to an RNeasy mini kit column (Qiagen). RNA was washed and eluted according to the manufacturer's instructions and the concentration was measured using a NanoDrop 2000 (Thermo Scientific).

RNA (20-200 ng) was reverse transcribed with Superscript II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions using gene-specific reverse (R) primers located in the exon downstream of the regulated exon of interest (S1 Table). The resulting cDNA was amplified by 28-31 cycles of PCR using Taq polymerase and the forward (F) and reverse (R) primers indicated in S1 Table. Samples were separated by electrophoresis on a 6% *bis*-Acrylamide (19:1) gel that was stained overnight with 1X SYBR green I nucleic acid gel stain (Invitrogen). The gel was imaged and quantified using an AlphaImagerHP (Alpha Innotech). The background-corrected sum of each band was measured and the percent exon inclusion was calculated using the following formula: ((exon included sum) /(exon included sum + exon excluded sum)) X 100. Data were analyzed by ordinary one-way ANOVA with multiple comparison correction.

RNA-Seq analysis

RNA was extracted using the RiboPure RNA Purification Kit (Invitrogen AM1924) from epaxial skeletal muscle from the tails of three biological replicates of adult WT and mutant fish. RNA quality was examined using the Fragment Analyzer RNA Analysis DNF-471 kit (Advanced Analytical) and all RQN values were > 8.0. Ribosomal RNA was depleted from 200 µg of the RNA using the NEBNext rRNA Depletion Kit (NEB E6310X) and then a cDNA library was prepared using the NEBNext Ultra RNA Library Prep Kit

for Illumina (NEB E7530L). The libraries were checked for quality using the Fragment Analyzer NGS Analysis DNF-474 kit (Advanced Analytical) and quantitated using the KAPA Library Quantification Kit (KAPA Code KK4824). The completed libraries were then pooled in equimolar amounts and sequenced on an Illumina Next-Seq 500. A minimum of 60 million paired-end 75 x 75 reads were obtained for each library.

BCL files were demultiplexed and converted to fastq files using BCL2Fastq (version 2.16.0.10). The fastq files were then checked for quality using FastQC (version 0.11.8) and aligned to the GRCz11 zebrafish genome using STAR (version 2.5.1b). The alternative splicing events were then analyzed using rMATS (version 4.0.2), grouping the biological replicates of each mutant and comparing them against WT. A similar preliminary analysis of the same sequencing data was performed using GRCz10 and rMATS (version 4.0.2). Significant mis-splicing events were categorized as having a false discovery rate (FDR) < 0.10. Cluster analysis was performed with R and visualized with the factoextra package (version 1.0.5) using the top 800 most variable cassette exon splicing events in PSI values. Data were first tested using the gap statistic and then K-Means clustering was performed with two centers. The data for human DM1 tibialis anterior comparisons were downloaded from NCBI GEO, accession # GSE86356. Data for the DM1-derived myoblast cell line was downloaded from NCBI SRA, study # SRP158284. The data were processed as described above and then compared to the zebrafish data for orthologous mis-splicing events. Orthologous exons were found using a custom python script and using tblastx to confirm. Exons were counted as orthologous if the genes from which they were transcribed were at least 75% conserved between species, the exon was in the same place in the transcript

(equal exon number in at least one annotated transcript), and if the exon had an evalue from tblastx of lower than 0.05.

Enrichment of putative YGCY Mbnl protein binding motifs was analyzed for the sets of cassette exons whose inclusion was significantly increased (FDR < 0.05, ΔPSI > 0.2) or decreased (FDR < 0.05, ΔPSI < −0.2) compared to WT in $1^{B/B};2^{A/A}$ mutants, $1^{B/B};2^{A/A};3^{C/C}$ mutants, or in both ($1^{B/B};2^{A/A} \cap 1^{B/B};2^{A/A};3^{C/C}$). For each set of cassette exons, the frequency of YGCY motifs (GCTT, CGCT, TGCT, and GCGC) was compared to the frequency of control 4-mers with identical A+T and CpG content within the following regions: cassette exons (normalized for analyzed sequence length), intronic sequences 1-250 bp immediately 5' or 3' of the cassette exons, intronic sequences 1-250 bp 3' of the upstream constitutive exons, and intronic sequences 1-250 bp 5' of the downstream constitutive exons. The $\log_2$ ((enrichment of YGCY motifs relative to control motifs in each region associated with regulated cassette exons)/(enrichment of YGCY motifs relative to control motifs in each region associated with non-regulated exons)) was plotted in heatmap form.

Analysis of *mbnl* gene expression

RT-qPCR was used to measure mbnl RNA levels in zebrafish tissues. Total RNA was prepared from 5 dpf larvae and adult tissues using the same procedure as for alternative splicing analysis. RNA was treated with TURBO DNase (Ambion) according to manufacturer's instructions then reverse transcribed with an oligo$(dT)_{20}$ primer using the Superscript III cDNA First Strand Synthesis kit (Invitrogen). The qPCR reaction was set up using the KAPA SYBR FAST ABI Prism kit (KAPA Biosystems) according to manufacturer's instructions and run on a Quant Studio 3 System (ThermoFisher) using

the default settings for SYBR Green reagents and the fast run mode in the QuantStudio Design & Analysis Software v.1.4.2. The comparative C_T ($\Delta\Delta C_T$) method was used to calculate relative mRNA levels. Five biological samples of each genotype or tissue type were run in triplicate and data were normalized to the expression of the housekeeping gene *eef1a1l1*[54]. Primers used for RT-qPCR are shown in S1 Table.

Western blots analyses were performed using 15-50 μg of protein lysates that were prepared in RIPA buffer (Boston Bioproducts) with complete mini EDTA-free protease inhibitor (Roche) from HeLa cells (positive control for MBNL1 and MBNL2 protein expression), whole 5 dpf zebrafish larvae, adult skeletal muscle, heart, brain, intestine, and cornea, or purchased from Santa Cruz Biotechnology (human placenta extract used as a positive control for MBNL3 protein expression). The antibodies tested included Santa Cruz anti-MBNL1(D-4) sc-515374 (1:250), Santa Cruz anti-MBNL1 (4A8) sc-136165 (1:500), Millipore anti-MBNL1 ABE-241 (1:1000), Abcam anti-MBNL2 ab105331 (1:250), Sigma-Aldrich anti-MBNL3 SAB1411751 (1:150), and Sigma-Aldrich anti-actin A5060 (1:1000). Santa Cruz goat anti-rabbit IgG-HRP sc-2004 (1:2000) and Sigma goat anti-mouse IgG FC specific HRP (1:2000) were used as secondary antibodies.

Zebrafish behavior assays

Age-matched single adult zebrafish were placed into a novel environment consisting of a custom glass aquarium measuring 18 cm deep X 25.4 cm long X 7.6 cm wide (Fig 3A). The fish were monitored with a Logitech camera during a 5-minute time frame to characterize basic exploratory swimming behavior. The raw tracking data was analyzed with custom software (DaniOPEN, https://github.com/stednitzs/daniopen) to

measure swim distance and distance from the bottom of the tank[55]. Following the test period, the fish were returned to their home tanks.

Histology

Ten-month-old WT and mutant adult zebrafish were euthanized by hypothermic shock and fixed in Bouin's solution (Sigma), then washed in 70% ethanol. Fixed fish were bisected parasagitally or transversely, processed for paraffin embedding, sectioned at 7 µm thickness, and stained with hematoxylin and eosin by a histologist from the University of Oregon Institute of Neuroscience. Muscle morphology was analyzed by a zebrafish pathologist from the Zebrafish International Resource Center who was blinded to genotype. Representative DIC images of muscle transverse sections were taken using a Leica DMLB microscope.

Acknowledgements

We thank the staff of the University of Oregon Zebrafish Facility for maintaining fish and providing expertise in performing zebrafish procedures, and Dr. Poh Kheng Loi for histological services. We thank Ellie Melancon, Camila Morales Fenero, and Michelle Massaquoi for advice in performing experiments and data analysis, and members of the labs of Karen Guillemin, Andy Berglund, and Judith Eisen for helpful discussions.

References

1. Oddo JC, Saxena T, McConnell OL, Berglund JA, Wang ET. Conservation of context-dependent splicing activity in distant Muscleblind homologs. Nucleic Acids Res. 2016;44(17):8352-62. doi: 10.1093/nar/gkw735. PubMed PMID: 27557707; PubMed Central PMCID: PMCPMC5041496.

2. Warf MB, Berglund JA. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. RNA. 2007;13(12):2238-51. doi: 10.1261/rna.610607. PubMed PMID: 17942744; PubMed Central PMCID: PMCPMC2080590.

3. Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell. 2012;150(4):710-24. doi: 10.1016/j.cell.2012.06.041. PubMed PMID: 22901804; PubMed Central PMCID: PMCPMC3428802.

4. Ho TH, Charlet BN, Poulos MG, Singh G, Swanson MS, Cooper TA. Muscleblind proteins regulate alternative splicing. EMBO J. 2004;23(15):3103-12. doi: 10.1038/sj.emboj.7600300. PubMed PMID: 15257297; PubMed Central PMCID: PMCPMC514918.

5. Rau F, Freyermuth F, Fugier C, Villemin JP, Fischer MC, Jost B, et al. Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. Nat Struct Mol Biol. 2011;18(7):840-5. doi: 10.1038/nsmb.2067. PubMed PMID: 21685920.

6. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, et al. circRNA biogenesis competes with pre-mRNA splicing. Mol Cell. 2014;56(1):55-66. doi: 10.1016/j.molcel.2014.08.019. PubMed PMID: 25242144.

7. Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, et al. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. Mol Cell. 2014;56(2):311-22. doi:

10.1016/j.molcel.2014.08.027. PubMed PMID: 25263597; PubMed Central PMCID: PMCPMC4224598.

8. Goers ES, Purcell J, Voelker RB, Gates DP, Berglund JA. MBNL1 binds GC motifs embedded in pyrimidines to regulate alternative splicing. Nucleic Acids Res. 2010;38(7):2467-84. doi: 10.1093/nar/gkp1209. PubMed PMID: 20071745; PubMed Central PMCID: PMCPMC2853123.

9. Grammatikakis I, Goo YH, Echeverria GV, Cooper TA. Identification of MBNL1 and MBNL3 domains required for splicing activation and repression. Nucleic Acids Res. 2011;39(7):2769-80. doi: 10.1093/nar/gkq1155. PubMed PMID: 21109529; PubMed Central PMCID: PMCPMC3074124.

10. Du H, Cline MS, Osborne RJ, Tuttle DL, Clark TA, Donohue JP, et al. Aberrant alternative splicing and extracellular matrix gene expression in mouse models of myotonic dystrophy. Nat Struct Mol Biol. 2010;17(2):187-93. doi: 10.1038/nsmb.1720. PubMed PMID: 20098426; PubMed Central PMCID: PMCPMC2852634.

11. Gates DP, Coonrod LA, Berglund JA. Autoregulated splicing of muscleblind-like 1 (MBNL1) Pre-mRNA. J Biol Chem. 2011;286(39):34224-33. doi: 10.1074/jbc.M111.236547. PubMed PMID: 21832083; PubMed Central PMCID: PMCPMC3190765.

12. Terenzi F, Ladd AN. Conserved developmental alternative splicing of muscleblind-like (MBNL) transcripts regulates MBNL localization and activity. RNA Biol. 2010;7(1):43-55. PubMed PMID: 20009516.

13. Tran H, Gourrier N, Lemercier-Neuillet C, Dhaenens CM, Vautrin A, Fernandez-Gomez FJ, et al. Analysis of exonic regions involved in nuclear localization, splicing

activity, and dimerization of Muscleblind-like-1 isoforms. J Biol Chem. 2011;286(18):16435-46. doi: 10.1074/jbc.M110.194928. PubMed PMID: 21454535; PubMed Central PMCID: PMCPMC3091249.

14. Brook JD, McCurrach ME, Harley HG, Buckler AJ, Church D, Aburatani H, et al. 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. 1992;68(4):799-808. PubMed PMID: 1310900.

15. Fu YH, Pizzuti A, Fenwick RG, Jr., King J, Rajnarayan S, Dunne PW, et al. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science. 1992;255(5049):1256-8. PubMed PMID: 1546326.

16. Mahadevan M, Tsilfidis C, Sabourin L, Shutler G, Amemiya C, Jansen G, et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science. 1992;255(5049):1253-5. PubMed PMID: 1546325.

17. Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, et al. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science. 2001;293(5531):864-7. doi: 10.1126/science.1062125. PubMed PMID: 11486088.

18. Fardaei M, Rogers MT, Thorpe HM, Larkin K, Hamshere MG, Harper PS, et al. 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. 2002;11(7):805-14. PubMed PMID: 11929853.

19. Miller JW, Urbinati CR, Teng-Umnuay P, Stenberg MG, Byrne BJ, Thornton CA, et al. Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. EMBO J. 2000;19(17):4439-48. doi:

10.1093/emboj/19.17.4439. PubMed PMID: 10970838; PubMed Central PMCID: PMCPMC302046.

20. Yum K, Wang ET, Kalsotra A. Myotonic dystrophy: disease repeat range, penetrance, age of onset, and relationship between repeat size and phenotypes. Curr Opin Genet Dev. 2017;44:30-7. doi: 10.1016/j.gde.2017.01.007. PubMed PMID: 28213156; PubMed Central PMCID: PMCPMC5447468.

21. Wenninger S, Montagnese F, Schoser B. Core Clinical Phenotypes in Myotonic Dystrophies. Front Neurol. 2018;9:303. doi: 10.3389/fneur.2018.00303. PubMed PMID: 29770119; PubMed Central PMCID: PMCPMC5941986.

22. Hilbert JE, Barohn RJ, Clemens PR, Luebbe EA, Martens WB, McDermott MP, et al. High frequency of gastrointestinal manifestations in myotonic dystrophy type 1 and type 2. Neurology. 2017;89(13):1348-54. doi: 10.1212/WNL.0000000000004420. PubMed PMID: 28855409; PubMed Central PMCID: PMCPMC5649763.

23. Bellini M, Biagi S, Stasi C, Costa F, Mumolo MG, Ricchiuti A, et al. Gastrointestinal manifestations in myotonic muscular dystrophy. World J Gastroenterol. 2006;12(12):1821-8. PubMed PMID: 16609987; PubMed Central PMCID: PMCPMC4087506.

24. Tarnopolsky MA, Pearce E, Matteliano A, James C, Armstrong D. Bacterial overgrowth syndrome in myotonic muscular dystrophy is potentially treatable. Muscle Nerve. 2010;42(6):853-5. doi: 10.1002/mus.21787. PubMed PMID: 21104859.

25. Tieleman AA, van Vliet J, Jansen JB, van der Kooi AJ, Borm GF, van Engelen BG. Gastrointestinal involvement is frequent in Myotonic Dystrophy type 2.

Neuromuscul Disord. 2008;18(8):646-9. doi: 10.1016/j.nmd.2008.05.010. PubMed PMID: 18602828.

26. Thomas JD, Oliveira R, Sznajder LJ, Swanson MS. Myotonic Dystrophy and Developmental Regulation of RNA Processing. Compr Physiol. 2018;8(2):509-53. doi: 10.1002/cphy.c170002. PubMed PMID: 29687899.

27. Kimura T, Nakamori M, Lueck JD, Pouliquin P, Aoike F, Fujimura H, et al. Altered mRNA splicing of the skeletal muscle ryanodine receptor and sarcoplasmic/endoplasmic reticulum Ca2+-ATPase in myotonic dystrophy type 1. Hum Mol Genet. 2005;14(15):2189-200. doi: 10.1093/hmg/ddi223. PubMed PMID: 15972723.

28. Zhao Y, Ogawa H, Yonekura S, Mitsuhashi H, Mitsuhashi S, Nishino I, et al. Functional analysis of SERCA1b, a highly expressed SERCA1 variant in myotonic dystrophy type 1 muscle. Biochim Biophys Acta. 2015;1852(10 Pt A):2042-7. doi: 10.1016/j.bbadis.2015.07.006. PubMed PMID: 26170059.

29. Kanadia RN, Johnstone KA, Mankodi A, Lungu C, Thornton CA, Esson D, et al. A muscleblind knockout model for myotonic dystrophy. Science. 2003;302(5652):1978-80. doi: 10.1126/science.1088583. PubMed PMID: 14671308.

30. Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, et al. Compound loss of muscleblind-like function in myotonic dystrophy. EMBO Mol Med. 2013;5(12):1887-900. doi: 10.1002/emmm.201303275. PubMed PMID: 24293317; PubMed Central PMCID: PMCPMC3914532.

31. Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, Yuan Y, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in

myotonic dystrophy. Neuron. 2012;75(3):437-50. doi: 10.1016/j.neuron.2012.05.029. PubMed PMID: 22884328; PubMed Central PMCID: PMCPMC3418517.

32. Poulos MG, Batra R, Li M, Yuan Y, Zhang C, Darnell RB, et al. Progressive impairment of muscle regeneration in muscleblind-like 3 isoform knockout mice. Hum Mol Genet. 2013;22(17):3547-58. doi: 10.1093/hmg/ddt209. PubMed PMID: 23660517; PubMed Central PMCID: PMCPMC3736872.

33. Choi J, Dixon DM, Dansithong W, Abdallah WF, Roos KP, Jordan MC, et al. Muscleblind-like 3 deficit results in a spectrum of age-associated pathologies observed in myotonic dystrophy. Sci Rep. 2016;6:30999. doi: 10.1038/srep30999. PubMed PMID: 27484195; PubMed Central PMCID: PMCPMC4971533.

34. Thomas JD, Sznajder LJ, Bardhi O, Aslam FN, Anastasiadis ZP, Scotti MM, et al. Disrupted prenatal RNA processing and myogenesis in congenital myotonic dystrophy. Genes Dev. 2017;31(11):1122-33. doi: 10.1101/gad.300590.117. PubMed PMID: 28698297; PubMed Central PMCID: PMCPMC5538435.

35. Liu YF, Liu HY, Tu LC, Lin CW, Hsiao KM, Pan H. Zebrafish muscleblind-like genes: identification, structural features and expression. Comp Biochem Physiol B Biochem Mol Biol. 2008;151(1):118-24. doi: 10.1016/j.cbpb.2008.06.005. PubMed PMID: 18606240.

36. Todd PK, Ackall FY, Hur J, Sharma K, Paulson HL, Dowling JJ. Transcriptional changes and developmental abnormalities in a zebrafish model of myotonic dystrophy type 1. Dis Model Mech. 2014;7(1):143-55. doi: 10.1242/dmm.012427. PubMed PMID: 24092878; PubMed Central PMCID: PMCPMC3882056.

37. Machuca-Tzili LE, Buxton S, Thorpe A, Timson CM, Wigmore P, Luther PK, et al. Zebrafish deficient for Muscleblind-like 2 exhibit features of myotonic dystrophy. Dis Model Mech. 2011;4(3):381-92. doi: 10.1242/dmm.004150. PubMed PMID: 21303839; PubMed Central PMCID: PMCPMC3097459.

38. Kierkegaard M, Tollback A. Reliability and feasibility of the six minute walk test in subjects with myotonic dystrophy. Neuromuscul Disord. 2007;17(11-12):943-9. doi: 10.1016/j.nmd.2007.08.003. PubMed PMID: 17869516.

39. Park D, Lee SH, Shin JH, Park JS. Lower limb muscle magnetic resonance imaging in myotonic dystrophy type 1 correlates with the six-minute walk test and CTG repeats. Neuromuscul Disord. 2018;28(1):29-37. doi: 10.1016/j.nmd.2017.08.005. PubMed PMID: 29066035.

40. Cachat J, Stewart A, Grossman L, Gaikwad S, Kadri F, Chung KM, et al. Measuring behavioral and endocrine responses to novelty stress in adult zebrafish. Nat Protoc. 2010;5(11):1786-99. doi: 10.1038/nprot.2010.140. PubMed PMID: 21030954.

41. Wagner SD, Struck AJ, Gupta R, Farnsworth DR, Mahady AE, Eichinger K, et al. Dose-Dependent Regulation of Alternative Splicing by MBNL Proteins Reveals Biomarkers for Myotonic Dystrophy. PLoS Genet. 2016;12(9):e1006316. doi: 10.1371/journal.pgen.1006316. PubMed PMID: 27681373; PubMed Central PMCID: PMCPMC5082313.

42. Freyermuth F, Rau F, Kokunai Y, Linke T, Sellier C, Nakamori M, et al. Splicing misregulation of SCN5A contributes to cardiac-conduction delay and heart arrhythmia in myotonic dystrophy. Nat Commun. 2016;7:11067. doi: 10.1038/ncomms11067. PubMed PMID: 27063795; PubMed Central PMCID: PMCPMC4831019.

43. El-Brolosy MA, Kontarakis Z, Rossi A, Kuenne C, Gunther S, Fukuda N, et al. Genetic compensation triggered by mutant mRNA degradation. Nature. 2019;568(7751):193-7. doi: 10.1038/s41586-019-1064-z. PubMed PMID: 30944477; PubMed Central PMCID: PMCPMC6707827.

44. Minier L, Lignier B, Bouvet C, Gallais B, Camart N. A Review of Psychopathology Features, Personality, and Coping in Myotonic Dystrophy Type 1. J Neuromuscul Dis. 2018;5(3):279-94. doi: 10.3233/JND-180310. PubMed PMID: 30040740; PubMed Central PMCID: PMCPMC6087440.

45. Delaporte C. Personality patterns in patients with myotonic dystrophy. Arch Neurol. 1998;55(5):635-40. PubMed PMID: 9605719.

46. Mootha VV, Hussain I, Cunnusamy K, Graham E, Gong X, Neelam S, et al. TCF4 Triplet Repeat Expansion and Nuclear RNA Foci in Fuchs' Endothelial Corneal Dystrophy. Invest Ophthalmol Vis Sci. 2015;56(3):2003-11. doi: 10.1167/iovs.14-16222. PubMed PMID: 25722209; PubMed Central PMCID: PMCPMC4373545.

47. Wieben ED, Aleff RA, Tang X, Butz ML, Kalari KR, Highsmith EW, et al. Trinucleotide Repeat Expansion in the Transcription Factor 4 (TCF4) Gene Leads to Widespread mRNA Splicing Changes in Fuchs' Endothelial Corneal Dystrophy. Invest Ophthalmol Vis Sci. 2017;58(1):343-52. doi: 10.1167/iovs.16-20900. PubMed PMID: 28118661; PubMed Central PMCID: PMCPMC5270622.

48. Winkler NS, Milone M, Martinez-Thompson JM, Raja H, Aleff RA, Patel SV, et al. Fuchs' Endothelial Corneal Dystrophy in Patients With Myotonic Dystrophy, Type 1. Invest Ophthalmol Vis Sci. 2018;59(7):3053-7. doi: 10.1167/iovs.17-23160. PubMed PMID: 30025114; PubMed Central PMCID: PMCPMC6005624.

49. Mootha VV, Hansen B, Rong Z, Mammen PP, Zhou Z, Xing C, et al. Fuchs' Endothelial Corneal Dystrophy and RNA Foci in Patients With Myotonic Dystrophy. Invest Ophthalmol Vis Sci. 2017;58(11):4579-85. doi: 10.1167/iovs.17-22350. PubMed PMID: 28886202; PubMed Central PMCID: PMCPMC5590687.

50. Lueck JD, Lungu C, Mankodi A, Osborne RJ, Welle SL, Dirksen RT, et al. Chloride channelopathy in myotonic dystrophy resulting from loss of posttranscriptional regulation for CLCN1. Am J Physiol Cell Physiol. 2007;292(4):C1291-7. doi: 10.1152/ajpcell.00336.2006. PubMed PMID: 17135300.

51. Lueck JD, Mankodi A, Swanson MS, Thornton CA, Dirksen RT. Muscle chloride channel dysfunction in two mouse models of myotonic dystrophy. J Gen Physiol. 2007;129(1):79-94. doi: 10.1085/jgp.200609635. PubMed PMID: 17158949; PubMed Central PMCID: PMCPMC2151606.

52. Jao LE, Wente SR, Chen W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A. 2013;110(34):13904-9. doi: 10.1073/pnas.1308335110. PubMed PMID: 23918387; PubMed Central PMCID: PMCPMC3752207.

53. Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. Dev Dyn. 2009;238(12):2975-3015. doi: 10.1002/dvdy.22113. PubMed PMID: 19891001; PubMed Central PMCID: PMCPMC3030279.

54. Vanhauwaert S, Van Peer G, Rihani A, Janssens E, Rondou P, Lefever S, et al. Expressed repeat elements improve RT-qPCR normalization across a wide range of zebrafish gene expression studies. PLoS One. 2014;9(10):e109091. doi:

10.1371/journal.pone.0109091. PubMed PMID: 25310091; PubMed Central PMCID: PMCPMC4195698.

55. Stednitz SJ, McDermott EM, Ncube D, Tallafuss A, Eisen JS, Washbourne P. Forebrain Control of Behaviorally Driven Social Orienting in Zebrafish. Curr Biol. 2018;28(15):2445-51 e3. doi: 10.1016/j.cub.2018.06.016. PubMed PMID: 30057306; PubMed Central PMCID: PMCPMC6082377.

Figure legends

Fig 1. Zebrafish DM models were generated by mutating *mbnl1, mbnl2,* and *mbnl3.*

(A) Primary molecular mechanism of the human genetic disorder myotonic dystrophy (DM). In unaffected individuals, free MBNL1, MBNL2, and MBNL3 proteins bind to target pre-mRNAs and regulate the inclusion of alternative exons in mRNAs. MBNL proteins suppress alternative exon inclusion in the example shown here, but they promote inclusion of other target alternative exons. In individuals with DM, MBNL proteins are sequestered by long CUG (DM1) or CCUG (DM2) repeat RNAs, which decreases their availability to bind to target pre-mRNAs and alters mRNA isoform production and downstream protein function. (B) Diagram of WT and mutant zebrafish *mbnl1*, *mbnl2,* and *mbnl3* predicted coding sequences. Mutant sequences are shown in S2 Table. *mbnl1* mutant alleles are denoted as 1^{A} and 1^{B}, *mbnl2* as 2^{A}, and *mbnl3* as $3^{A/A}$, $3^{B/B}$, and $3^{C/C}$.

Fig 2. Double and triple homozygous *mbnl* mutant zebrafish were decreased in size.

(A) Standard length of young adult 60 days post fertilization (dpf) WT and *mbnl* mutant zebrafish that were raised in the same tank. On the right are representative images of fish of each genotype taken at 76 dpf. (B) & (C) Standard length of 7 dpf WT and double *mbnl* mutant zebrafish (B) or clutchmates from an incross of $1^{+/B};2^{A/A};3^{C/C}$ fish (C). For all graphs, *mbnl1* mutant alleles are denoted as 1^{A} and 1^{B}, *mbnl2* alleles as 2^{A}, and *mbnl3* alleles as 3^{A}, 3^{B}, and 3^{C}. Each dot represents one fish and error bars indicate standard deviation. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

Fig 3. Zebrafish *mbnl* mutants exhibited altered movement.

(A) Example still image taken from video of swim tests. Software was used to track the center of mass of each fish in the X and Y directions over 5 minutes. (B) Representative traces of individual fish of each genotype taken during the swim tests. X and Y axes are the same as in (A). (C) to (E) Total distance that adult WT and (C) single, (D) double, and (E) triple homozygous *mbnl* mutant fish swam during a five-minute swim test. (F) to (H) Average relative distance from the bottom of the tank of fish during a five-minute swim test. Zero represents the bottom of the tank and 100 represents the top of the tank. For all sub-figures, *mbnl1* mutant alleles are denoted as 1^{A} and 1^{B}, *mbnl2* alleles as 2^{A}, and *mbnl3* alleles as 3^{A}, 3^{B}, and 3^{C}. Each dot represents one fish and error bars indicate standard error. Single and double mutant data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test, while triple mutant data were analyzed by a Student's t-test. Data bars that do not share a

letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

Fig 4. Alternative splicing of *mbnl1* exon 5 was misregulated across tissues in zebrafish *mbnl* mutants.

(A) to (F) RT-PCR analysis showing percent spliced in (PSI) of *mbnl1* exon 5 in WT and *mbnl* mutant whole 5 dpf larvae (A), and in adult skeletal muscle (B), heart (C), brain (D), cornea (E), and intestine (F). For all graphs, *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents RNA from one adult fish or a pool of five larval fish. Error bars indicate standard error of the mean. Representative RT-PCR gels are shown above each graph with band sizes in bp on the left. Dividing lines indicate samples run on separate gels. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

Fig 5. Zebrafish *mbnl* mutations led to widespread changes in adult skeletal muscle RNA alternative splicing.

(A) Graph showing the total number of RNA alternative splicing events of different types that were significantly misregulated between WT and *mbnl* mutant adult skeletal muscle, as identified by RNA-Seq. (B) The average absolute value of the change in percent spliced in (ΔPSI) is shown for the set of 62 cassette exons whose inclusion was significantly misregulated in RNA-Seq analysis of the skeletal muscle of at least four of seven *mbnl* mutant fish lines compared to WT. Each dot represents one fish. Error bars indicate standard error of the mean. Data were analyzed by ordinary one-way

ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. (C) K-means cluster analysis based on changes in cassette exon inclusion showing that $1^{B/B};2^{A/A}$ and $1^{B/B};2^{A/A};3^{C/C}$ mutants cluster closely with each other, while other mutants cluster more closely with WT. Each small circle or triangle represents an individual fish, and the large circle and triangle represent the centers of the clusters. (D) Heat maps showing the enrichment of the previously identified Mbnl protein binding sequence YGCY (where Y is a pyrimidine) within significantly misregulated cassette exons, in the intronic sequences 250 bp immediately upstream and downstream of those exons, and in the 250 bp of intronic sequences immediately adjacent to the flanking constitutive exons. Activated cassette exons are those in which Mbnl proteins regulated inclusion positively, and repressed cassette exons are those for which inclusion is decreased by Mbnl proteins. The YGCY enrichment analysis was performed for the set of cassette exons that were significantly misregulated in $1^{B/B};2^{A/A}$ fish, in $1^{B/B};2^{A/A};3^{C/C}$ fish, and for the set of misregulated exons that overlapped between the two ($1^{B/B};2^{A/A} \cap 1^{B/B};2^{A/A};3^{C/C}$). For all graphs, the *mbnl1* mutant allele is denoted as1^{B}, the *mbnl2* allele as 2^{A}, and the *mbnl3* allele as 3^{C}. S5 Table contains the raw data used to generate this figure.

Fig 6. Misregulation of many DM-associated alternative splicing events was conserved in *mbnl* mutants.

(A) Change in percent spliced in values (ΔPSI) between mutant and WT are shown for orthologous exons in zebrafish $1^{B/B};2^{A/A};3^{C/C}$ skeletal muscle and in tibialis muscle from human DM1 patients. ρ is the Spearman's rank correlation coefficient. (B) to (H) RNA Seq ((B), (C), (E), and (G)) and RT-PCR ((D), (F), and (H)) analyses showing percent

spliced in (PSI) of *mbnl1* exon 5 (B), *atp2a1l* exon 23 ((C) and (D)), *ryr1b* exon 72 ((E) and (F)), and *ank3b* exon 36 ((G) and (H)) in WT and *mbnl* mutant adult zebrafish skeletal muscle. There is published evidence of misregulation of the human orthologs of each of these exons in DM patient-derived tissues [27, 28, 42]. For all graphs, *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents RNA from one fish. Error bars indicate standard error of the mean. Representative RT-PCR gels are shown above each graph with band sizes in bp shown on the left. Dividing lines indicate samples run on separate gels. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

Supporting information

S1 Fig. Effects of *mbnl* mutations on larval *mbnl* mRNA levels.

(A) to (C) RT-qPCR showing the levels of (A) *mbnl1,* (B) *mbnl2,* and (C) *mbnl3* mRNAs in WT and *mbnl* mutant 5 days post fertilization whole larval zebrafish. For all graphs, *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents RNA pooled from five larval fish and error bars indicate standard error of the mean. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

S2 Fig. Single *mbnl* mutant zebrafish were not decreased in size.

(A) to (C) Standard length of 83 dpf clutchmates from an incross of *1*$^{+/B}$ fish (A), 83 dpf clutchmates from an incross of *2*$^{+/A}$ fish (B), and 78 dpf clutchmates from an incross of *3*$^{+/C}$ fish (C). (D) Standard length of 7 dpf larval fish. For all graphs, *mbnl1* mutant alleles are denoted as *1*^{A} and *1*^{B}, *mbnl2* alleles as *2*^{A}, and *mbnl3* alleles as *3*^{A}, *3*^{B}, and *3*^{C}. Each dot represents one fish. Error bars indicate standard deviation. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

S3 Fig. Adult zebrafish WT and *mbnl* mutant muscle fibers were structurally similar.

Representative hematoxylin and eosin stained transverse sections of epaxial muscle fibers from the tails of adult zebrafish.

S4 Fig. The introns surrounding zebrafish *mbnl1* exon 5 contain potential Mbnl protein binding sites.

Zebrafish *mbnl1* exon5 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S5 Fig. Zebrafish *mbnl2* exon 5 and its surrounding introns contain potential Mbnl protein binding sites.

Zebrafish *mbnl2* exon5 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the

EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S6 Fig. Zebrafish *mbnl* mRNA levels were highest in adult skeletal muscle and brain.

(A) to (C) RT-qPCR showing relative levels of (A) *mbnl1* mRNA, (B) *mbnl2* mRNA, and (C) *mbnl3* mRNA in WT 5 dpf whole larvae, adult skeletal muscle, heart, brain, cornea, and intestine. Each dot represents one adult fish or a pool of 5 larval fish, and error bars indicate standard error of the mean. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

S7 Fig. Alternative splicing of *mbnl2* exon 5 was misregulated across tissues in zebrafish *mbnl* mutants.

(A) to (F) RT-PCR analysis showing percent spliced in (PSI) of *mbnl2* exon 5 in WT and *mbnl* mutant whole 5 dpf larvae (A), and in adult skeletal muscle (B), heart (C), brain (D), cornea (E), and intestine (F). For all graphs, *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents RNA from one fish or a pool of five larval fish. Error bars indicate standard error of the mean. Representative RT-PCR gels are shown above each graph with band sizes in bp shown on the left. Dividing lines indicated samples run on separate gels. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

S8 Fig. *mbnl* mutations did not affect adult skeletal muscle *mbnl* mRNA levels. (A) to (C) RNA Seq data showing the levels of (A) *mbnl1,* (B) *mbnl2,* and (C) *mbnl3* mRNAs in WT and *mbnl* mutant adult zebrafish skeletal muscle. For all graphs, homozygous *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents one fish and error bars indicate standard error of the mean. Data were analyzed by ordinary one-way ANOVA with Tukey's multiple comparisons test. Data bars that do not share a letter above them are significantly different from one another. S5 Table contains the raw data used to generate this figure.

S9 Fig. Misregulation of many DM-associated alternative splicing events was conserved in zebrafish *mbnl* mutants.

(A) Change in percent spliced in values (ΔPSI) between mutant and WT are shown for orthologous exons in zebrafish $1^{B/B};2^{A/A};3^{C/C}$ skeletal muscle and in human DM1 patient-derived myoblasts. ρ is the Spearman's rank correlation coefficient. (B) to (F) RNA Seq ((B), (C), and (E)) and RT-PCR ((D) and (F)) analyses showing percent spliced in (PSI) of *mbnl2* exon 5 (B), *aplp2* exon 7 ((C) and (D)), and *atp6v1h* exon 6 ((E) and (F)) in WT and *mbnl* mutant adult zebrafish skeletal muscle. For all graphs, homozygous *mbnl1* mutant alleles are denoted as 1^A and 1^B, *mbnl2* alleles as 2^A, and *mbnl3* alleles as 3^A, 3^B, and 3^C. Each dot represents RNA from one fish. Error bars indicate standard error of the mean. Representative RT-PCR gels are shown above each RT-PCR graph. S5 Table contains the raw data used to generate this figure.

S10 Fig. Zebrafish *aplp2* exon 7 and its surrounding introns contain potential Mbnl protein binding sites.

Zebrafish *aplp2* exon 7 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S11 Fig. Zebrafish *atp6v1h* exon 6 and its surrounding introns contain potential Mbnl protein binding sites.

Zebrafish *atp6v1h* exon 6 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S12 Fig. The introns surrounding zebrafish *atp2a1l* exon 23 contain potential Mbnl protein binding sites.

Zebrafish *atp2a1l* exon 23 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Needle program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S13 Fig. The introns surrounding zebrafish *ryr1b* exon 72 contain potential Mbnl protein binding sites.

Zebrafish *ryr1b* exon 72 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S14 Fig. The introns surrounding zebrafish *ank3b* exon 36 contain potential Mbnl protein binding sites.

Zebrafish *ank3b* exon 36 (bold upper case letters) and its surrounding intronic sequences (lower case letters) were aligned with orthologous human sequences using the EMBOSS Water program. Potential YGCY (Y=pyrimidine) Mbnl protein binding sites are indicated in red.

S1 Table. Primers used in this study.

This table lists the name, sequence, target gene ID from ZFIN, target alternative exon coordinates from the GRCz11 genome assembly, and purpose of each primer used in this study. F: forward primer, R: reverse or antisense primer, gRNA: guide RNA.

S2 Table. Zebrafish *mbnl* mutant sequences.

For each zebrafish *mbnl* mutant used in this study, this table lists the University of Oregon allele number, exon that was mutated within the target gene, guide RNA (gRNA) used to generate the mutant, forward primers, reverse primers, and restriction sites used to genotype the mutants by restriction fragment length polymorphism, ZFIN Gene ID for the targeted gene, GRCz11 genome coordinates and coding sequence of the targeted exon, and predicted protein coding sequence. Sequences that are changed compared to WT are indicated in red. Sequences of primers from this table are listed in S1 Table.

S3 Table. Zebrafish single, double, and triple homozygous *mbnl* mutants are viable to adulthood.

For each zebrafish single, double, and triple homozygous mutant *mbnl* genotype, this table lists the cross that was used to generate the genotype, the percent of adult

offspring expected of the genotype according to Mendelian genetics, and the actual percent and number of adult offspring observed with the genotype.

S4 Table. Significantly misregulated alternative splicing events identified by RNA-Seq in zebrafish *mbnl* mutants.

Each tab contains a list of alternative splicing events (cassette exon, mutually exclusive exon, retained intron, alternative 5' splice site, or alternative 3' splice site) that were identified by RNA-Seq analysis as being significantly misregulated in adult zebrafish *mbnl* mutant skeletal muscle compared to WT. All genome coordinates are from the GRCz11 genome assembly, except for *WT* vs. $1^{B/B};2^{A/A};3^{C/C}$, for which a list of significantly misregulated cassette exons identified from a preliminary analysis of the GRCz10 genome assembly is also shown. Complete RNA-Seq data are available from NCBI GEO #GSE145270.

S5 Table. Data used to generate figures.

Each tab contains the raw data that were used to generate a figure, as well as details of the statistical analyses that were performed for that figure.

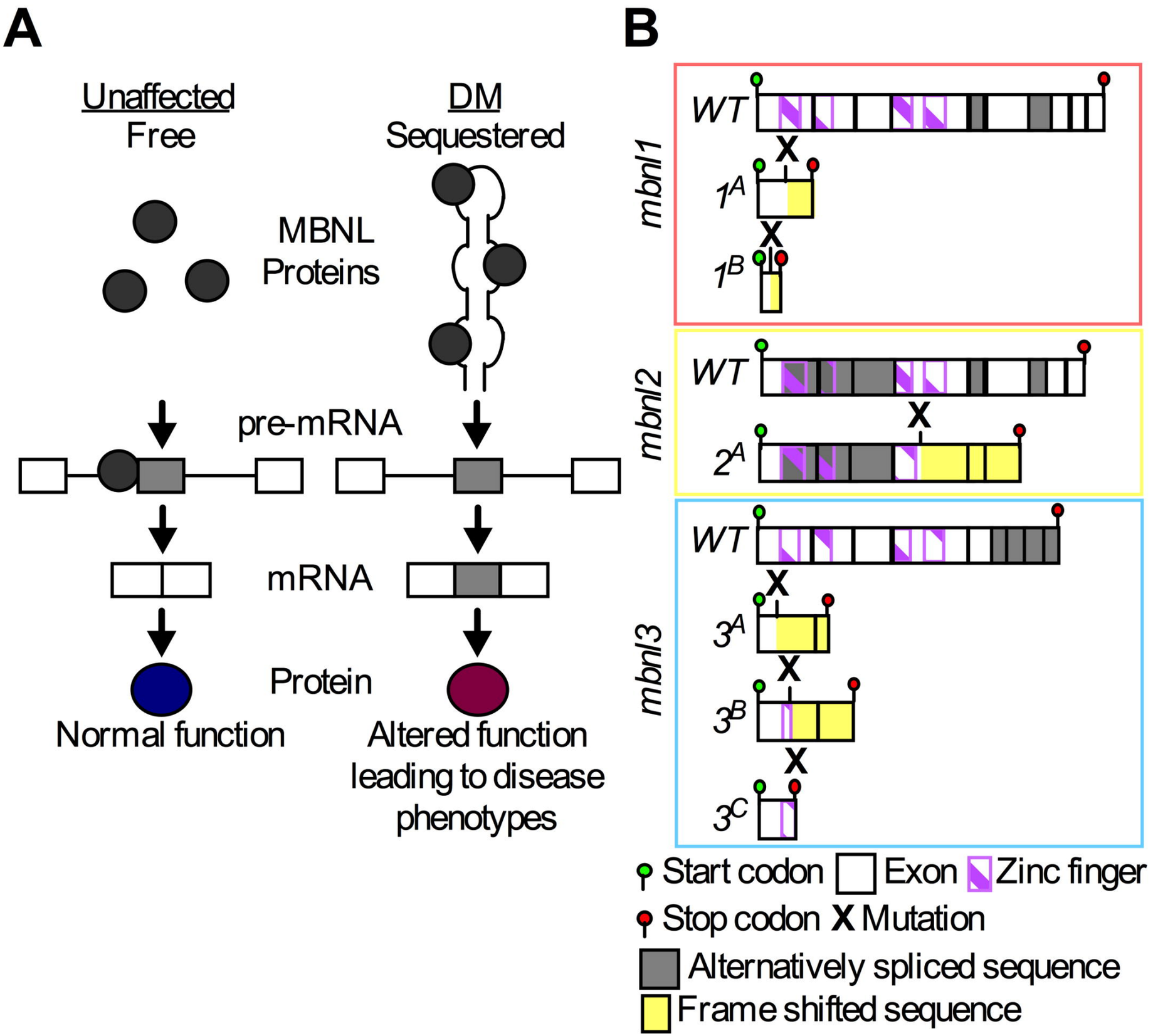
A
Unaffected
Free
DM
Sequestered
MBNL
Proteins
pre-mRNA
mRNA
Protein
Normal function
Altered function
leading to disease
phenotypes
B
mbnl1
WT
1A
1B
mbnl2
WT
2A
mbnl3
WT
3A
3B
3C
Start codon
Exon
Zinc finger
Stop codon
X Mutation
Alternatively spliced sequence
Frame shifted sequence

Fig 1

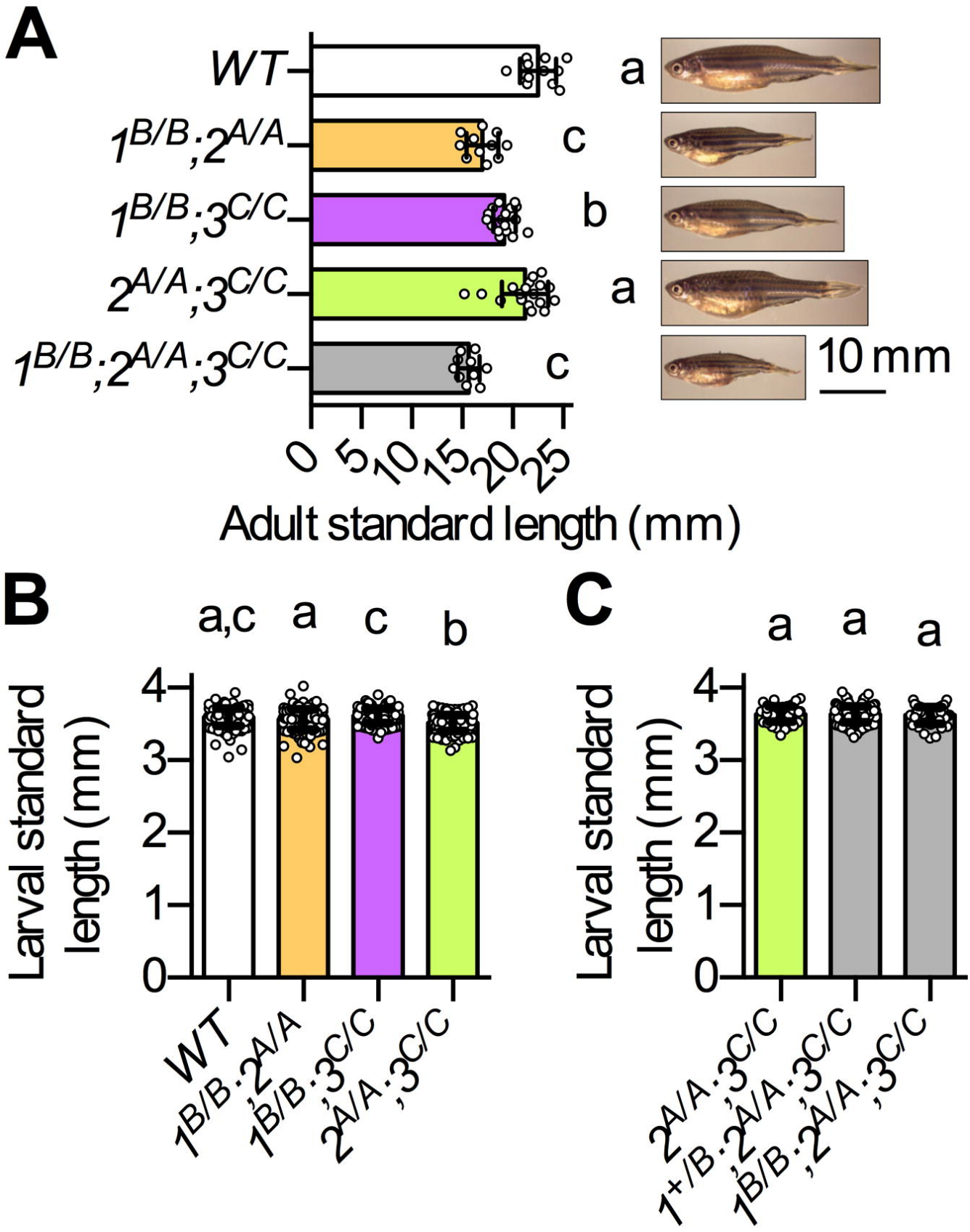
A
WT
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
1B/B;2A/A;3C/C
a
c
b
a
c
10 mm
0 5 10 15 20 25
Adult standard length (mm)
B
a,c a c b
Larval standard length (mm)
0 1 2 3 4
WT
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
C
a a a
Larval standard length (mm)
0 1 2 3 4
2A/A;3C/C
1+/B;2A/A;3C/C
1B/B;2A/A;3C/C

Fig 2

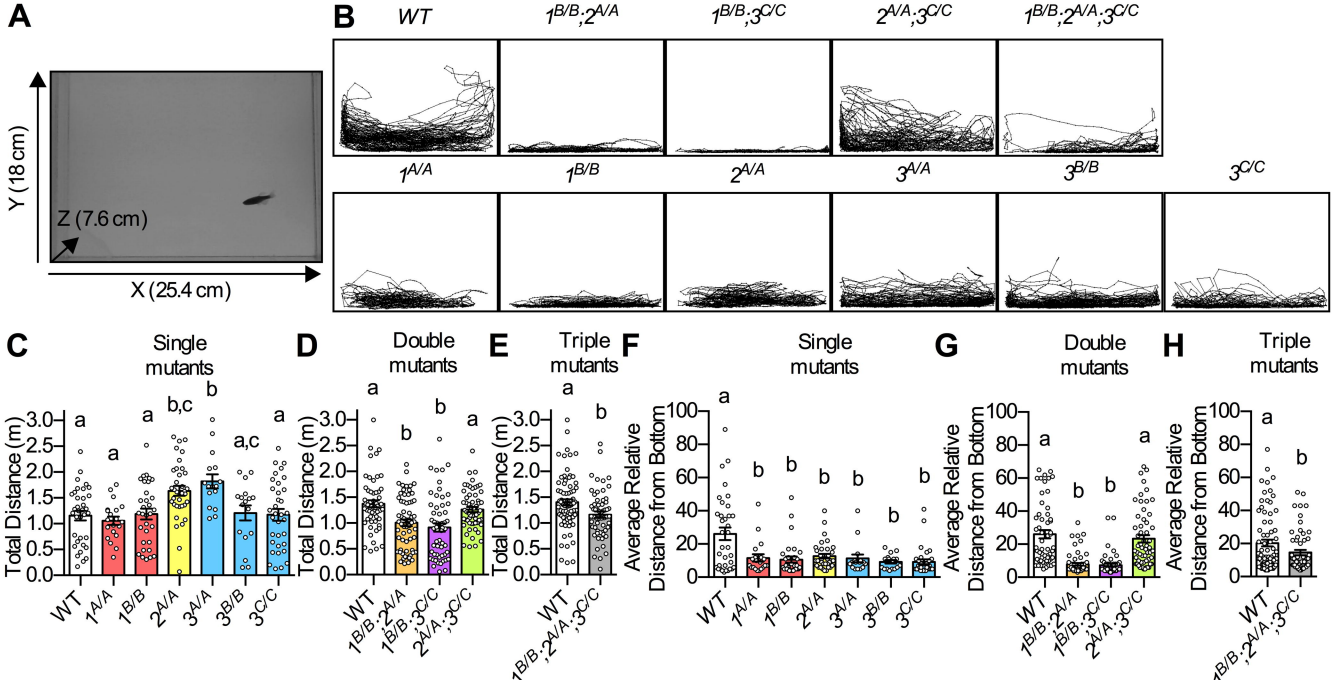
A
Y (18 cm)
Z (7.6 cm)
X (25.4 cm)
B
WT
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
1B/B;2A/A;3C/C
1A/A
1B/B
2A/A
3A/A
3B/B
3C/C
C
Single mutants
Total Distance (m)
D
Double mutants
E
Triple mutants
F
Average Relative Distance from Bottom
G
H
WT
1A/A
1B/B
2A/A
3A/A
3B/B
3C/C
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
1B/B;2A/A;3C/C

Fig 3

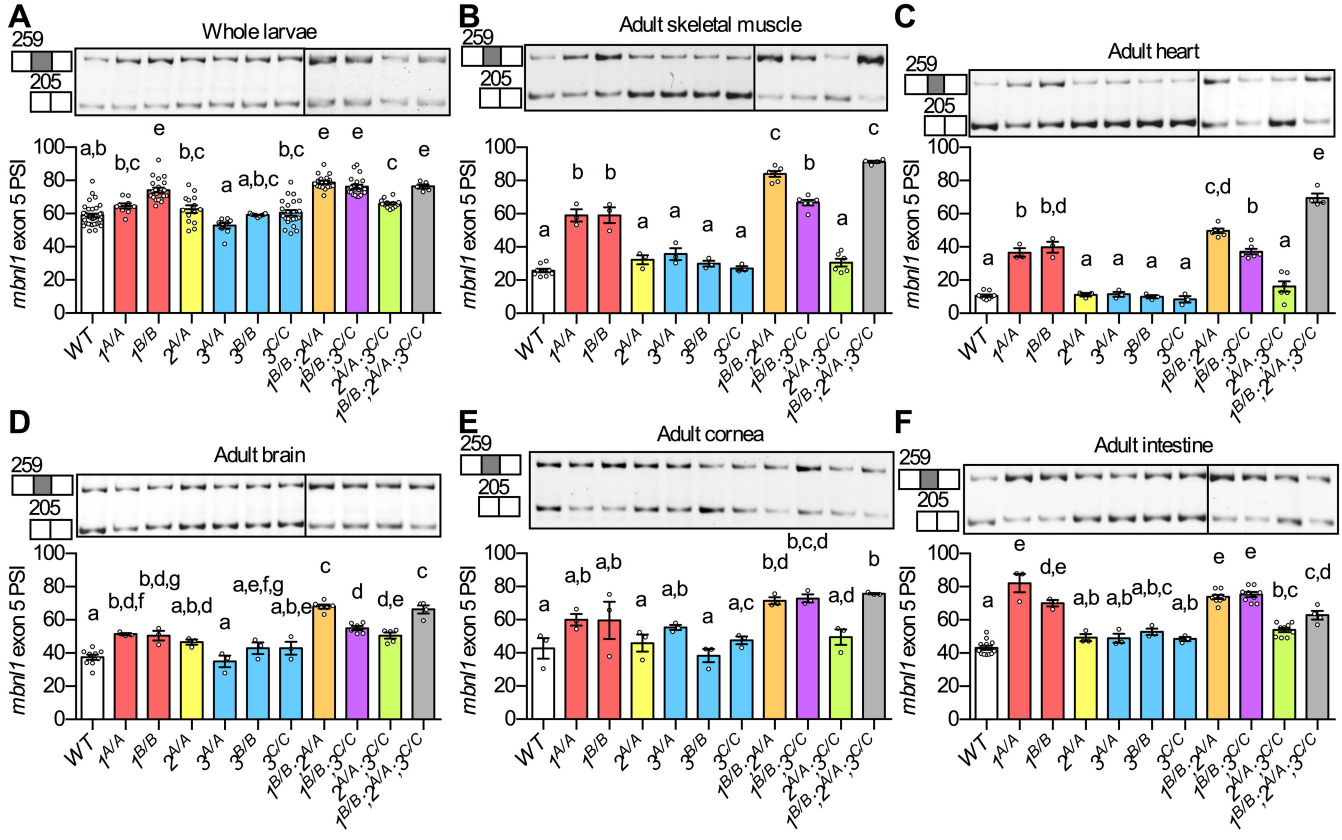
A
Whole larvae
B
Adult skeletal muscle
C
Adult heart
D
Adult brain
E
Adult cornea
F
Adult intestine
259
205
mbnl1 exon 5 PSI
WT
1A/A
1B/B
2A/A
3A/A
3B/B
3C/C
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
1B/B;2A/A;3C/C

Fig 4

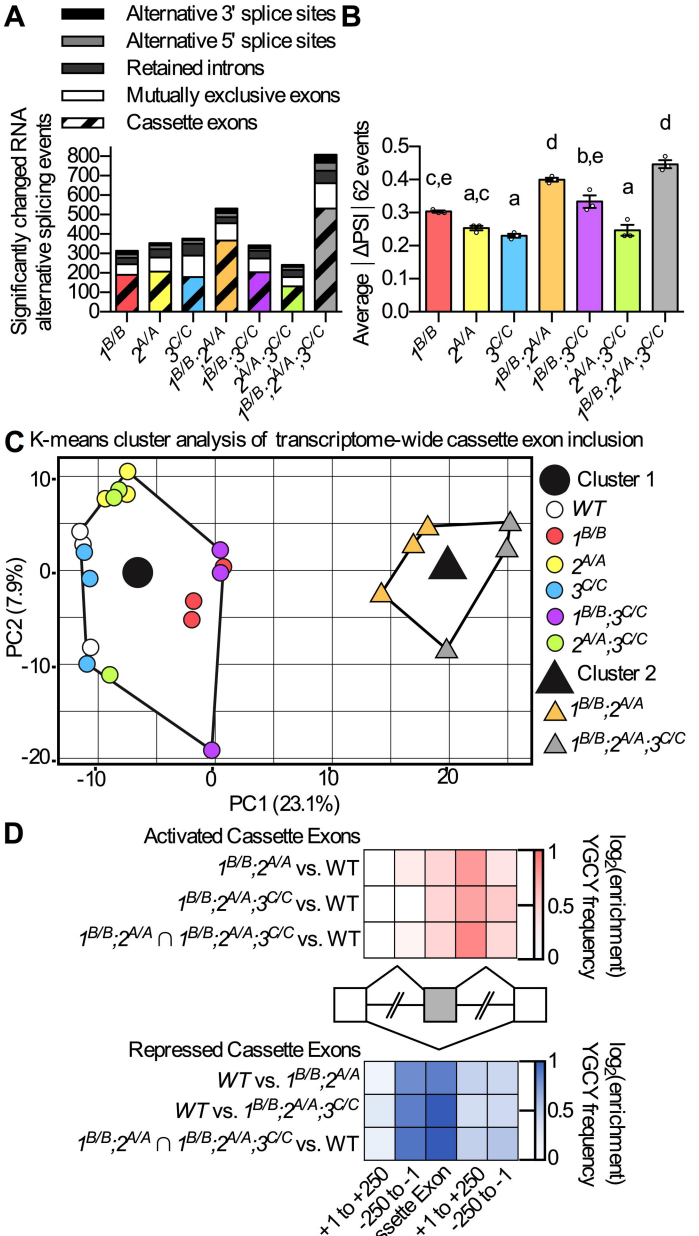
A
Alternative 3' splice sites
Alternative 5' splice sites
Retained introns
Mutually exclusive exons
Cassette exons
Significantly changed RNA alternative splicing events
B
Average |ΔPSI| 62 events
c,e
a,c
a
d
b,e
a
d
C K-means cluster analysis of transcriptome-wide cassette exon inclusion
PC2 (7.9%)
PC1 (23.1%)
Cluster 1
WT
Cluster 2
D
Activated Cassette Exons
Repressed Cassette Exons
log2(enrichment) YGCY frequency
+1 to +250
-250 to -1
Cassette Exon

Fig 5

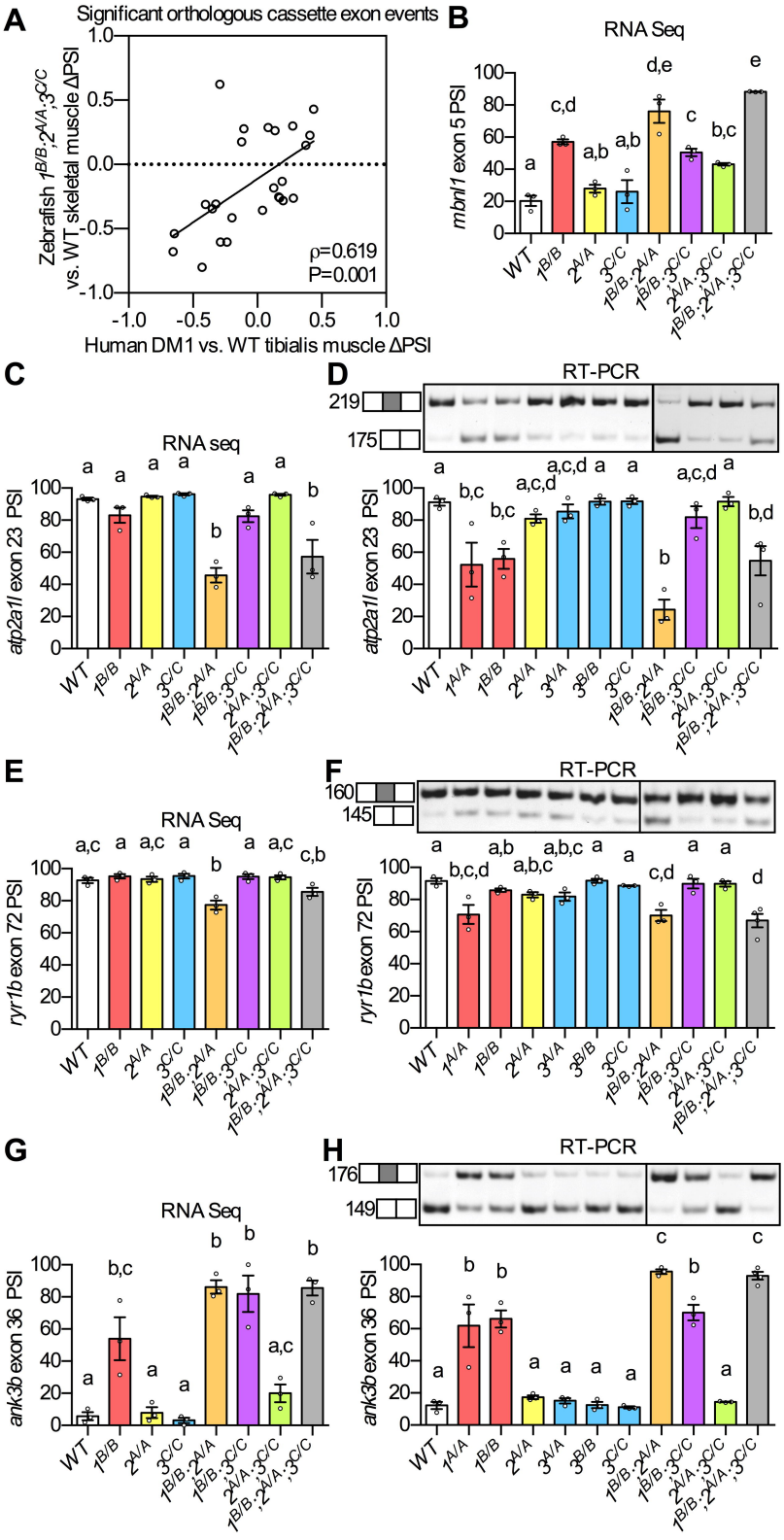
A
Significant orthologous cassette exon events
Zebrafish 1B/B;2A/A;3C/C vs. WT skeletal muscle ΔPSI
Human DM1 vs. WT tibialis muscle ΔPSI
ρ=0.619
P=0.001
B
RNA Seq
mbnl1 exon 5 PSI
C
RNA seq
atp2a1 exon 23 PSI
D
RT-PCR
219
175
atp2a1 exon 23 PSI
E
RNA Seq
ryr1b exon 72 PSI
F
RT-PCR
160
145
ryr1b exon 72 PSI
G
RNA Seq
ank3b exon 36 PSI
H
RT-PCR
176
149
ank3b exon 36 PSI
WT
1B/B
2A/A
3C/C
1B/B;2A/A
1B/B;3C/C
2A/A;3C/C
1B/B;2A/A;3C/C
1A/A
3A/A
3B/B

Fig 6