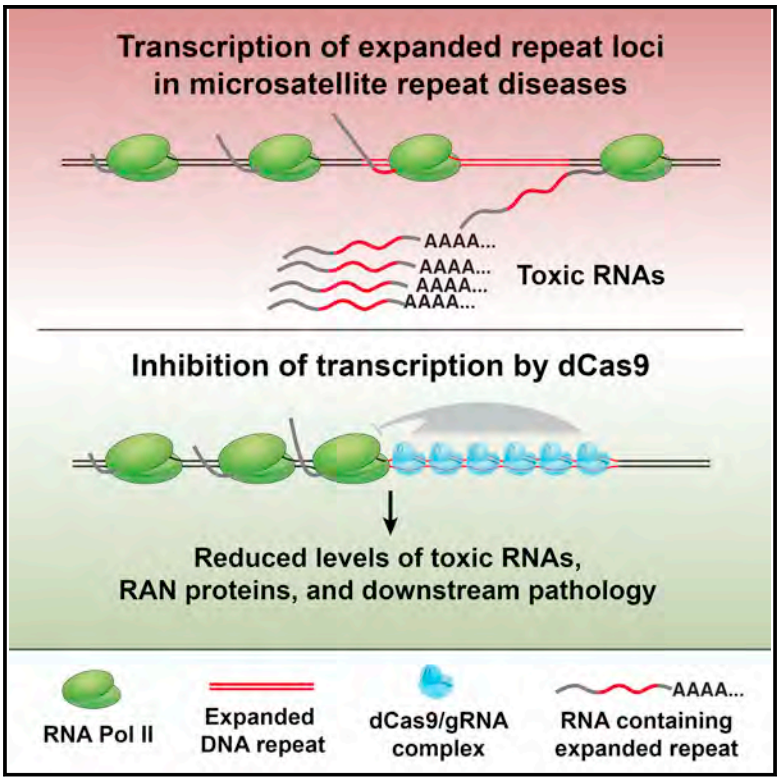
Article

Molecular Cell

Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9

Graphical Abstract



Highlights

- Deactivated Cas9 (dCas9) impedes transcription of expanded microsatellite repeats
- Repeat length, PAM sequence, and DNA strand influence extent of inhibition
- dCas9 rescues molecular and cellular phenotypes in patient-derived cell lines
- dCas9 rescues muscle phenotypes in a mouse model of myotonic dystrophy

Authors

Belinda S. Pinto, Tanvi Saxena, Ruan Oliveira, ..., Guangbin Xia, Maurice S. Swanson, Eric T. Wang

Correspondence

eric.t.wang@ufl.edu

In Brief

Pinto, Saxena, et al. demonstrate that deactivated Cas9 efficiently blocks transcription of the expanded microsatellite repeats causing diseases such as myotonic dystrophy types 1 and 2 and *C9orf72*-ALS/FTD, thereby reversing toxic downstream molecular and cellular effects of repeat expression *in vitro* and *in vivo*.

Pinto et al., 2017, Molecular Cell *68*, 1–12
November 2, 2017
https://doi.org/10.1016/j.molcel.2017.09.033

CellPress

Please cite this article in press as: Pinto et al., Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9, Molecular Cell (2017), https://doi.org/10.1016/j.molcel.2017.09.033

Molecular Cell
Article

Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9

Belinda S. Pinto,[1,3,4] Tanvi Saxena,[1,3,4] Ruan Oliveira,[1,3] Héctor R. Méndez-Gómez,[1] John D. Cleary,[1,3] Lance T. Denes,[1,3] Ona McConnell,[1,3] Juan Arboleda,[1,3] Guangbin Xia,[2,3] Maurice S. Swanson,[1,3] and Eric T. Wang[1,3,5,*]
[1]Department of Molecular Genetics & Microbiology
[2]Department of Neurology
[3]Center for NeuroGenetics
University of Florida, Gainesville, FL 32610, USA
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: eric.t.wang@ufl.edu
https://doi.org/10.1016/j.molcel.2017.09.033

SUMMARY

Transcription of expanded microsatellite repeats is associated with multiple human diseases, including myotonic dystrophy, Fuchs endothelial corneal dystrophy, and *C9orf72*-ALS/FTD. Reducing production of RNA and proteins arising from these expanded loci holds therapeutic benefit. Here, we tested the hypothesis that deactivated Cas9 enzyme impedes transcription across expanded microsatellites. We observed a repeat length-, PAM-, and strand-dependent reduction of repeat-containing RNAs upon targeting dCas9 directly to repeat sequences; targeting the non-template strand was more effective. Aberrant splicing patterns were rescued in DM1 cells, and production of RAN peptides characteristic of DM1, DM2, and *C9orf72*-ALS/FTD cells was drastically decreased. Systemic delivery of dCas9/gRNA by adeno-associated virus led to reductions in pathological RNA foci, rescue of chloride channel 1 protein expression, and decreased myotonia. These observations suggest that transcription of microsatellite repeat-containing RNAs is more sensitive to perturbation than transcription of other RNAs, indicating potentially viable strategies for therapeutic intervention.

INTRODUCTION

Microsatellite expansion diseases are a class of genetically inherited conditions associated with destabilization and expansion of short repetitive sequences in the genome, which cause pathogenic effects via multiple mechanisms, including epigenetic silencing, RNA gain of function, and/or protein gain of function (Nelson et al., 2013). These diseases, including myotonic dystrophy types 1 and 2 (DM1 and DM2), Fuchs endothelial corneal dystrophy (FECD), Huntington disease (HD), *C9orf72*-ALS/FTD (C9ALS/FTD), and spinocerebellar ataxias (SCAs), are often multi-systemic and can affect the central nervous system, muscle, and the heart. Somatic instability causes repeat expansion throughout the lifetime of an individual, with the most dramatic expansions reaching thousands of nucleotides in length in post-mitotic tissues (Kennedy et al., 2003; Thornton et al., 1994). Cellular toxicity in these diseases occurs partly due to transcription of the expanded repeat tract. For example, in DM1 and DM2, expanded CUG or CCUG repeat RNAs, respectively, sequester Muscleblind-like (MBNL) proteins from their endogenous RNA targets (Miller et al., 2000), leading to aberrant splicing patterns (Kanadia et al., 2003) and altered RNA stability/localization (Du et al., 2010; Wang et al., 2012), among other effects. In C9ALS/FTD, expanded G_4C_2 and C_4G_2 repeat RNAs sequester RNA binding proteins, as well as undergo repeat-associated non-ATG (RAN) translation to produce toxic dipeptide polymers (Gendron et al., 2014; Zu et al., 2011, 2013). Taken together, these findings suggest that silencing of expanded repeat loci holds therapeutic value.

Various approaches have been taken to silence toxic RNA or protein in microsatellite expansion diseases, including antisense oligonucleotides (Wheeler et al., 2012; Lagier-Tourenne et al., 2013; Kanadia et al., 2003; Sah and Aronin, 2011), small RNAs (Furling et al., 2003; François et al., 2011; Langlois et al., 2003; Harper et al., 2005), and small molecules (Siboni et al., 2015; Rzuczek et al., 2017). Perturbation to co-factors of RNA polymerase II (Pol II) reduces transcription through expanded repeats in HD (Liu et al., 2012) and C9ALS/FTD (Kramer et al., 2016) models, and treatment with Actinomycin D at nanomolar doses preferentially impedes transcription of CTG repeats in DM models (Siboni et al., 2015). A hypothesis is that efficiency of transcription through expanded repeats is decreased relative to non-repetitive sequences. This provides a therapeutic window through which to impede transcription of these sequences in a repeat length-dependent manner, resulting in premature termination and nascent transcript turnover.

A deactivated version of the Cas9 enzyme (dCas9) of the CRISPR system can be used to impair transcription of specific

Please cite this article in press as: Pinto et al., Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9, Molecular Cell (2017), https://doi.org/10.1016/j.molcel.2017.09.033

loci, as well as visualize, tether, and/or isolate DNA in a sequence-specific manner. In prokaryotes, dCas9 efficiently inhibits transcriptional initiation and elongation when bound to gene bodies or promoters (Qi et al., 2013). In eukaryotes, dCas9 inhibits transcriptional initiation when fused to an inhibitory domain and targeted near the transcription start site (TSS) (Gilbert et al., 2013; Lawhorn et al., 2014; Konermann et al., 2013; Gilbert et al., 2014). However, elongation inhibition by targeting dCas9 alone to the gene body (>1 kb from TSS) has been largely ineffective, even when recruiting dCas9 to >90 possible sites (Chen et al., 2013).

Here, we test the hypothesis that expanded microsatellite repeats are highly sensitive to transcriptional blockade by dCas9, even in the context of elongating Pol II. We find that efficiency of inhibition follows rules similar to those observed in non-repetitive contexts, with clear dependencies on proto-spacer adjacent motif (PAM) sequence and targeted DNA strand. In addition, application of this approach to cell and animal models of disease rescues downstream pathogenic consequences. These observations describe application of the CRISPR/Cas9 system to impede transcription of expanded microsatellites in a strand-dependent manner and suggest that targeting transcription in a repeat length-dependent manner may be a viable therapeutic strategy for these diseases.

RESULTS

dCas9-gRNA Complexes Reduce Abundance of Repeat-Containing RNAs in a Length-, PAM-, and Strand-Dependent Manner

We hypothesized that dCas9 might impede transcription of expanded microsatellite repeats more potently than non-repetitive sequences because expanded repeats may: (1) present challenges for Pol II elongation even under normal conditions and (2) allow for high levels of dCas9 recruitment using a single guide RNA (gRNA) sequence, forming a substantial block to the elongating polymerase (Figure 1A). To test this hypothesis, we employed a plasmid-based strategy in cell culture to examine the effects of dCas9 recruitment on transcription of CTG/CAG repeats that occur in DM1, FECD, HD, and SCAs. We utilized plasmids containing 0, 12, 40, 240, 480, or 960 CTG repeats located >1.5 kb downstream of the TSS in the *DMPK* 3′ UTR. To recruit dCas9 to these repeats, we designed gRNAs targeting CTG/CAG repeats in each of three possible nucleotide phases of the non-template strand and one phase of the template strand (Figure 1B, left). Each PAM was, therefore, constrained to CAG, AGC, GCA, or CTG. To precisely measure the expression of the repeat transcripts, we developed an amplicon-based, deep-sequencing assay called ''Measurement of Barcoded Transcripts by Amplicon Sequencing'' (MBTA-seq) (Figure S1). For this assay, we modified plasmids by introducing distinct 8 nt barcodes downstream of each CTG repeat tract. RNA was harvested from cells transfected with these plasmids, polyA+ selected, amplified by RT-PCR across the barcode region but avoiding the repeat-containing region, and deep sequenced (Figure S1). This assay was highly reproducible and allowed quantitation of multiple repeat-containing transcripts in the same pool of cells (Table S1). In the presence of dCas9 and $(CAG)_6$ gRNA, we observed a dramatic reduction in expression of RNAs containing expanded CUG repeats (Figure 1B, right). Knockdown efficiency increased with the number of repeats, presumably due to the increased number of dCas9-gRNA complexes recruited. Twelve CTG repeats, likely recruiting at most a single dCas9-gRNA complex, showed ~50% repeat-containing RNA remaining, and 40 CTG repeats showed ~25% remaining. Repeat lengths > 240 CTG showed only ~5% remaining. $(AGC)_6$ and $(GCA)_6$ gRNAs showed poor knockdown efficiency, consistent with previously described SpCas9 PAM preferences, where NGG is best, NAG is second best, and NCG/NTG are equally disfavored (Kuscu et al., 2014). $(CUG)_6$ gRNA showed little to no knockdown of transcripts containing expanded CUG repeats, potentially due to a weak PAM as well as targeting to the template strand. Importantly, dCas9 was required for decreased expression of repeat-containing transcripts, as presence of gRNAs alone did not lead to knockdown (Figure S2A).

These observations suggest that recruitment of dCas9-gRNA complexes can impede transcription of expanded microsatellite repeat tracts in a repeat length-dependent manner. However, these experiments were performed using transiently transfected plasmids, which may not accurately model all aspects of transcriptional regulation in a genomic context. Furthermore, experiments to assess binding of protein complexes to DNA loci are most commonly performed using genomic targets and not plasmids. Therefore, we established a HeLa cell line in which transgenes for each of the six CTG repeat lengths, with associated barcodes, were stably integrated (Figure S2B). To measure DNA binding by dCas9-gRNA complexes, we performed chromatin immunoprecipitation (ChIP) against dCas9 in the presence of $(CAG)_6$ gRNA or non-targeting control gRNA. Similar to MBTA-seq, we deep sequenced the barcodes to quantitate abundance of immunoprecipitated DNA encoding each repeat length. To confirm transcriptional repression in this cell line, we measured RNA abundance by MBTA-seq following transfection of dCas9-gRNA complexes. We observed that dCas9-gRNA binding to DNA increased as a function of repeat length, concomitant with a decrease in relative RNA abundance, when comparing $(CAG)_6$ gRNA to control (Figure 1C). In a genomic context, dCas9-gRNA reduced the abundance of transcripts containing 960 CUG repeats to ~20% and was associated with ~8-fold increased binding relative to genomic loci without CTG repeat tracts. These observations suggest that binding of dCas9-gRNA complexes to DNA impedes transcription of repeat tracts and that both binding and potency of transcriptional inhibition increases with the number of repeats.

To test whether dCas9-gRNA complexes can impede transcription of other expanded repeats, we assayed CCTG repeat tracts, which cause DM2, and observed similar reduction in expression of CCUG repeat-containing RNAs (Figure 1D). Here, the tetranucleotide repeat allowed testing of four distinct gRNAs targeting the non-template strand—$(CAGG)_5$, $(AGGC)_5$, $(GGCA)_5$, and $(GCAG)_5$—corresponding to AGG, GGC, GCA, and CAG PAMs, respectively. Consistent with PAM preferences of SpCas9, knockdown was most efficient with the $(CAGG)_5$ gRNA, although the $(AGGC)_5$ and $(GCAG)_5$ gRNAs were also quite effective, in contrast to the $(CCUG)_5$ gRNA. Similar to our

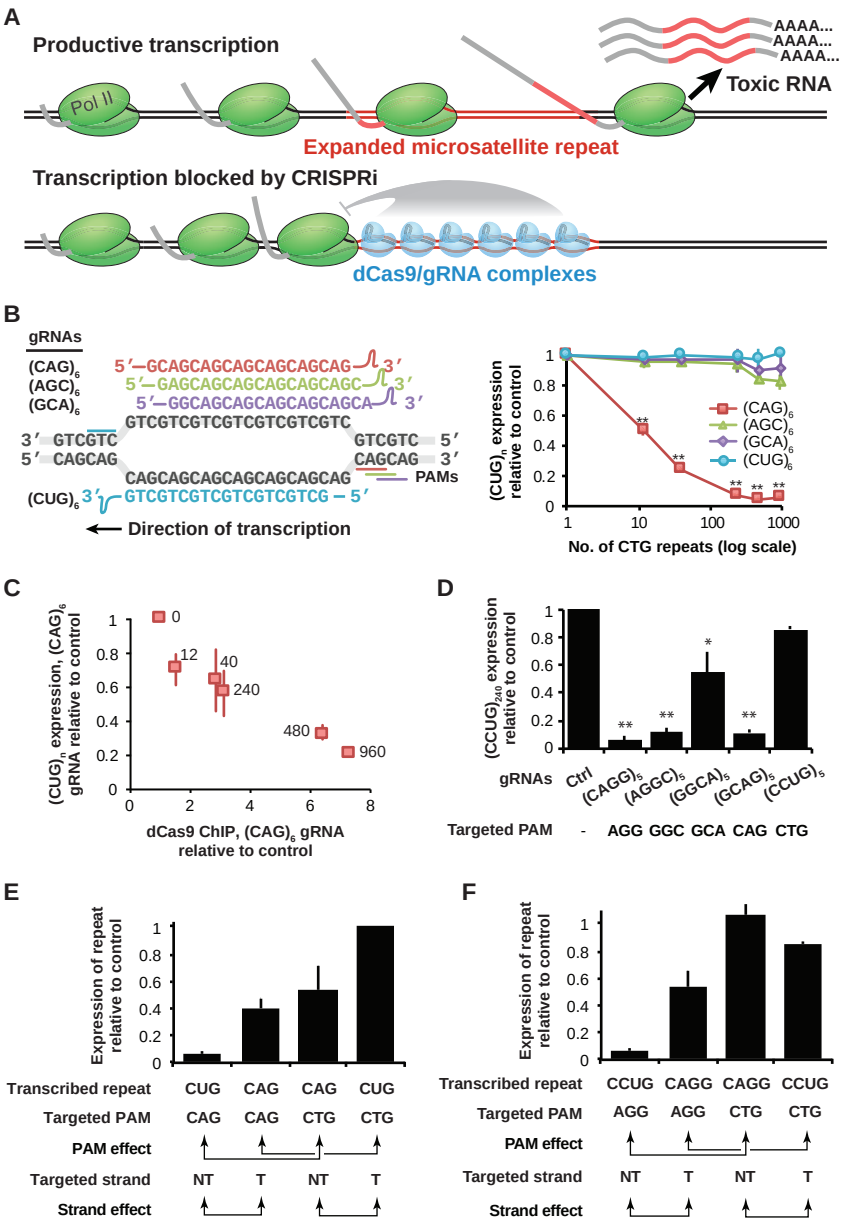


Figure 1. Deactivated SpCas9 Impedes Transcription of Expanded Microsatellite Repeats in a Length-, PAM-, and Strand-Dependent Manner

(A) Proposed model for how recruitment of dCas9/gRNA complexes to expanded microsatellite repeats impedes transcription by Pol II.

(B) Schematic of gRNAs used to target transcription of CTG repeats (left). Abundance of CUG repeat RNA in the presence of dCas9/gRNAs targeting the repeat tracts of various lengths in HeLa cells, relative to the same RNA species with zero repeats (right). Error bars show standard deviation.

(C) Relative dCas9 ChIP signal across all repeat lengths versus percent RNA remaining following $(CAG)_6$ gRNA relative to control gRNA treatment. dCas9 ChIP signal is computed as dCas9 IP divided by input chromatin with $(CAG)_6$ gRNA treatment, divided by dCas9 IP divided by input chromatin with control gRNA treatment. Relative abundance of repeat-containing loci and RNAs was assessed by deep sequencing of the barcodes. Error bars show SEM.

(D) Abundance of CCUG repeat RNA in the presence of dCas9 and gRNAs targeting the $(CCTG)_{240}$ repeat tract in HeLa cells, relative to the same RNA species with zero repeats.

(E) Abundance of RNAs containing 960 CUG or CAG repeats in the presence of dCas9 and $(CUG)_6$ or $(CAG)_6$ gRNAs in HeLa cells, relative to RNA species with zero repeats. Arrows denote comparisons relevant for assessing PAM- or strand-dependent effects on efficacy.

(F) Abundance of RNAs containing 240 CCUG or CAGG repeats in the presence of dCas9 and $(CCUG)_5$ or $(CAGG)_5$ gRNAs in HeLa cells, relative to RNA species with zero repeats. Arrows denote comparisons relevant for assessing PAM- or strand-dependent effects on efficacy. All significance tests are by two-tailed t test, $^{**}p < 0.0005$, $^{*}p < 0.005$.

studies of expanded CUG repeats, presence of gRNAs alone did not lead to knockdown (Figure S2C).

In both bacterial and mammalian systems, blockade of Pol II is most effective when targeting dCas9 to the non-template strand of transcribed DNA, likely due to Pol II helicase activity that unwinds nucleic acids hybridized to the template strand (Qi et al., 2013). Our observations are consistent with this model, but thus far, we have not made comparisons that exclusively control for either the PAM or targeted strand. In addition, some of the most efficient gRNAs are those that could also facilitate dCas9 targeting of RNA via complementarity between the gRNA and the repeat-containing RNA. To study how the PAM and targeted strand separately influence the efficacy of transcriptional blockade, and to clarify DNA versus RNA-targeting mechanisms, we generated $(CAG)_{960}$ and $(CAGG)_{240}$ plasmids suitable for MBTA-seq. These constructs are identical to their $(CTG)_{960}$ and $(CCTG)_{240}$ repeat-containing counterparts except for the repeat tract and barcode. We then used MBTA-seq to measure the efficiency of CUG_{960} or CAG_{960} knockdown in the presence of either $(CAG)_6$ or $(CUG)_6$ gRNA (Figure 1E). By testing all four combinations, we could separate PAM-dependent effects from strand-dependent effects. Targeting the non-template DNA strand reduced expression more effectively than targeting the template strand. Specifically, $(CUG)_{960}$ was reduced to ~5% by a $(CAG)_6$ gRNA, while $(CAG)_{960}$ was reduced to ~40% by the same gRNA. These two conditions effectively compare targeting of non-template versus template strand, controlling for the CAG PAM. Similarly, $(CAG)_{960}$ was reduced to ~50% by the $(CUG)_6$ gRNA, and $(CUG)_{960}$ remained relatively high at ~97%. Again, targeting of the non-template strand was more efficient than targeting of the template strand, controlling for the CTG PAM.

PAM-dependent effects were quantified by comparing the abundance of $(CUG)_{960}$ in the presence of $(CAG)_6$ gRNA relative to the abundance of $(CAG)_{960}$ in the presence of $(CUG)_6$ gRNA. Here, both gRNAs target the non-template strand but use different PAMs. We observed ~5% RNA remaining with the CAG PAM and ~55% RNA remaining with the CTG PAM, controlling for targeted strand, consistent with the CAG PAM being more effective. Similarly, measurement of $(CAG)_{960}$ RNA in the presence of $(CAG)_6$ gRNA and $(CUG)_{960}$ RNA in the presence of $(CUG)_6$ gRNA allows comparison of both PAMs, controlling for targeted strand. Again, we observed more effective silencing with a CAG PAM (~40% RNA remaining) as compared to the CTG PAM (no change in RNA). We observed similar trends with CCTG/CAGG repeats (Figure 1F). $(CCUG)_{240}$ was reduced to ~5% by $(CAGG)_5$ gRNA and $(CAGG)_{240}$ to ~55% by $(CAGG)_5$ gRNA (non-template versus template strand, AGG PAM). Although we did not observe the same template versus non-template effect when silencing with a CTG PAM, the $(CCUG)_5$ gRNA was largely ineffective, showing little to no silencing.

Overall, these results separate the effects of PAM sequence and targeted strand in the context of transcriptional blockade. Furthermore, they support a model in which dCas9/gRNA complexes target repeat-containing DNA, because reductions in RNA abundance are achieved even when using gRNAs that are not complementary to transcribed RNAs.

dCas9-Mediated Transcriptional Inhibition Rescues Splicing Defects and Blocks RAN Translation in Cell-Based Models of DM and C9ALS/FTD

Downstream symptoms of many repeat expansion diseases are caused by the expression of toxic RNA species, which sequesters RNA binding proteins necessary for cellular functions, causing downstream changes to the transcriptome and proteome (Kanadia et al., 2003; Nelson et al., 2013; Miller et al., 2000; Du et al., 2010; Wang et al., 2012; Gendron et al., 2014; Zu et al., 2011, 2013). To assess whether dCas9-mediated transcriptional silencing can rescue molecular and cellular phenotypes in disease models, we used a HeLa cell-based model of DM1, in which $(CTG)_{480}$ repeats are expressed from a plasmid (Ho et al., 2005). Co-expression of dCas9 and $(CAG)_6$ gRNA led to a reduction in the percentage of cells showing CUG-containing RNA foci, as well as a reduction in the number of foci per cell (Figures 2A and 2B). We measured rescue of MBNL-dependent splicing misregulation, characteristic of DM1 cells (Ho et al., 2004; Charlet-B et al., 2002; Savkur et al., 2001), using a splicing minigene reporter containing MBNL1 exon 5 (Parkesh et al., 2012). The percent spliced in, psi (Ψ), of this exon is regulated by MBNL proteins and changes from 10% in healthy HeLa cells to ~70% in cells expressing $(CUG)_{480}$ repeats. Co-transfection of dCas9 and $(CAG)_6$ gRNA partially rescued splicing dysregulation, reducing Ψ to ~35% (Figure 2C). This rescue required both dCas9 and $(CAG)_6$ gRNA, indicating that $(CAG)_6$ gRNA expression alone cannot neutralize toxic CUG RNA by displacing MBNL protein or facilitating degradation. Splicing defects in DM1 depend on total CUG repeat load; Ψ values for MBNL1 exon 5 increased with the length of the CTG repeat tract transfected into cells (Figure 2D) (Wagner et al., 2016). Interestingly, co-expression of dCas9 and $(CAG)_6$ gRNA yielded a similar rescue of splicing across all repeat lengths exhibiting misplicing, i.e., 240, 480, and 960 CUG repeats.

The CTG, CCTG, and G_4C_2 repeats associated with DM1, DM2, and C9ALS/FTD, respectively, undergo RAN translation (Zu et al., 2011, 2013, 2017). We hypothesized that dCas9-mediated transcriptional repression would also yield reduced RAN protein abundance. Western blot analyses of RAN peptides, facilitated by downstream protein tags, revealed reduction in CTG, CCTG, and G_4C_2 RAN product (Figure 2E). In all three western analyses, RAN proteins were observed only with repeat-containing plasmids and were slightly reduced upon expression of a control, non-targeting gRNA, potentially due to competition of plasmids for transcriptional machinery. Consistent with measurements of CUG-containing RNA levels (Figure 1B), the FLAG-tagged poly-Ala peptide translated from CUG repeats was dramatically reduced in the presence of $(CAG)_6$ gRNA and modestly reduced in the presence of $(CUG)_6$ gRNA (Figure 2E). Importantly, RAN protein was not reduced in the presence of $(CAG)_6$ gRNA alone without dCas9 (Figure S2D). Hemagglutinin (HA)-tagged poly-LPAC translated from CCUG repeats was dramatically reduced in the presence of $(CAGG)_5$, $(AGGC)_5$, and $(GCAG)_5$ gRNAs (Figure 2F), but not in the presence of $(CCUG)_5$ gRNA (Figure S2E). To test effects of dCas9 on levels of poly-GlyPro translated from G_4C_2 repeats, we designed two gRNAs with NGG PAMs: $(C_4G_2)_3$ that would target the non-template strand and $(G_4C_2)_3$ that would target the template strand. Consistent with our studies of CTG and CCTG repeats, dCas9-gRNA complexes targeted to $(G_4C_2)_{120}$ exhibited silencing of RAN peptide in a strand-dependent manner (Figure 2G) and did not occur in the presence of gRNA alone (Figure S2F). While reduction in RAN protein was achieved with both gRNAs, silencing was more effective with gRNA targeting the non-template strand.

dCas9-Mediated Transcriptional Inhibition Reduces Nuclear RNA Foci and Rescues Splicing Defects in Human DM1 Myoblasts

We next tested whether this approach could impede transcription of expanded repeats in the native *DMPK* locus and in primary human DM1 myoblasts and chose to deliver dCas9 by adeno-associated virus (AAV). Due to AAV genome size limits, we turned to the smaller *S. aureus* Cas9 (SaCas9), packaged with a U6 promoter-driven gRNA (Ran et al., 2015) and deactivated via D10A and H557A mutations (Nishimasu et al., 2015). Since dSaCas9 exhibits PAM preferences distinct than dSpCas9 (NNGRR versus NGG), we used MBTA-seq to confirm that dSaCas9 with a $(CAG)_6$ gRNA could impede transcription of expanded CTG repeats in a length-dependent manner (Figure S3A). Next, we packaged dSaCas9 with control or $(CAG)_6$ gRNA into AAV2/6, as we found that the AAV6 capsid efficiently infects myoblasts in culture (Figure S3B). dSaCas9 protein carrying five nuclear localization signals was found to be nuclear, although at times also cytoplasmic (Figure 3A; Figure S3B). Similar to our studies in HeLa, the number of CUG repeat foci per cell was decreased in the presence of dSaCas9 and $(CAG)_6$ gRNA relative to control gRNA (Figures 3A and 3B).

Although splicing events in DM1 patient muscle have been well characterized, cultured myoblasts do not express many transcripts present in mature muscle. To identify molecular

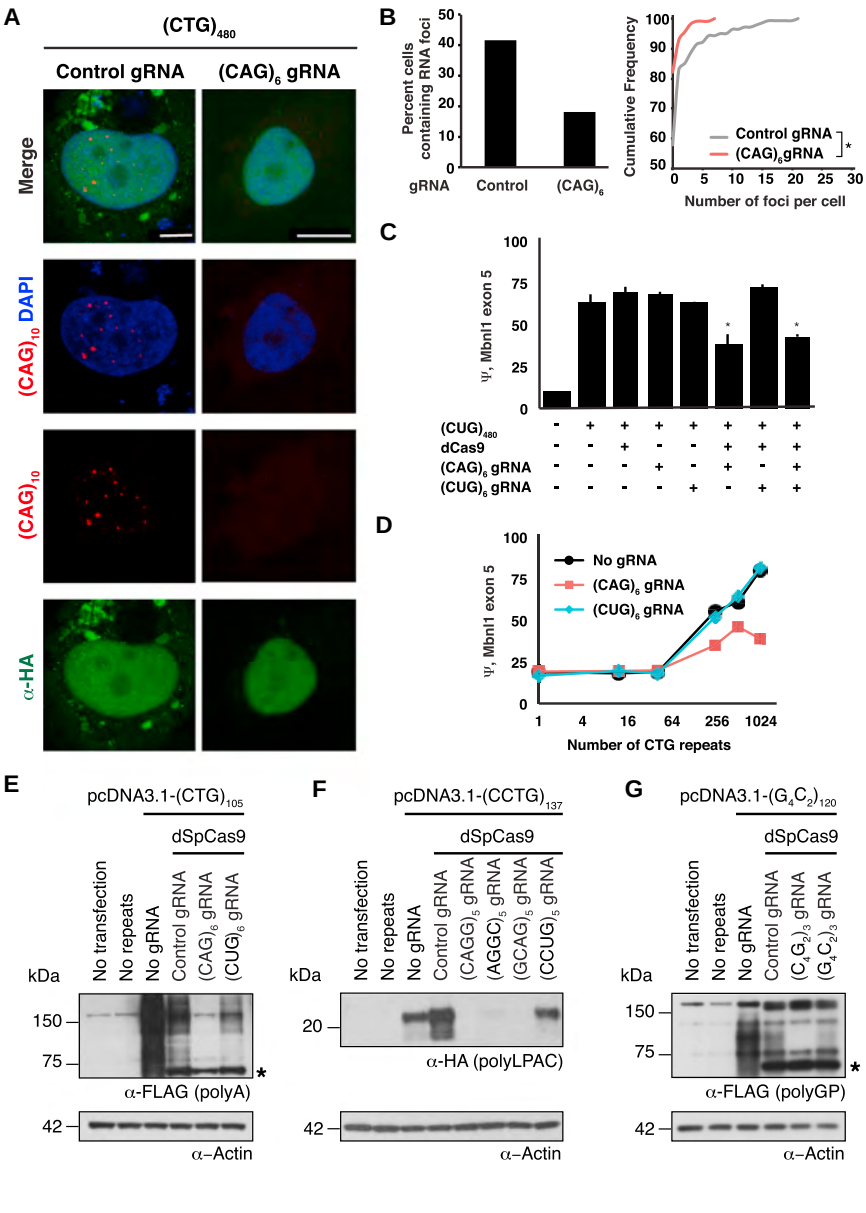


Figure 2. Transcriptional Inhibition Rescues Molecular and Cellular Phenotypes in Cell Culture Models of DM1, DM2, and C9ORF72/ALS/FTD

(A) Representative FISH-IF images of HeLa cells transfected with plasmids expressing $(CTG)_{480}$ repeats, HA-dCas9, and control or $(CAG)_6$ gRNAs. RNA foci are shown in red, dCas9 in green, and DNA (DAPI) in blue. Scale bar, 10 μm.

(B) Quantitation of HA-positive cells showing nuclear RNA foci (left) and the cumulative density function of HA-positive cells with a given number of RNA foci (right) in the presence of dCas9 and control or $(CAG)_6$ gRNAs (Control: 139 cells, $(CAG)_6$: 111 cells, Kolmogorov-Smirnov test, $^*p < 0.005$).

(C) Ψ of MBNL1 exon 5 expressed in a minigene context in HeLa cells in the presence of combinations of plasmids encoding $(CTG)_{480}$ repeats, dCas9, and $(CAG)_6$ or $(CUG)_6$ gRNAs. $n > 3$ for all conditions (two-tailed t test, $^*p < 0.005$).

(D) Ψ of the MBNL1 exon 5 minigene in HeLas in the presence of plasmids encoding 0, 12, 40, 240, 480, or 960 CTG repeats, as well as dCas9 and $(CAG)_6$ or $(CUG)_6$ gRNAs. $n > 3$ for all conditions. Error bars are too small to be visible.

(E) Western blot to detect FLAG-tagged poly-Ala RAN peptides expressed from DM1 CTG repeats. Cells were transfected with various combinations of plasmids encoding no or $(CTG)_{105}$ repeats with combinations of dCas9 and control, $(CAG)_6$, or $(CUG)_6$ gRNAs. The peptides migrate as a 75 and 150 kD smear. The asterisk indicates a cross-reacting protein produced by the dSpCas9 plasmid.

(F) Western blot against the HA-tagged LPAC RAN protein expressed from DM2 CCTG repeats. Cells were transfected with various combinations of plasmids encoding no or $(CCTG)_{137}$ repeats with combinations of dCas9 and control, $(CAGG)_5$, $(AGGC)_5$, $(GCAG)_5$, or $(CCUG)_5$ gRNAs. The peptide migrates at 20 kD.

(G) Western blot against the FLAG-tagged poly-GlyPro RAN protein expressed from ALS G_4C_2 repeats. Cells were transfected with various combinations of plasmids encoding no or $(G_4C_2)_{120}$ repeats with combinations of dCas9 and control, $(C_4G_2)_3$, or $(G_4C_2)_3$ gRNAs. The peptides migrate as a 75 and 120 kD smear. The asterisk indicates a cross-reacting protein produced by the dSpCas9 plasmid; β-Actin serves as the loading control and n = 3 for (E), (F), and (G) (representative blot shown).

changes appropriately modeling *in vivo* DM1 biology, we further characterized the DM1 myoblasts described above as well as an unaffected line. Four RNA sequencing (RNA-seq) libraries were created from the DM1 cells and two from the unaffected line, each grown in low- or high-serum conditions. To identify splicing changes appropriately modeling those observed in human DM1 muscle, we first defined a high-confidence set of DM1-relevant splicing events by re-analyzing a set of 55 transcriptomes (44 DM1 and 11 unaffected) from human tibialis biopsies (GEO: GSE86356). We identified splicing events whose inclusion level, Ψ, strongly correlated to the concentration of free, functional MBNL protein in the affected tissue, as previously described (Wagner et al., 2016). We fit sigmoid curves describing the relationship between MBNL concentration and Ψ, for example, for MBNL1 exon 5 (Figure 3C). The best-fitting events were selected (see STAR Methods; Data S1), and their dysregulation (ΔΨ, tibialis anterior, DM1 minus unaffected) was plotted against dysregulation observed in the DM1 myoblasts (ΔΨ, primary myoblast, DM1 minus unaffected) (Figure 3D). These filtering steps yielded a set of myoblast splicing events that models splicing events in human DM1 muscle.

To assess splicing rescue by dCas9, we first performed RT-PCR to measure the inclusion level of MBNL1 exon 5. While viral infection with an AAV encoding dSaCas9 and control gRNA yielded a Ψ of ~50%, infection with AAV encoding dSaCas9 and $(CAG)_6$ gRNA rescued Ψ to ~20% (Figure 3E). We assessed

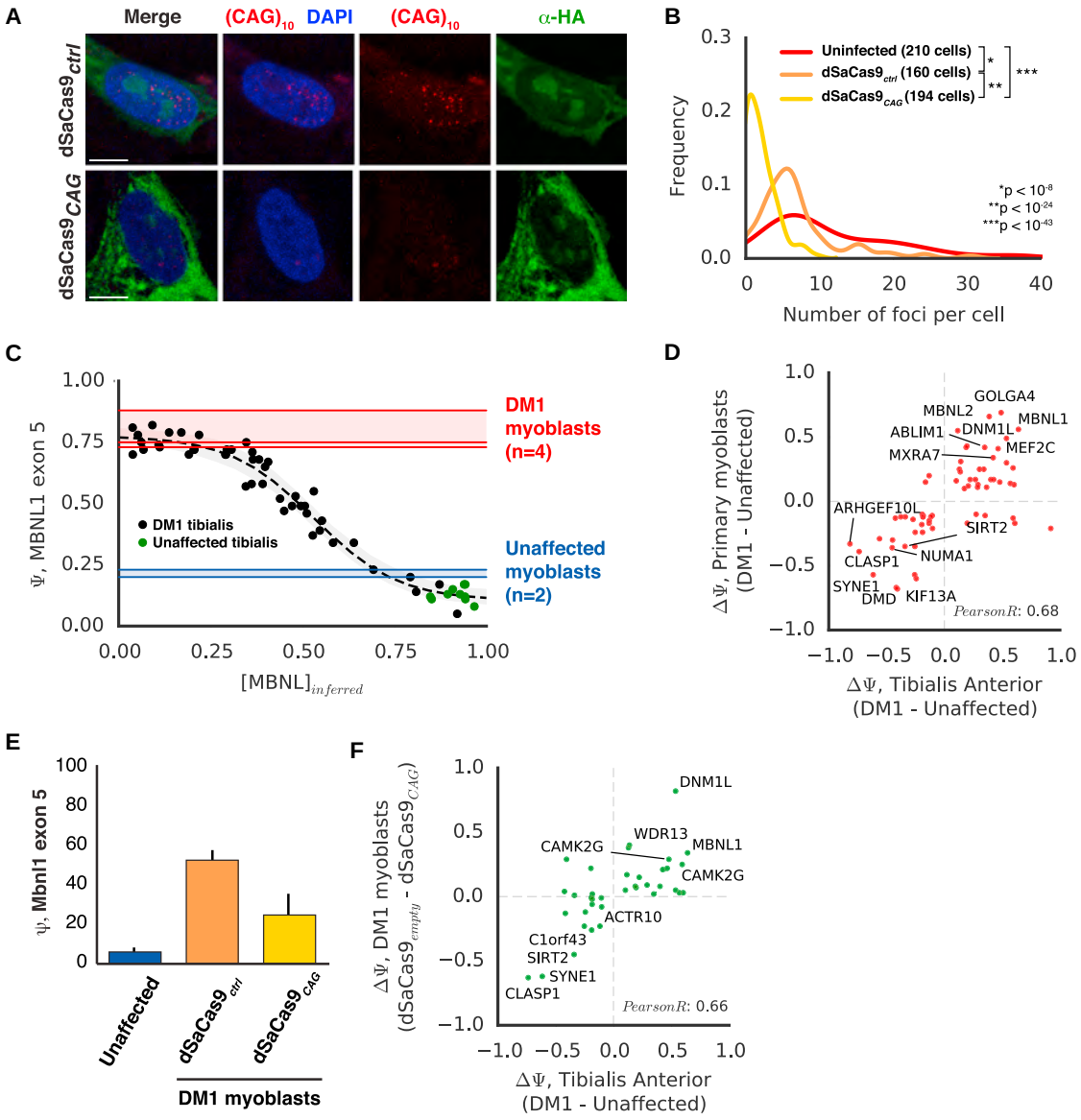


Figure 3. AAV-dSaCas9 Rescues Molecular and Cellular Phenotypes in Human DM1 Myoblasts

(A) Representative FISH-IF images of DM1 myoblasts infected with AAV-dSaCas9-control or $(CAG)_6$ gRNAs. RNA foci are shown in red, dCas9 in green, and DAPI in blue. Scale bar, 10 μm.

(B) Probability density function of cells with a given number of RNA foci in the presence of control or $(CAG)_6$ gRNAs (Kolmogorov-Smirnov test for statistical significance).

(C) MISO estimates of MBNL1 exon 5 Ψ in 11 unaffected and 44 DM1 TA biopsies plotted in order of $[MBNL]_{inferred}$ as previously described. A sigmoid curve was fit to these points and shown (dashed line) with 90% confidence intervals (gray shading). MISO estimates of MBNL1 exon 5 Ψ values in unaffected and DM1 primary myoblast lines (2 and 4 replicates, respectively) are shown in blue and red lines, respectively, with ranges also shaded.

(D) Scatterplot for splicing events regulated in both TA biopsies and the human myoblast lines described in (C). 115 splicing events were selected from TA biopsies based on best sigmoid fits with $[MBNL]_{inferred}$ as in (C). The x axis of the scatterplot is the mean Ψ in the most severely affected biopsies (<0.33 $[MBNL]_{inferred}$) minus the mean Ψ across unaffected individuals. The y axis of the scatterplot is the mean Ψ in the DM1 myoblast minus the mean Ψ in the unaffected myoblast line. Labeled points have $\Delta\Psi > 0.3$ in both conditions, a sigmoid fit < 1.3, and a y axis monotonicity score > 1 (see STAR Methods). Pearson correlation is listed.

(E) MBNL1 exon 5 Ψ, assessed by RT-PCR in unaffected myoblasts and DM1 myoblasts infected with AAV-dSaCas9-control or $(CAG)_6$ gRNAs.

(F) Scatterplot illustrating changes in Ψ in response to AAV-dSaCas9 $(CAG)_6$ gRNA for splicing events exhibiting concordant behavior in (D). The x axis is as in (D), but the y axis is ΔΨ in DM1 myoblasts, AAV-dSaCas9-control gRNA treatment minus AAV-dSaCas9-$(CAG)_6$ gRNA treatment. Labeled points have $\Delta\Psi > 0.1$ in both conditions and a y axis Bayes Factor > 5 (see STAR Methods). Pearson correlation is listed.

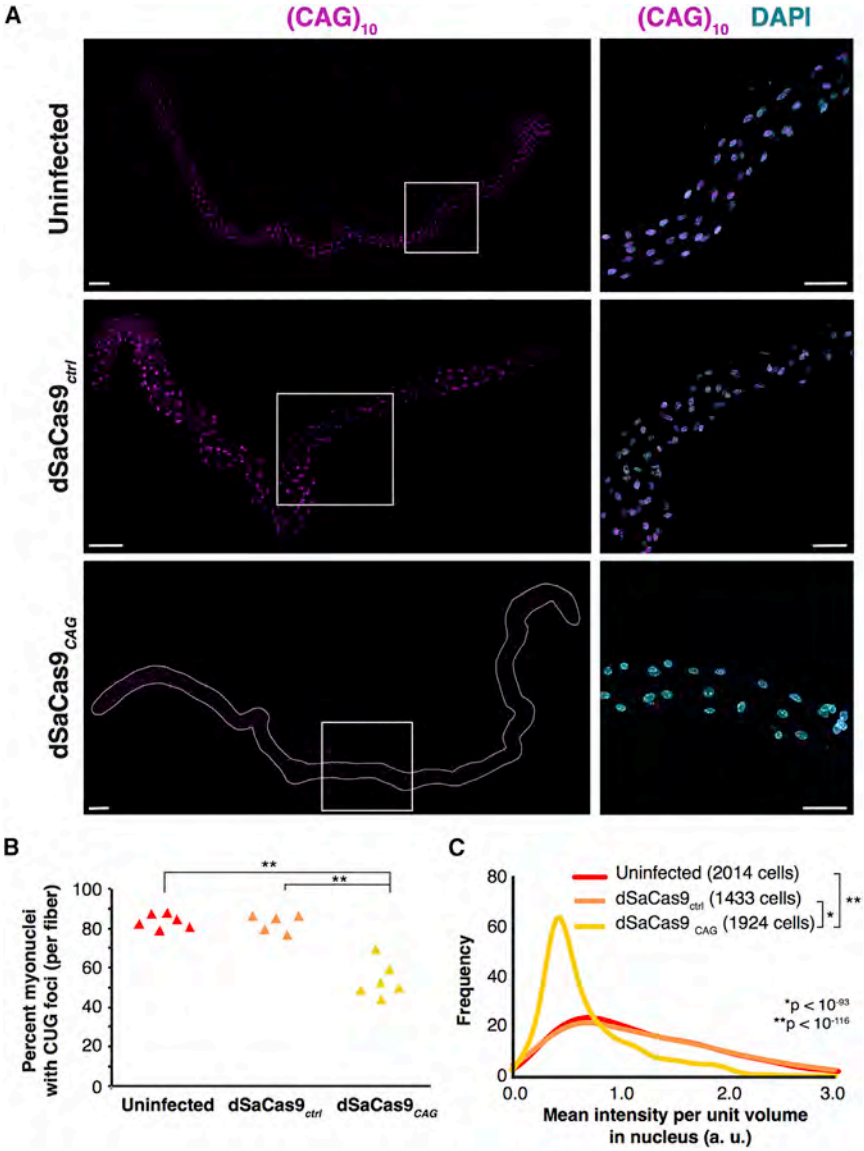


Figure 4. AAV-dSaCas9-$(CAG)_6$ Reduces RNA Foci in *HSA*^{LR} Muscle Fibers

(A) FISH to detect nuclear RNA foci (magenta) in myonuclei of *HSA*^{LR} EDL muscle fibers that were untreated (top) or infected with AAV-dSaCas9-control gRNA (middle) or $(CAG)_6$ gRNA (bottom). A representative fiber for each condition is shown (scale bar, 100 μm). Insets from each fiber (white boxes) are on the right (scale bar, 50 μm). DAPI is in cyan. Images are tile-scanned Z projections (see STAR Methods).

(B) The percentage of myonuclei per fiber showing RNA foci was quantitated across all three conditions with 5–6 fibers per condition, 400–500 nuclei per fiber (two-tailed t test, $^{**}p < 0.0005$).

(C) Probability density function of intensity of FISH signal in myonuclei from untreated fibers (red) or fibers infected with AAV-dSaCas9-control gRNA (orange) or AAV-dSaCas9-$(CAG)_6$ gRNA (yellow). A Kolmogorov-Smirnov test was performed for statistical significance.

splicing transcriptome-wide by RNA-seq and analyzed those exons in the myoblasts whose changes are concordant with changes observed in tibialis biopsies (lower left and upper right quadrants, Figure 3D) and whose baseline Ψ values are less than 0.33 units apart when comparing tibialis to myoblasts. Several exons were successfully rescued, and a correlation of ~0.66 was observed between dysregulation in tibialis (ΔΨ, DM1 minus unaffected) versus rescue by dSaCas9 and $(CAG)_6$ gRNA (ΔΨ, control gRNA minus $(CAG)_6$ gRNA) (Figure 3F; Table S3). To assess potential off-target changes in gene expression for other transcripts with genomic CTG or CAG repeat tracts, we treated unaffected myoblasts with AAV-dSaCas9 and $(CAG)_6$ gRNA or control gRNA and performed RNA-seq. We enumerated the longest contiguous CTG or CAG repeat tracts in all pre-mRNAs in the human genome and found none to exceed 24 repeats; maximums were 24 CTG in TCF4 and 22 in AR. We observed no differences in gene expression that depended on the length of repeat tracts for both CTG and CAG repeats (Figures S3C and S3D).

These observations suggest that dSaCas9 targeted to CTG repeat tracts can impede transcription of expanded CUG repeats, relieve MBNL sequestration, and restore splicing homeostasis in human cells containing DM1 repeat expansions, with selectivity that depends on the extreme repeat lengths commonly found in symptomatic tissue.

dCas9-Mediated Transcriptional Inhibition Reduces RNA Foci, Rescues Clcn1 Expression, and Decreases Myotonia in a Mouse Model of DM1

We next assessed whether dCas9 could impede transcription of expanded CTG repeats in a well-established mouse model of DM1, *HSA*^{LR} (Mankodi et al., 2000). These mice carry a human skeletal actin transgene containing 250 CTG repeats in the 3′ UTR and exhibit molecular, cellular, and phenotypic properties characteristic of DM1 patients. To assess efficacy of this approach independent of potential *in vivo* drug delivery challenges, we dissected extensor digitorum longus muscle (EDL) fibers and cultured them *ex vivo*. As CTG repeat expression in this model is driven by the *HSA* promoter, RNA foci are ubiquitous and numerous in myonuclei, showing fluorescence *in situ* hybridization (FISH) signal in ~80% of myonuclei (Figure 4A, top). This remained at similar levels following 2 days of AAV treatment with dSaCas9 and control gRNA but fell to ~50% with dSaCas9 and $(CAG)_6$ gRNA. (Figure 4B). FISH signal intensity, quantitated in >1,400 nuclei per condition (see STAR Methods; Figure S4), was dramatically decreased in fibers treated with dSaCas9-$(CAG)_6$ gRNA relative to control gRNA and untreated fibers (Figure 4C).

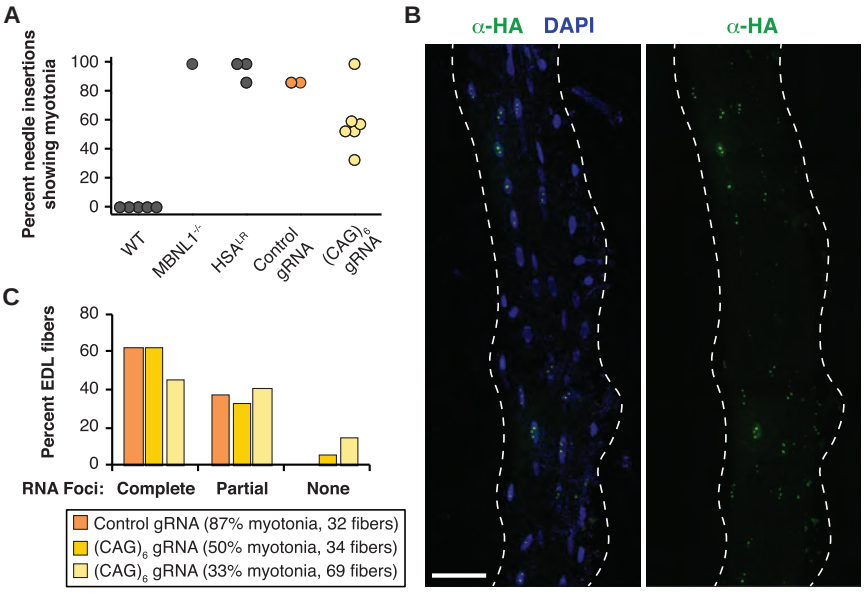


Figure 5. AAV6-dSaCas9-$(CAG)_6$ Rescues Muscle Phenotypes in HSA^{LR} Mice

(A) Percent needle insertions showing a myotonic run in wild-type FVB, $MBNL1^{-/-}$, HSA^{LR} (black), and HSA^{LR} treated with AAV6-dSaCas9-control gRNA (orange) or $(CAG)_6$ gRNA (yellow). Each point represents a muscle from a single animal, but in some animals, two muscles were assayed ($p < 0.027$ for AAV6-dSaCas9-control versus AAV6-dSaCas9-CAG6 treated animals, two-tailed t test).

(B) IF against HA-dSaCas9 (green) in an EDL fiber from a mouse treated with AAV6-dSaCas9-$(CAG)_6$ gRNA. DAPI is in blue. Scale bar, 50 μm.

(C) Quantitation of EDL fibers with complete, partial, or no CUG RNA foci upon treatment with AAV6-dSaCas9-control or $(CAG)_6$ gRNA. Myotonia levels and numbers of fibers analyzed are also listed.

(D) IF against Clcn1 in TA muscle sections from FVB, HSA^{LR}, and HSA^{LR} treated with AAV-dSaCas9-control or $(CAG)_6$ gRNA. DAPI is shown in blue. Scale bar, 50 μm.

These observations suggest that effective delivery of dSaCas9-$(CAG)_6$ to muscle fibers *ex vivo* is sufficient to significantly reduce RNA foci load within 2 days.

Having observed robust elimination of RNA foci *ex vivo*, we turned to *in vivo* experiments. Given previous reports of immune reactivity against Cas9, especially following intramuscular injection (Chew et al., 2016), we chose to administer AAV-dSaCas9-gRNA by temporal vein injection at postnatal day 2, prior to full establishment of immune tolerance. AAV2/6 or AAV2/9 carrying dSaCas9 and $(CAG)_6$ or control gRNA was injected. 5 weeks following injection, we performed electromyography to analyze myotonia in tibialis anterior and gastrocnemius muscles (Figure 5A). Wild-type FVB mice showed no myotonia, an *Mbnl1* knockout mouse showed myotonia in 100% of needle insertions, and HSA^{LR} mice showed myotonia in 87%–100% of insertions. HSA^{LR} injected with AAV6-dSaCas9-control gRNA showed myotonia levels similar to uninjected HSA^{LR} (87%). However, HSA^{LR} mice injected with AAV6-dSaCas9-$(CAG)_6$ gRNA showed a reduction in myotonia, with some showing myotonia in only 33%–50% of insertions. Interestingly, nearly all mice injected with AAV2/6 showed rescue while only one mouse injected with AAV9 showed rescue (Figure S4D).

Because mis-splicing of chloride channel 1 (*Clcn1*) is well established to mediate myotonia in DM1 and in HSA^{LR} mice, we measured splicing patterns of *Clcn1* mRNA in treated mice. Surprisingly, when analyzing RNA extracted from bulk muscle tissue, we could not detect significant rescue in *Clcn1* splicing or other MBNL-dependent splicing events (data not shown). Because myotonia is measured by assaying individual bundles of fibers, we hypothesized that some regions of muscle may be rescued and not others and that any potential change in isoform composition in these regions may be diluted in analyses of bulk tissue. This is consistent with a roughly ~2.5-fold increase in myonuclei per fiber over the first 4 weeks of post-natal development, which could also dilute the proportion of myonuclei containing AAV episomes (Duddy et al., 2015). AAV6-delivered dSaCas9 distribution in EDL fibers revealed mosaic expression, with rare fibers showing region-specific nuclear signal for dCas9 (Figure 5B), and most fibers showing absence of dCas9. Control immunofluorescence experiments against lamin A showed ubiquitous and consistent labeling of all myonuclei, ruling out potential staining artifacts (Figure S4E). In spite of mosaic dSaCas9 distribution, 5%–15% of fibers showed complete loss of CUG RNA foci (Figure 5C). Furthermore, extent of reduction in RNA foci correlated to strength of myotonia rescue across animals. Reduction of RNA foci in some nuclei raised the possibility that proper *Clcn1* splicing could be achieved in a subset of nuclei and that these mRNAs may spread locally throughout the fiber to produce functional Clcn1 protein. Indeed, mice treated with $(CAG)_6$ gRNA

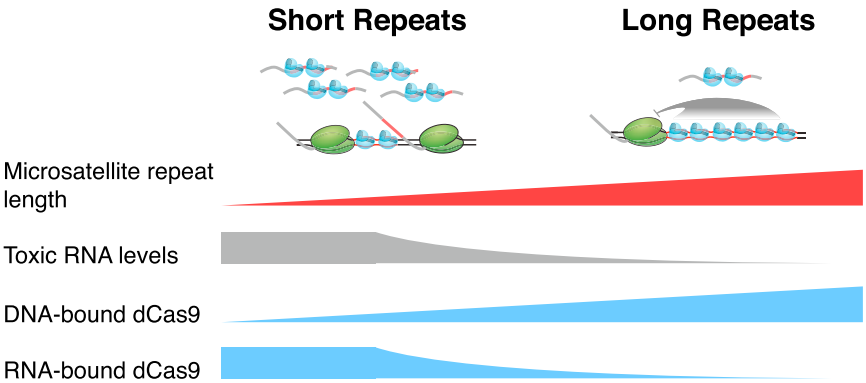


Figure 6. Model for Microsatellite Repeat Expansion Targeting by dCas9

dCas9 binding to DNA impedes transcription of long, expanded microsatellite repeats. In the context of shorter repeats, fewer copies of DNA-bound dCas9 may be insufficient to fully inhibit elongation by Pol II, permitting production of RNA that can also be targeted by dCas9.

showed increased Clcn1 staining at muscle membranes relative to mice treated with control gRNA (Figure 5D), but Clcn1 protein was localized only to a subset of fibers per muscle section. These observations are consistent with observations of myotonia elimination in a subset of fibers, yet absence of splicing rescue when assessing bulk tissue. In summary, these experiments suggest that dCas9 can, in principle, rescue disease phenotypes via transcriptional repression but that widespread rescue of molecular events in muscle will require efficient delivery to a large proportion of myonuclei expressing toxic RNA.

DISCUSSION

CRISPR/Cas9 has been previously applied to microsatellite expansion diseases, to remove expanded repeat tracts (van Agtmaal et al., 2017), or to cause repeat contraction (Cinesi et al., 2016). Here, we have applied a deactivated form of Cas9 to these diseases. By testing disease-associated repeat sequences across multiple repeat lengths *in vitro* and in disease models, we demonstrate that dCas9 can substantially reduce repeat-containing transcript abundance. Importantly, repression efficiency is proportional to repeat length, because longer repeats recruit an increased number of dCas9/gRNA complexes, leading to greater transcriptional blockade (Figures 1B and 1C).

dCas9 is well established to target DNA (Qi et al., 2013) but has also been shown to bind repeat-containing RNA in a gRNA-dependent manner (Batra et al., 2017). Indeed, when we immunoprecipitated dCas9 and assessed RNA binding, we observed increased binding to transcripts containing longer repeats (Figure S2G). However, as repeat length increases, the amount of successfully transcribed RNA decreases, suggesting that at long repeat lengths, less RNA remains to be targeted. In the context of disease, microsatellite repeat tracts undergo dramatic somatic expansion in DM1, DM2, C9ALS/FTD, and HD, and evidence suggests that repeat lengths in tissues of symptomatic individuals reach thousands of nucleotides in length (Kennedy et al., 2003; Thornton et al., 1994). Therefore, DNA-based targeting by dCas9 may play a primary role in driving potential therapeutic benefits in individuals with full expansions (Figure 6). At short and intermediate repeat lengths, therapeutic benefit may derive from both DNA- and RNA-based targeting by dCas9 to not only impede transcription of repeat-containing DNA, but also modulate downstream consequences of transcribed "escaper" RNAs. Taken together, we demonstrate that dCas9 is an effective tool to limit downstream effects of toxic RNAs.

While this approach mitigates a key safety concern of CRISPR, which is the unwanted cleavage of off-targets, it raises questions about whether long-term exposure to dCas9 can be achieved and tolerated. Effective, safe delivery in a multi-systemic fashion to sufficient numbers of post-mitotic cells will be necessary for this approach to be therapeutically viable. In DM and C9ALS/FTD, it is unclear what fraction of nuclei must be rescued in muscle or central nervous system to halt or improve disease symptoms. Here, we observed that rescue of a subset of nuclei was sufficient to yield partial rescue of myotonia. However, myotonia can be modeled as a loss-of-function event in which Clcn1 protein is lost, and there is evidence that in other recessive muscle diseases, gene editing of a small subset of nuclei can restore sufficient protein expression across muscle fibers to yield therapeutic benefit (Kemaladewi et al., 2017; Tabebordbar et al., 2016). Conversely, mis-splicing events yielding pathogenic isoforms with dominant-negative behavior may require full elimination to mitigate deleterious consequences.

Our approach used systemically delivered AAV, which revealed regional variation in transduction efficiency, likely because NLS-tagged Cas9 protein remained restricted to myonuclear domains transduced by AAV episomes, and many new myonuclei are recruited to fibers throughout post-natal muscle development (Duddy et al., 2015). Previous studies with AAV-mediated gene therapy typically evaluate expression of proteins that can spread throughout muscle fibers, precluding measurement of transduction efficiency to all myonuclei. It is possible that localized delivery would have facilitated transduction of a higher proportion of myonuclei. Nevertheless, reduction of toxic RNA repeat load in a subset of nuclei was sufficient to increase production of Clcn1 protein, with consequences for distant regions of muscle potentially naive to dCas9. These results emphasize how differences in mechanism of action and delivery efficiency between various therapeutic approaches should significantly influence how we interpret changes in response to those therapies, both at the molecular and phenotypic level.

In summary, the ability of dCas9 to impede transcription of microsatellite expansions defines a window in which transcriptional inhibition of expanded repeats is feasible yet does not interfere with that of typical RNAs. This approach may serve as a baseline with which to compare alternative therapeutic approaches, as well as a tool to identify mechanisms and principles by which

transcriptional activity of RNA polymerase II can be modulated in a sequence-specific manner.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Lines
  - Animals
- METHOD DETAILS
  - Cell Culture, Cell Lines, and Transfection
  - Cloning Barcoded Repeats and gRNA Plasmids
  - Generating Barcoded Repeat Expressing HeLas
  - dCas9 Chromatin and RNA IP
  - FISH and Immunofluorescence Analyses
  - Western Analyses for RAN Peptides
  - Recombinant AAV Production
  - Muscle Fiber Isolation
  - Transduction of Myoblasts and Muscle Fibers
  - Quantitating Signal Intensity of Nuclear Foci
  - Analysis of RNA-Seq Data
  - Electromyography
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILTY

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, three tables, and one datafile and can be found with this article online at https://doi.org/10.1016/j.molcel.2017.09.033.

AUTHOR CONTRIBUTIONS

Conceptualization, T.S., B.S.P., and E.T.W.; Methodology, T.S., B.S.P., and E.T.W.; Software, T.S. and E.T.W.; Formal Analysis, T.S., B.S.P., and E.T.W.; Investigation, T.S., B.S.P., R.O., J.D.C., H.R.M.-G., L.T.D., O.M., J.A., and E.T.W.; Initiated Project, T.S. and E.T.W.; Resources, H.R.M.-G., J.D.C., G.X., R.O., and M.S.S.; Writing – Original Draft, T.S., B.S.P., and E.T.W.; Writing – Review & Editing, T.S., B.S.P., E.T.W., M.S.S., H.R.M.-G., G.X., and R.O.; Visualization, T.S., B.S.P., and E.T.W.; Funding Acquisition, E.T.W.; Supervision, E.T.W.

ACKNOWLEDGMENTS

We would like to thank L. Ranum, T. Zu, and J. Cleary for generously providing us with the plasmids to assess RAN translation, C. Chamberlain for assisting with EMG, G. Aslanidi for providing the AAV6 capsid plasmid, and J. Singh for advice on AAV production. We thank T. Wang for construction of the $CCTG_{240}$ repeat tract. We thank C. Thornton for RNA samples from DM1 patients, which were obtained from the BioBank of the University of Rochester Wellstone Center (NIH NS048843). E.T.W. and T.S. were supported by DP5 OD017865. L.T.D. was supported by a Thomas H. Maren Fellowship.

Received: April 13, 2017
Revised: September 5, 2017
Accepted: September 22, 2017
Published: October 19, 2017

REFERENCES

Batra, R., Nelles, D.A., Pirie, E., Blue, S.M., Marina, R.J., Wang, H., Chaim, I.A., Thomas, J.D., Zhang, N., Nguyen, V., et al. (2017). Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. Cell *170*, 899–912.e10.

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

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., and Huang, B. (2013). Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell *155*, 1479–1491.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. *23*, 1163–1171.

Chew, W.L., Tabebordbar, M., Cheng, J.K.W., Mali, P., Wu, E.Y., Ng, A.H.M., Zhu, K., Wagers, A.J., and Church, G.M. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. Nat. Methods *13*, 868–874.

Cinesi, C., Aeschbach, L., Yang, B., and Dion, V. (2016). Contracting CAG/CTG repeats using the CRISPR-Cas9 nickase. Nat. Commun. *7*, 13272.

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

Duddy, W., Duguez, S., Johnston, H., Cohen, T.V., Phadke, A., Gordish-Dressman, H., Nagaraju, K., Gnocchi, V., Low, S., and Partridge, T. (2015). Muscular dystrophy in the mdx mouse is a severe myopathy compounded by hypotrophy, hypertrophy and hyperplasia. Skelet. Muscle *5*, 16.

François, V., Klein, A.F., Beley, C., Jollet, A., Lemercier, C., Garcia, L., and Furling, D. (2011). Selective silencing of mutated mRNAs in DM1 by using modified hU7-snRNAs. Nat. Struct. Mol. Biol. *18*, 85–87.

Furling, D., Doucet, G., Langlois, M.-A., Timchenko, L., Belanger, E., Cossette, L., and Puymirat, J. (2003). Viral vector producing antisense RNA restores myotonic dystrophy myoblast functions. Gene Ther. *10*, 795–802.

Gendron, T.F., Belzil, V.V., Zhang, Y.-J., and Petrucelli, L. (2014). Mechanisms of toxicity in C9FTLD/ALS. Acta Neuropathol. *127*, 359–376.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell *154*, 442–451.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-scale CRISPR-mediated control of gene repression and activation. Cell *159*, 647–661.

Gombash Lampe, S.E., Kaspar, B.K., and Foust, K.D. (2014). Intravenous injections in neonatal mice. J. Vis. Exp. Published online November 11, 2014. https://doi.org/10.3791/52037.

Harper, S.Q., Staber, P.D., He, X., Eliason, S.L., Martins, I.H., Mao, Q., Yang, L., Kotin, R.M., Paulson, H.L., and Davidson, B.L. (2005). RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. Proc. Natl. Acad. Sci. USA *102*, 5820–5825.

Ho, T.H., Charlet-B, N., Poulos, M.G., Singh, G., Swanson, M.S., and Cooper, T.A. (2004). Muscleblind proteins regulate alternative splicing. EMBO J. *23*, 3103–3112.

Ho, T.H., Savkur, R.S., Poulos, M.G., Mancini, M.A., Swanson, M.S., and Cooper, T.A. (2005). Colocalization of muscleblind with RNA foci is separable from mis-regulation of alternative splicing in myotonic dystrophy. J. Cell Sci. *118*, 2923–2933.

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

Kemaladewi, D.U., Maino, E., Hyatt, E., Hou, H., Ding, M., Place, K.M., Zhu, X., Bassi, P., Baghestani, Z., Deshwar, A.G., et al. (2017). Correction of a splicing defect in a mouse model of congenital muscular dystrophy type 1A using a homology-directed-repair-independent mechanism. Nat. Med. *23*, 984–989.

Kennedy, L., Evans, E., Chen, C.-M., Craven, L., Detloff, P.J., Ennis, M., and Shelbourne, P.F. (2003). Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. Hum. Mol. Genet. *12*, 3359–3367.

Konermann, S., Brigham, M.D., Trevino, A., Hsu, P.D., Heidenreich, M., Cong, L., Platt, R.J., Scott, D.A., Church, G.M., and Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. Nature *500*, 472–476.

Kramer, N.J., Carlomagno, Y., Zhang, Y.-J., Almeida, S., Cook, C.N., Gendron, T.F., Prudencio, M., Van Blitterswijk, M., Belzil, V., Couthouis, J., et al. (2016). Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. Science *353*, 708–712.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. *32*, 677–683.

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H.-R., Jiang, J., Watt, A.T., Chun, S., Katz, M., et al. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc. Natl. Acad. Sci. USA *110*, E4530–E4539.

Langlois, M.-A., Lee, N.S., Rossi, J.J., and Puymirat, J. (2003). Hammerhead ribozyme-mediated destruction of nuclear foci in myotonic dystrophy myoblasts. Mol. Ther. *7*, 670–680.

Lawhorn, I.E.B., Ferreira, J.P., and Wang, C.L. (2014). Evaluation of sgRNA target sites for CRISPR-mediated repression of TP53. PLoS ONE *9*, e113232.

Liu, C.-R., Chang, C.-R., Chern, Y., Wang, T.-H., Hsieh, W.-C., Shen, W.-C., Chang, C.-Y., Chu, I.-C., Deng, N., Cohen, S.N., and Cheng, T.-H. (2012). Spt4 is selectively required for transcription of extended trinucleotide repeats. Cell *148*, 690–701.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

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

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

Nelson, D.L., Orr, H.T., and Warren, S.T. (2013). The unstable repeats–three evolving faces of neurological disease. Neuron *77*, 825–843.

Nishimasu, H., Cong, L., Yan, W.X., Ran, F.A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., and Nureki, O. (2015). Crystal structure of Staphylococcus aureus Cas9. Cell *162*, 1113–1126.

Parkesh, R., Childs-Disney, J.L., Nakamori, M., Kumar, A., Wang, E., Wang, T., Hoskins, J., Tran, T., Housman, D., Thornton, C.A., and Disney, M.D. (2012). Design of a bioactive small molecule that targets the myotonic dystrophy type 1 RNA via an RNA motif-ligand database and chemical similarity searching. J. Am. Chem. Soc. *134*, 4731–4742.

Pasut, A., Jones, A.E., and Rudnicki, M.A. (2013). Isolation and culture of individual myofibers and their satellite cells from adult skeletal muscle. J. Vis. Exp. Published online March 22, 2013. https://doi.org/10.3791/50074.

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

Piedra, J., Ontiveros, M., Miravet, S., Penalva, C., Monfar, M., and Chillon, M. (2015). Development of a rapid, robust, and universal picogreen-based method to titer adeno-associated vectors. Hum. Gene Ther. Methods *26*, 35–42.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173–1183.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Rutledge, E.A., Halbert, C.L., and Russell, D.W. (1998). Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. J. Virol. *72*, 309–319.

Rzuczek, S.G., Colgan, L.A., Nakai, Y., Cameron, M.D., Furling, D., Yasuda, R., and Disney, M.D. (2017). Precise small-molecule recognition of a toxic CUG RNA repeat expansion. Nat. Chem. Biol. *13*, 188–193.

Sah, D.W.Y., and Aronin, N. (2011). Oligonucleotide therapeutic approaches for Huntington disease. J. Clin. Invest. *121*, 500–507.

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

Siboni, R.B., Nakamori, M., Wagner, S.D., Struck, A.J., Coonrod, L.A., Harriott, S.A., Cass, D.M., Tanner, M.K., and Berglund, J.A. (2015). Actinomycin D specifically reduces expanded CUG repeat RNA in myotonic dystrophy models. Cell Rep. *13*, 2386–2394.

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407–411.

Thornton, C.A., Johnson, K., and Moxley, R.T., 3rd (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. Ann. Neurol. *35*, 104–107.

van Agtmaal, E.L., André, L.M., Willemse, M., Cumming, S.A., van Kessel, I.D.G., van den Broek, W.J.A.A., Gourdon, G., Furling, D., Mouly, V., Monckton, D.G., et al. (2017). CRISPR/Cas9-induced (CTG·CAG)n repeat instability in the myotonic dystrophy type 1 locus: implications for therapeutic genome editing. Mol. Ther. *25*, 24–43.

Wagner, S.D., Struck, A.J., Gupta, R., Farnsworth, D.R., Mahady, A.E., Eichinger, K., Thornton, C.A., Wang, E.T., and Berglund, J.A. (2016). Dose-dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. PLoS Genet. *12*, e1006316.

Wang, E.T., Cody, N.A.L., Jog, S., Biancolella, M., Wang, T.T., Treacy, D.J., Luo, S., Schroth, G.P., Housman, D.E., Reddy, S., et al. (2012). Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. Cell *150*, 710–724.

Wang, E.T., Ward, A.J., Cherone, J.M., Giudice, J., Wang, T.T., Treacy, D.J., Lambert, N.J., Freese, P., Saxena, T., Cooper, T.A., and Burge, C.B. (2015). Antagonistic regulation of mRNA expression and splicing by CELF and MBNL proteins. Genome Res. *25*, 858–871.

Wheeler, T.M., Leger, A.J., Pandey, S.K., MacLeod, A.R., Nakamori, M., Cheng, S.H., Wentworth, B.M., Bennett, C.F., and Thornton, C.A. (2012). Targeting nuclear RNA for in vivo correction of myotonic dystrophy. Nature *488*, 111–115.

Xia, G., Santostefano, K.E., Goodwin, M., Liu, J., Subramony, S.H., Swanson, M.S., Terada, N., and Ashizawa, T. (2013). Generation of neural cells from DM1 induced pluripotent stem cells as cellular model for the study of central nervous system neuropathogenesis. Cell. Reprogram. *15*, 166–177.

Xiao, X., Li, J., and Samulski, R.J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. J. Virol. *72*, 2224–2232.

Zolotukhin, S., Potter, M., Zolotukhin, I., Sakai, Y., Loiler, S., Fraites, T.J., Jr., Chiodo, V.A., Phillipsberg, T., Muzyczka, N., Hauswirth, W.W., et al. (2002). Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. Methods *28*, 158–167.

Zu, T., Gibbens, B., Doty, N.S., Gomes-Pereira, M., Huguet, A., Stone, M.D., Margolis, J., Peterson, M., Markowski, T.W., Ingram, M.A.C., et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc. Natl. Acad. Sci. USA *108*, 260–265.

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T.M., Harms, M.B., Falchook, A.E., Subramony, S.H., et al. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc. Natl. Acad. Sci. USA *110*, E4968–E4977.

Zu, T., Cleary, J.D., Liu, Y., Bañez-Coronel, M., Bubenik, J.L., Ayhan, F., Ashizawa, T., Xia, G., Clark, H.B., Yachnis, A.T., et al. (2017). RAN translation regulated by muscleblind proteins in myotonic dystrophy type 2. Neuron *95*, 1292–1305.e5.

Please cite this article in press as: Pinto et al., Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9, Molecular Cell (2017), https://doi.org/10.1016/j.molcel.2017.09.033

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-HA tag	Cell Signaling Technologies	Cat# 3724; RRID: AB_1549585
Rabbit anti-Clcn1	Alpha Diagnostic International	Cat#CLC11-S; RRID: AB_2079253
Critical Commercial Assays		
KAPA Stranded RNA-Seq Library Preparation Kit	Kapa Biosystems	Cat# KK8401
Quant-iT Picogreen dsDNA assay kit	Life Technologies	Cat# P11496
NEBNext Ultra II Directional RNA Library Prep with Sample Purification Beads	New England Biolabs	E7765S
NEBNext rRNA Depletion Kit (Human/Mouse/Rat)	New England Biolabs	E6310S
Deposited Data		
Raw and analyzed data	This paper	GEO: GSE103997
Experimental Models: Cell Lines		
Human DM1 primary myoblast cells	Provided by Guangbin Xia	N/A
Experimental Models: Organisms/Strains		
Mouse: *HSA*^{LR}: FVB-*HSA*^{LR}	N/A	N/A
Oligonucleotides		
PCR primers for amplicon library generation and sequencing, see Figure S1	This paper	N/A
CalFluor 610 conjugated $(CAG)_{10}$ probe	This paper, Biosearch Technologies	N/A
Recombinant DNA		
Plasmid: CTG_0	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80418
Plasmid: CTG_{12}	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80416
Plasmid: CTG_{40}	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80415
Plasmid: CTG_{240}	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80414
Plasmid: CTG_{480}	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80413
Plasmid: CTG_{960}	Gift from Tom Cooper; Ho et al., 2004	Addgene Cat#80412
Plasmid: CAG_{960}	Gift from Tom Cooper	CAG_{960}
Plasmid: CTG_0 + barcode	This paper	N/A
Plasmid: CTG_{12} + barcode	This paper	N/A
Plasmid: CTG_{40} + barcode	This paper	N/A
Plasmid: CTG_{240} + barcode	This paper	N/A
Plasmid: CTG_{480} + barcode	This paper	N/A
Plasmid: CTG_{960} + barcode	This paper	N/A
Plasmid: CAG_{960} + barcode	This paper	N/A
Plasmid: $CCTG_{240}$	This paper	N/A
Plasmid: $CAGG_{240}$	This paper	N/A
Plasmid: $CCTG_{240}$ + barcode	This paper	N/A
Plasmid: $CAGG_{240}$ + barcode	This paper	N/A
Plasmid: pXdCas9	Gift from Albert Cheng; Cheng et al., 2013	pXdCas9
Plasmid: pCR-Blunt II-TOPO- U6-gRNA	Addgene	Cat #41824
Plasmid: pCR-Blunt II-TOPO- U6-$(CAG)_6$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(AGC)_6$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(GCA)_6$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(CTG)_6$ gRNA	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Plasmid: pCR-Blunt II-TOPO- U6-$(CAGG)_5$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(AGGC)_5$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(GCAG)_5$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(CCUG)_5$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(C_4G_2)_3$ gRNA	This paper	N/A
Plasmid: pCR-Blunt II-TOPO- U6-$(G_4C_2)_3$ gRNA	This paper	N/A
Plasmid: pAAV-CMV::NLS-dSaCas9-NLS-3xHA-bGHpA;U6::sgRNA	This paper	N/A
Plasmid: pAAV-CMV::NLS-dSaCas9-NLS-3xHA-bGHpA;U6::$(CAG)_6$ sgRNA	This paper	N/A
Plasmid: pAAV6 packaging plasmid	Gift from George Aslanidi	N/A
Plasmid: pXX6 helper plasmid	Gift from Nick Muzyczka	N/A
Plasmid: pcDNA-6XStop-$(CTG)_{150}$-3X(FLAG-HA-cMyc-His)	Gift from Laura Ranum; Zu et al., 2011	N/A
Plasmid: pcDNA-6XStop-$(CCTG)_{137}$-3X(FLAG-HA-cMyc-His)	Gift from Laura Ranum; Zu et al., 2017	N/A
Plasmid: pcDNA-6XStop-$(G_4C_2)_{120}$-3X(FLAG-HA-cMyc-His)	Gift from Laura Ranum; Zu et al., 2013	N/A
Plasmid: pAC156	Gift from Albert Cheng	N/A
Software and Algorithms		
Custom Python scripts	This paper	https://github.com/etwang/dCas9paper

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Eric T. Wang (eric.t.wang@ufl.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines

Human cell lines and primary cells were utilized in this study. Patient myoblast cell lines were derived from muscle biopsies under a University of Florida-approved IRB protocol with informed consent from all patients. Information on the genders of the patient cells is unavailable. All cells were grown under standard conditions of 37°C and 5% CO_2 in media supplemented with 10% fetal bovine serum.

Animals

In vivo animal studies were performed on *HSA*^{LR} transgenic mice in this study. Both males and females were used in EMG analysis, which was performed at 5 weeks of age, as by this time point mice display appreciable myotonia. All these studies were performed in accordance with guidelines and regulations of the Institutional Animal Care and Use Committee (IACUC) at the University of Florida.

METHOD DETAILS

Cell Culture, Cell Lines, and Transfection

HeLa and HEK293T cells were cultured in 1X DMEM supplemented with 10% fetal bovine serum and 1% Penicillin/Streptomycin at 37°C and 5% CO_2. The DM1 primary myoblast cells were obtained by G. Xia. The clinical data of research subjects from who muscle biopsies were obtained, have been described before (Xia et al., 2013). Muscle biopsies were performed using 7G UCH Muscle Biopsy Needle (#8066, http://www.cadencescience.com). Samples were cut into small pieces and seeded into a 6cm dish. Myoblasts were expanded at a 1:2 ratio in myoblast growth media (Skeletal Muscle Cell Growth Medium-2, Lonza, #CC 3245) at 37°C and 5% CO_2. For viral transductions, these primary myoblasts were maintained in the growth medium for 3 days and then switched to Differentiation media (DMEM-F12 supplemented with 2% horse serum) for 3 days.

Plasmid DNA transfections of HeLa and HEK293T cells were performed using Trans-IT LT1 (MirusBio) as per manufacturer's instructions. For RNA and FISH analyses, HeLa cells were harvested or processed 72 hours post transfection.

Cloning Barcoded Repeats and gRNA Plasmids

Non-barcoded plasmids carrying 0, 12, 40, 240, 480, 960 CTG repeats (CTG_n) and 960 CAG repeats (CAG_{960}) were kind gifts from Tom Cooper (Baylor College of Medicine). Plasmids with 0 and 12 repeats were modified so that their vector backbones were identical to the others, and carried the ampicillin resistance gene. The $(CCTG)_{240}$ plasmid was created as in Philips et al. (1998). In brief, oligonucleotide fragments 5′-TCGA$(CCTG)_{20}$C-3′ and 5′-TCGAG$(CAG)_{20}$-3′, were phosphorylated, annealed, gel isolated, and concatemerized by T4 DNA ligase. Concatemers not in a head to tail orientation were digested by SalI and XhoI. Concatemers were gel-isolated and cloned into the SalI site of CTG_0. To create $CAGG_{240}$, CTG_{960} was digested with SalI and HindIII and a new fragment was introduced which included HindIII and AgeI restriction sites (DT_MCS). $CCTG_{240}$ was digested with XmaI and HindIII and ligated with DT_MCS digested with HindIII and AgeI to reverse orientation of the repeats and form $CAGG_{240}$.

For MBTA-seq, each repeat-containing plasmid was barcoded by introducing a fragment containing a random 8-nt sequence at the PflMI restriction site located downstream of the repeats via In-Fusion cloning (Clontech). Clones were sequenced to confirm that each repeat containing plasmid carried a unique barcode (Figure S1A). Barcoded CTG_{960} and CAG_{960} plasmids were generated by digesting $(CTG)_{960}$ and $(CAG)_{960}$ with HindIII and AccIII and ligating gBlocks containing unique barcodes with the digested plasmids.

The dCas9 gRNAs were cloned into an AflII- digested U6 expression vector by annealing oligos as previously described (Mali et al., 2013). The dSaCas9 gRNAs were cloned into BsaI - digested vector by annealing oligos as previously described (Ran et al., 2015).

Generating Barcoded Repeat Expressing HeLas

To generate stable cell lines, the *DMPK* expression cassette was removed from each of the 6 barcoded plasmids and inserted into pAC156 (gift from Albert Cheng), a plasmid with Piggybac transposon terminal repeats as well as a puromycin selection cassette. All 6 plasmids were transiently transfected together with the Piggybac mPB transposase into HeLa, and selected by puromycin. Single cells were isolated by flow cytometry, and colonies were cultured in 96 well plates. 48 colonies were expanded and subjected to MBTA-seq to screen for integration and expression of all 6 plasmids. Counts for all barcode deep sequencing experiments can be found in Table S1.

dCas9 Chromatin and RNA IP

dCas9 ChIP and RIP experiments were performed on the HeLa cell line containing the six different repeat lengths (CTG0- 960) integrated into genome. 10cm plates of ~80% confluent cells were transfected with the pXdCas9 and U6 expression vectors expressing the dCas9 protein and the control or $(CAG)_6$ gRNAs respectively (Cheng et al., 2013; Mali et al., 2013), using the manufacturers guidelines. For ChIP, 48 hours after transfection cells were crosslinked with 1% formaldehyde for 20 minutes at room temperature. Crosslinking reactions were stopped by addition of glycine to a final concentration of 0.125 M. Cells were then harvested, washed with phosphate buffered saline and pelleted. 1 mL of Lysis Buffer 1 (50mM HEPES [pH 7.5], 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% Igepal, 0.25% Triton X-100) was added to the cells and rocked at 4°C for 10 minutes. After spinning, the cells were incubated in Lysis Buffer 2 (10mM Tris-HCl [pH 8.0], 200mM NaCl, 1mM EDTA, 0.5mM EGTA) for 10 minutes at RT. Nuclei were pelleted and resuspended in 1mL of Lysis Buffer 3 (10mM Tris-HCl [pH 8.0], 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% Na-deoxycholate and 0.5% N-lauroylsarcosine) and subjected to sonication in a Covaris S220 to obtain DNA fragments averaging 4kb in length. One-twentieth of the total chromatin served as input. The remaining material was used in the IP, which was performed using the HA-tag (C29F4) rabbit mAb conjugated to magnetic beads (Cell Signaling Technologies) at 4°C overnight to pull down the HA-tagged dSpCas9 and interacting DNA. Beads were washed 7 times with wash buffer (50mM HEPES [pH 7.6], 500mM LiCl, 1mM EDTA, 1% Igepal and 0.7% Na-Deoxycholate). Immunocomplexes were eluted from the beads with elution buffer (50mM Tris-HCl, 10mM EDTA and 1% SDS) at 65°C for 15 minutes. Crosslinks in the IP and input were reversed overnight at 65°C and treated with RNase A and proteinase K to remove RNAs and proteins. DNA was extracted with phenol-chloroform and precipitated with ethanol. The barcoded region associated with each repeat length was amplified from the isolated DNA fragments using primers containing adapters facilitating deep-sequencing.

For the RNA IP experiments, 48 hours after transfection, cells were harvested in 1mL of lysis buffer (100mM Kcl, 5 mM MgCl2, 10mM HEPES [pH 7.0], 0.5% Igepal, 1mM DTT, 100 U/mL SUPERase In RNase inhibitor (Thermo Fisher), 2mM vanadyl ribonucleoside complexes solution, 25uL/mL protease inhibitor cocktail). One-twentieth of the resulting lysate was used as input and the remaining lysate was incubated with HA-tag (C29F4) rabbit mAb conjugated to magnetic beads (Cell Signaling Technologies) at 4°C overnight to pull down the HA-tagged dSpCas9 and interacting RNA. Beads were washed with the lysis buffer four times at 4°C and immunocomplexes were eluted off with 0.1% SDS and proteins removed using proteinase K at 50°C for 30 minutes. RNA was isolated using the Direct-zol RNA miniprep kit (Zymo Research) and contaminating DNA was eliminated using TURBO DNase (Thermo Fisher). cDNA was generated using Superscript IV Reverse Transcriptase (Thermo Fisher) and subsequently barcoded regions were amplified using flanking primers carrying sequences suitable for deep sequencing.

FISH and Immunofluorescence Analyses

To detect nuclear RNA foci, cells or muscle fibers were fixed with 4% PFA for 10 minutes at room temperature followed by ice cold RNase free 70% ethanol for 30 mins. Fixed samples were washed with a 25% formamide wash buffer at 30°C for 30 mins and then hybridized with a CalFluor 610 conjugated $(CAG)_{10}$ oligonucleotide (Biosearch Technologies) in a 25% formamide hybridization buffer overnight at 30°C. Finally samples were washed two times at 30°C with wash buffer for 30 minutes to an hour, incubated with DAPI

(1mg/mL) and mounted in Vectashield. Further IF analysis was performed on HeLa cells and DM1 myoblasts to detect the presence of the dCas9-HA protein. After excess oligonucleotide was washed off, the cells were blocked in 3% normal goat serum in 1% Triton X-100-PBS for 30 mins at room temperature, incubated with anti-HA antibody (1:500, #3724, Cell Signaling Technologies) overnight at 4°C, washed with 1X PBS, incubated with Alexa Fluor 488 conjugated anti-rabbit secondary antibody (1:500, Life Technologies) for 2 hours, washed and incubated with DAPI for 5 mins and mounted in Vectashield. Slides were imaged using the Zeiss LSM 880 Confocal Laser Scanning Microscope.

IF analysis to detect dCas9 in mouse muscle fibers from the EDL of mice injected with AAV6-dSaCas9-$(CAG)_6$ was conducted as described above, except samples were fixed with 100% isopropanol at −20°C for 10 mins.

Clcn1 was detected in mouse muscle by performing IF on frozen muscle sections of the TA. Frozen muscle was sectioned into 10uM slices, fixed with 100% acetone at −20°C for 20 minutes, washed with 0.3% Triton X-100-PBS and incubated with rabbit anti Clcn1 (1:100, #CLC11-S, Alpha Diagnostic International) overnight at 4°C. Samples were incubated with goat anti-rabbit Alexa Fluor 568 (1:500, Thermo Fisher) for 2 hours at RT and then treated with DAPI and mounted in Vectashield.

Western Analyses for RAN Peptides

HEK293Tcells were transfected with one of the following plasmids that express tagged RAN translated products: pcDNA-6XStop-$(CTG)_{150}$-3X(FLAG-HA-cMyc-His), pcDNA-6XStop-$(CCTG)_{137}$-3X(FLAG-HA-cMyc-His) or pcDNA-6XStop-$(G_4C_2)_{120}$-3X(FLAG-HA-cMyc-His) (Zu et al., 2011, 2013, 2017). These cells were co-transfected with the pXdCas9 plasmid expressing the dCas9 protein (Cheng et al., 2013), and U6 expression vectors (Mali et al., 2013) expressing the control gRNA or gRNA's targeting either strand of the CTG, CCTG and G_4C_2 repeats ($(CAG)_6$ or $(CUG)_6$ gRNAs, $(CAGG)_5$, $(AGGC)_5$, $(GCAG)_5$,or $(CCUG)_5$ gRNAs and $(C_4G_2)_3$ or $(G_4C_2)_3$ gRNAs, respectively).

72 hours after transfection, cells in each well of a 12-well tissue culture plate were gently rinsed 1X with PBS and lysed in 200ul of RIPA Buffer (50 mM Tris-Cl pH 7.4, 150 mM NaCl, 0.1% Na-Deoxycholate, 1% NP-40, 0.5% SDS) with protease inhibitors for 30 minutes on ice. Genomic DNA was sheared by 8-10 passages through a 21-gauge needle. The resulting lysate was centrifuged at 18,000 x g for 15 min and the supernatant was collected. The protein concentration of the lysate was determined using Pierce BCA Protein Assay. Equal amounts of protein were loaded and separated on a 4%–12% NuPage Bis-Tris gel (Novex) and transferred to a nitrocellulose membrane (Amersham). The membrane was blocked in 5% milk in PBS-Tween20 (0.05%) for 1 hour and probed with anti-FLAG (1:2000) or anti-HA (1:1000) antibody in 1% milk solution in PBS-Tween20 (0.05%) overnight at 4°C. After the membrane was incubated with anti-mouse and anti-rabbit HRP (1:10,000) for 2 hours at room temperature, the bands were detected using the SuperSignal West Femto Maximum Sensitivity Substrate as per manufacturers protocol (Zu et al., 2011).

Recombinant AAV Production

Viral production was achieved through transfection of HEK293T cells cultured in 150mm plates with the pAAV6 serotype packaging plasmid (Rutledge et al., 1998), pXX6 helper plasmid that contains the adenovirus E4, VA and E2a helper regions (Xiao et al., 1998) and AAV2-ITR containing plasmid expressing dSaCas9 and the control or $(CAG)_6$ gRNA (generated from the SaCas9 plasmid pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA, Addgene #61591). Transfections were carried out using the TransIT-LT1 transfection reagent and recommended protocols (MirusBio). Cells were harvested between 48h and 72h post-transfection, recombinant AAV2/6 virus was purified by iodixanol step gradients followed by vector concentration and buffer exchange with lactated Ringer's in an Apollo 150kDa concentrator (Orbital Biosciences) (Zolotukhin et al., 2002). Virus titers were determined using the Quant-iT Picogreen dsDNA assay kit (Life Technologies) (Piedra et al., 2015) and found to be $\sim 10^{11}$ vg/mL.

Muscle Fiber Isolation

Single muscle fibers were isolated from 3-4 week old HSA^{LR} mice as described previously (Pasut et al., 2013). Briefly, the EDL was dissected and digested with a 0.2% Collagenase Type I in DMEM solution in a 37°C water bath for 1 hour without agitation. The digested muscle was then flushed with DMEM to separate out individual muscle fibers. Fibers were cultured in DMEM containing 20% FBS overnight before infection with AAV.

Transduction of Myoblasts and Muscle Fibers

Virus carrying dSaCas9 and control or $(CAG)_6$ gRNA was used to infect human DM1 primary myoblast cell lines and HSA^{LR} mouse EDL muscle fibers. To determine whether blocking expression of the CTG repeats in the human myoblasts affected the presence of RNA foci and splicing of MBNL targets, cells were grown to 60% confluency on CC2 chamber slides and infected for 6 days (3 days in growth media plus 3 days in differentiation media) with viral titers of 10^9. To analyze the effects on RNA foci in HSA^{LR} muscle fibers, 10-20 muscle fibers were cultured in wells of a 96 well plate and infected with a 10^9 viral titer for 48h.

Quantitating Signal Intensity of Nuclear Foci

Python scripts were written to quantitate intensity of FISH signal from RNA foci within nuclei of muscle fibers (python functions are listed below, and also see Figure S4). Regions of interest were defined across each tile-scanned z stack image obtained by confocal imaging. To identify and segment nuclei, 16-bit intensity values were scaled to lie between 0 and 1 (skimage.img_as_float). Then, a Gaussian filter of standard deviation 3 pixels was applied to the raw DAPI signal (skimage.filters.Gaussian), and a threshold of

mean + 5 * standard deviation of the Gaussian filtered signal was used to generate a binary mask. Holes were filled (scipy.ndimage.fill_holes), and a binary opening operation was performed to remove salt noise (skimage.morphology.binary_opening with ball structured element of radius 2). Nuclei were segmented and labeled (skimage.measure.label), and objects with pixel volume < 1000 or > 20000 were removed. FISH signal was scaled to lie between 0 and 1 as above, and a grayscale opening operation was performed to measure background (skimage.morphology.opening with a ball structured element of radius 3). This background intensity multipled by 3 was subtracted from the FISH signal to yield background-subtracted FISH signal. The binary nucleus mask was applied to this signal, and total intensity was measured within each nucleus, and divided by the nuclear volume in pixels, to obtain the final FISH signal per unit volume for each nucleus. This procedure was applied to all regions of interest across all fibers.

Analysis of RNA-Seq Data

100 ng of RNA was used to prepare RNA-seq libraries using the KAPA Ribo-Erase Strand-Specific kit. Samples were pooled and sequenced on the NextSeq 500 Version 2, using a High-Output 2x75 kit. Reads were mapped to hg19 by Hisat2, and splicing events were quantitated by MISO. Ψ values from DM tibialis biopsies were fit to sigmoid curves using 4-parameter estimation, where $\Psi = \Psi_{min} + (\Psi_{max} - \Psi_{min}) / (1 + e^{-slope * ([MBNL]inferred - EC_{50})})$, using python/scipy packages. The $[MBNL]_{inferred}$ value was taken from Wagner et al. (2016). The "fit error" was evaluated by taking the sum of squared errors between observed Ψ and Ψ as predicted by the sigmoid curves. Events consistently regulated between non-DM1 and DM1 myoblasts were identified using a modified monotonicity test (Wang et al., 2015), $\Delta\Psi > 0.1$, BF > 5) where the 2 non-DM1 libraries were grouped together, and 4 DM1 libraries were grouped together. For Figure 5B, events with < 1.3 sigmoid fit error and > 1 monotonicity Z-score were selected for display. For Figure 5F, only events identified in Figure 5B to lie in the upper right or lower left quadrants were further analyzed; in addition, events were required to exhibit < 0.33 difference in Ψ between cells treated with AAV-dSaCas9-control gRNA and non-DM1 tibialis biopsies. Raw RNA-seq reads for these libraries are publicly available (GEO accession number pending). Statistics on read coverage are in Table S2.

Electromyography

To determine whether expression of dCas9-$(CAG)_6$ rescued myotonia in the *HSA*^{LR} mice, mice were injected with AAV6-dSaCas9 and control or $(CAG)_6$ gRNA at 10^{10} viral genomes per mouse via the temporal vein at P2 (Gombash Lampe et al., 2014). Myotonia was assessed by electromyography (EMG) at 5 weeks of age as described previously (Kanadia et al., 2003). EMG was performed under general anesthesia (intraperitoneal ketamine, 100 mg/kg; xylazine, 10 mg/kg) using 30 gauge concentric needle electrodes with at least 15 needle insertions per muscle in the hindlimb muscles, gastrocnemius and TA. Myotonic discharges were denoted as a percentage of the total number of insertions. In Figure 5A, each point represents one of the hindlimb muscles from a single animal, but in some animals, both muscles were tested. N = 2 mice for control gRNA and N = 4 mice for $(CAG)_6$ gRNA.

QUANTIFICATION AND STATISTICAL ANALYSIS

Custom Python and R scripts were written to analyze data in this manuscript. Read counts for all MBTA-seq experiments were obtained by analysis of raw fastQ sequence. Statistical tests and numbers of cells/animals are described in Supplemental tables, figure legends, and text.

DATA AND SOFTWARE AVAILABILTY

The accession number for the data reported in this paper is GEO: GSE103997. Custom Python scripts can be found at: https://github.com/etwang/dCas9paper.