UC San Diego
UC San Diego Previously Published Works

Title

Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9

Permalink

Journal

Cell, 170(5)

ISSN

0092-8674

Authors

Batra, Ranjan
Nelles, David A
Pirie, Elaine
et al.

Publication Date

2017-08-01

DOI

10.1016/j.cell.2017.07.010

Peer reviewed

Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9

Ranjan Batra[1,2,3,10], **David A. Nelles**[1,2,3,10], **Elaine Pirie**[1,2,3], **Steven M. Blue**[1,2,3], **Ryan J. Marina**[1,2,3], **Harrison Wang**[1,2,3], **Isaac A. Chaim**[1,2,3], **James D. Thomas**[4], **Nigel Zhang**[1,2,3], **Vu Nguyen**[1,2,3], **Stefan Aigner**[1,2,3], **Sebastian Markmiller**[1,2,3], **Guangbin Xia**[5], **Kevin D. Corbett**[1,6,7], **Maurice S. Swanson**[4], and **Gene W. Yeo**[1,2,3,8,9,11,*]

[1]Department of Cellular and Molecular Medicine, University of California at San Diego, La Jolla, CA, USA

[2]Stem Cell Program, University of California at San Diego, La Jolla, CA, USA

[3]Institute for Genomic Medicine, University of California at San Diego, La Jolla, CA, USA

[4]Department of Molecular Genetics and Microbiology, Center for NeuroGenetics and the Genetics Institute, University of Florida, College of Medicine, Gainesville, FL, USA

[5]Department of Neurology, University of Florida, College of Medicine, Gainesville, FL, USA

[6]Ludwig Institute for Cancer Research, San Diego Branch, La Jolla, CA, USA

[7]Department of Chemistry, University of California, San Diego, La Jolla, CA, USA

[8]Molecular Engineering Laboratory, A[*]STAR, Singapore, Singapore

[9]Department of Physiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

SUMMARY

Microsatellite repeat expansions in DNA produce pathogenic RNA species that cause dominantly inherited diseases such as myotonic dystrophy type 1 and 2 (DM1/2), Huntington's disease, and *C9orf72*-linked amyotrophic lateral sclerosis (C9-ALS). Means to target these repetitive RNAs are required for diagnostic and therapeutic purposes. Here, we describe the development of a programmable CRISPR system capable of specifically visualizing and eliminating these toxic RNAs. We observe specific targeting and efficient elimination of micro-satellite repeat expansion RNAs both when exogenously expressed and in patient cells. Importantly, RNA-targeting Cas9

[*]Correspondence: geneyeo@ucsd.edu.
[10]These authors contributed equally
[11]Lead Contact

(RCas9) reverses hallmark features of disease including elimination of RNA foci among all conditions studied (DM1, DM2, C9-ALS, polyglutamine diseases), reduction of polyglutamine protein products, relocalization of repeat-bound proteins to resemble healthy controls, and efficient reversal of DM1-associated splicing abnormalities in patient myotubes. Finally, we report a truncated RCas9 system compatible with adeno-associated viral packaging. This effort highlights the potential of RCas9 for human therapeutics.

Graphical abstract
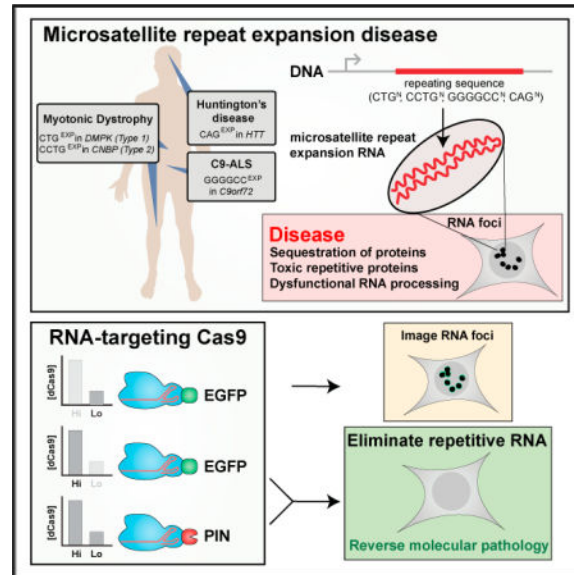
INTRODUCTION

Microsatellite repeat expansions (MREs) are repetitive 2–9 base pair (bp) DNA sequences that often encode toxic repetitive RNA and protein products that cause cellular damage and debilitating human disease. MREs can be located in non-coding regions of protein-coding genes, resulting in diseases such as the most common form of adult-onset muscular dystrophy (myotonic dystrophy type 1 [DM1]) (Batra et al., 2010; Turner and Hilton-Jones, 2010) and DM2 (Liquori et al., 2001), the most common form of familial amyotrophic lateral sclerosis (*C9orf72*-linked ALS [C9-ALS]) (DeJesus-Hernandez et al., 2011; Renton et al., 2011), fragile-X tremor ataxia syndrome (FXTAS) (Oberlé et al., 1991; Verkerk et al., 1991; Yu et al., 1991), spinocerebellar ataxias 8, 10, and 12 (Daughters et al., 2009; Holmes et al., 1999; Matsuura et al., 2000), Huntington's disease-like 2 (HDL2) (Wilburn et al., 2011), and Fuchs' corneal dystrophy (Du et al., 2015). MREs in coding regions lead to hereditary conditions such as Huntington's disease (HD) (The Huntington's Disease Collaborative Research Group, 1993), spinal bulbar and muscular atrophy (Kennedy Disease) (La Spada et al., 1991), spinocerebellar ataxias (SCAs) 1, 2, 3, 6, 7, and 17 (La Spada and Taylor, 2010), and oculopharyngeal muscular dystrophy (OPMD) (Batra et al., 2015; O'Rourke and Swanson, 2009). As these are prevalent conditions with limited recourse, technologies to visualize and eliminate these pathogenic transcripts in live cells would have tremendous diagnostic and therapeutic utility.

Current efforts to target repetitive RNAs linked to MREs typically involve antisense oligonucleotides (ASOs) to block (Wheeler et al., 2009) or cleave (Lagier-Tourenne et al., 2013; Lee et al., 2012) MRE RNAs. However, ASOs must be routinely administered for life, are large, do not efficiently cross the blood-brain barrier, and rely on passive targeting that contributed to inefficient accumulation in target tissue and clinical failure for a DM1 indication (Geary et al., 2015) (ClinicalTrials.gov Identifier: NCT02312011). Recent developments in the gene therapy fields indicate that long-term production of encoded therapeutic materials is possible using adeno-associated virus (AAV) (Naldini, 2015). RNAi systems can be encoded in viral vectors but typically do not target MRE RNAs efficiently. Overexpression of MRE RNA-bound proteins can compensate for some features of DM1 (Kanadia et al., 2006), but the tissue specificity of sequestered proteins and the toxicity of translated products from MRE RNAs reduces the therapeutic potential of this approach. Thus, an RNA-directed platform encoded in AAV may provide long-term therapeutic value for MRE diseases.

CRISPRs form the bases of bacterial immune systems that recognize and destroy invading genetic material. Reconstitution of these DNA-targeting systems in eukaryotic cells has supported powerful applications in genome editing and other DNA-level manipulations (Sheridan, 2017; Yu et al., 2017). A recent study demonstrated that MREs can be excised from DNA using genome engineering techniques (van Agtmaal et al., 2017), although simultaneous induction of a pair of double-stranded breaks flanking the MRE followed by DNA repair in the absence of the repeat has not been realized with sufficient efficiency. In contrast to RNA-directed systems, genome editing features the risk of permanent, off-target edits to DNA (Schaefer et al., 2017).

We recently established the ability of a repurposed type II CRISPR system to target and track RNA in human cells (Nelles et al., 2016) indicating the potential of other manipulations to RNA using RNA-targeting Cas9 (RCas9). Here, we describe the ability of CRISPR/Cas9 to both visualize and eliminate MRE RNAs in human cells. We observe efficient elimination of all MRE RNAs studied (CUG^{exp}, $CCUG^{exp}$, CAG^{exp}, $GGGGCC^{exp}$) and describe a reduced version of the RNA-targeting CRISPR/Cas9 system optimized for AAV packaging. We also observe reversal of disease-associated molecular defects including reduction of polyglutamine proteins and near-complete reversal of DM1-associated splicing defects to resemble healthy patient cells (93% reversal in patient myotubes). As the most prominent symptoms (myotonia, muscle weakness, and cardiac conduction defects) associated with DM1 are caused by dysfunctional RNA splicing, this efficient splicing reversal indicates the broad efficacy of this approach.

RESULTS

RNA-Targeting Cas9 Supports Imaging of Microsatellite Repeat Expansion RNAs that Cause DM1, DM2, and CAG^{exp} Diseases

We recently demonstrated that GFP-tagged, nuclease-dead Cas9 (dCas9) with an antisense oligonucleotide carrying a proto-spacer adjacent motif (PAMmer) and a single guide RNA (sgRNA) enables visualization of mRNAs in human cells (Nelles et al., 2016). To target repeat expansion RNAs for visualization (Figure 1), we expressed MRE RNA in COS-M6

cells with plasmids encoding 105 copies of CTG (CTG^{105}) found in DM1, 300 copies of CCTG ($CCTG^{300}$) from DM2, and 80 copies of CAG (CAG^{80}) from HD (Bañez-Coronel et al., 2015) and various SCAs. RNA fluorescence in situ hybridization (FISH) for these repeats revealed formation of RNA foci in a manner consistent with typical manifestation of these diseases (Figure 1B) (see probe sequences in Table S1) (Wojciechowska and Krzyzosiak, 2011). We next constructed plasmids encoding sgRNAs utilizing optimized scaffold sequences (Chen et al., 2013) targeting CUG, CCUG, and CAG MRE RNAs. Co-transfection with plasmids expressing repeat RNAs, their cognate sgRNAs, and dCas9-GFP resulted in colocalization of RCas9 signal and RNA foci (Figure 1B) that was absent in the presence of a non-targeting sgRNA (sequences in Table S2). RCas9 localization was dependent on repetitive RNA and not the plasmids, as demonstrated by treatment with DNase (did not alter FISH signal) and RNase (abolished the nuclear foci) (Figures S1A and S1B). Co-staining for nucleolin and FISH against CUG repeat RNA showed little overlap between nucleoli and CUG signal (Figure S1C), suggesting co-localization is not coincidental (Chen et al., 2013). In these experiments, we found that the PAMmer oligonucleotide was unnecessary to promote co-localization of RNA FISH and RCas9 signal that is consistent with previous studies that indicate RNA recognition does not require, but is enhanced by, the PAMmer (Nelles et al., 2016; O'Connell et al., 2014). We conclude that RCas9 enables the visualization of microsatellite repeat expansion RNA in live cells.

RNA-Targeting Cas9 Eliminates CUG^{exp}, $CCUG^{exp}$, $GGGGCC^{exp}$, and CAG^{exp} RNA Foci

Next, we evaluated if RNA foci can be dissipated by dCas9 targeting the repetitive RNAs. We constructed expression vectors carrying HA-tagged dCas9 fused to the PIN RNA endonuclease domain from *SMG6* (Takeshita et al., 2007) and measured the ability of this fusion and dCas9 alone to eliminate repeat expansion RNA by imaging and RNA dot blot analysis (Figure 2). We co-transfected COS-M6 cells with plasmids encoding dCas9 or PIN-dCas9, each repeat expansion (Bañez-Coronel et al., 2015; Zu et al., 2013) and the cognate sgRNA. FISH and HA immunofluorescence were used to detect MRE RNAs and dCas9, respectively. We observed efficient elimination of RNA foci in HA-positive cells (Figures 2A–2E). Importantly, nearby cells lacking expression of dCas9 or PIN-dCas9 frequently exhibit CUG^{exp} RNA foci (Figure 2B, white arrows indicate dCas9-transfected cells, red arrows are untransfected). The ability of dCas9 alone to eliminate CUG repeat RNA foci is consistent with studies involving blocking ASOs and engineered RNA binding proteins that indicate simple binding to CUG repeat RNAs is sufficient to attenuate their levels (Wheeler et al., 2009; Zhang et al., 2014).

We quantified the ability of CRISPR/Cas9 to promote loss of repeat expansion foci by counting the number of cells with at least one nuclear RNA focus in the presence of the RCas9 system and normalized to total number of cells transfected with repeat expansion RNAs (Figure 2F) and observed near-complete elimination of CUG, CCUG, CAG, and GGGGCC repeat RNA foci. We observed that the PAMmer is not required to promote efficient elimination of RNA foci (Figure 2C) and conducted all subsequent experiments without a PAMmer unless otherwise specified. To assess whether repeat expansion RNA levels were attenuated or foci were simply dispersed, we conducted RNA dot blots against CUG, CCUG, and GGGGCC repeat expansion RNAs in the presence of the RCas9 system

and observed that dCas9 fused to a non-specific RNA endonuclease (PIN-dCas9), dCas9-GFP, and wild-type (WT) nuclease-active Cas9 all reduced repeat RNA levels (Figures 2G–2I). We conclude that RCas9 eliminates repeat expansion RNAs.

RNA-Targeting Cas9 Binds and Promotes Cleavage of Microsatellite Repeat Expansion RNAs

To investigate RCas9 interaction with MRE RNAs and evaluate the value of the PIN domain, we conducted a set of binding, pull-down, and RNA cleavage experiments both in vitro and in cells (Figures 3A–3E). We first performed an electrophoretic mobility shift assay (EMSA) with increasing amounts of COS-M6 total cell extract (ranging from 0–40 μg of total protein) from cells co-transfected with dCas9-GFP and either a CUG-targeting sgRNA or a non-targeting control sgRNA to evaluate dCas9-GFP binding to 10 ng of ^{32}P-labeled $(CUG)^{12}$ RNA. Protein-RNA ternary complex formation in the presence of Mg^{2+} followed by native gel electrophoresis revealed that while CUG^{12} RNA did not associate with dCas9-GFP in the presence of non-targeting (NT) sgRNA, extract containing dCas9-GFP and a CUG-targeting sgRNA resulted in retarded $(CUG)^{12}$ migration that is dependent on the concentration of the extract (Figure 3A). All measurements were conducted in the absence of a PAMmer. Addition of an anti-GFP antibody prevented the $(CUG)^{12}$ RNA from entering the gel (supershift), indicating that dCas9-GFP is indeed bound to $(CUG)^{12}$ RNA. Further, immunoprecipitation of dCas9-GFP:CUG-targeting sgRNA (and not non-targeting -sgRNA) using an anti-GFP antibody yielded *in vitro* transcribed ^{32}P-labeled $(CUG)^{54}$ RNA (Figure 3B). These results indicate that RCas9 can directly interact with CUG MRE RNA in the absence of a cognate PAMmer.

We next investigated whether the PIN domain enhances cleavage of repetitive RNA. We combined increasing concentrations (10–40 μg) of COS-M6 cellular extract containing CUG-targeting sgRNA and either dCas9-GFP or PIN-dCas9 with radiolabeled $(CUG)^{12}$ RNA and observed that PIN-dCas9 with cognate sgRNA, but not with non-targeting -sgRNA, promoted cleavage of $(CUG)^{12}$ RNA while dCas9-GFP did not (Figure 3C). We next assessed the dose-dependence of this RNA cleavage activity in living cells via transfection of increasing amounts of plasmids encoding dCas9-GFP or PIN-dCas9 with the cognate sgRNA and a plasmid expressing $(CUG)^{105}$ RNA. We observed that CUG repeat RNA was more efficiently eliminated by lower levels of PIN-dCas9 compared to dCas9-GFP (Figures 3D and 3E). Therefore, subsequent repeat RNA elimination experiments were conducted with PIN-dCas9.

We next conducted transcriptional arrest experiments utilizing a drug-inducible reporter to evaluate whether RCas9-based RNA elimination occurs post-transcriptionally (Figures 3F–3I). We utilized a plasmid containing a tetracycline-inducible promoter (tetracycline responsive element or TRE) driving the expression of GFP and exons 11–15 of *DMPK*, where exon 15 carries 960 CTG repeats (Lee et al., 2012) (Figure 3F). We transfected COS-M6 cells with this plasmid along with a separate plasmid containing the tetracycline transactivator (tTA) and added doxycycline to induce expression of GFP and the repeat-harboring *DMPK* transcript. 24 hr later, we removed doxycycline to halt transcription of GFP and CUG repetitive RNA. We then transfected dCas9 (PIN-dCas9 or dCas9-GFP) and

sgRNA (CUG-targeting or non-targeting). We assessed the amount of CUG repeat expansion RNA by FISH and RT-qPCR after 24 hr. We observed that CUG repeat RNA foci appeared upon addition of doxycycline, and after transcription of the repeat was halted, the foci were efficiently eliminated by PIN-dCas9 in the presence of the CUG-targeting but not the non-targeting sgRNA (Figures 3G and 3H). After transcriptional arrest, using PCR primers that interrogate a portion of *DMPK* exon 15 just upstream of the repeat (Lee et al., 2012), we observed that both dCas9-GFP and PIN-dCas9 efficiently attenuated levels of the repeat-harboring transcript in the presence of CUG-targeting sgRNA (Figures 3I and 3J). This result shows that RCas9 eliminates CUG repeat RNA foci post-transcriptionally. We used a similar approach that involves doxycycline-inducible expression of the hexanucleotide repeat $(GGGGCC)^{90}$ to also demonstrate a contribution of post-transcriptional elimination of RNA foci (Figure S2A).

Truncated Cas9 Proteins Promote Efficient Elimination of Repeat Expansion RNAs

In contrast to genome engineering, RNA-level therapeutics must continuously engage pathogenic RNAs as they are transcribed. AAVs provide a means to generate therapeutic transgenes such as the RCas9 system for long periods after initial administration but have a limited packaging capacity of 4.7 kb. We therefore investigated the ability of various versions of the RCas9 system (truncated, AAV-compatible Cas9 and various fusions) to eliminate repeat expansion RNA via semi-quantitative dot blot measurements (Figure 4). We constructed 3 truncation mutants predicted to not disrupt folding of Cas9 protein. These mutants lack the HNH domain ("ΔHNH") or the HNH and REC2 domains ("ΔHNH, ΔREC2"). We also generated a mutant composed of only the REC-lobe ("REC-only") (Figure 4A). We compared the ability of the ΔHNH and REC-only truncation mutants fused to the PIN domain to eliminate CUG repeat expansion RNA by dot blot compared to PIN-dCas9 ("full length") and observed elimination comparable to full-length protein by the ΔHNH truncation mutant (Figure 4B). We conducted a similar measurement with the ΔHNH and ΔREC2 versions with and without a fused PIN domain and observed reduction of CUG repeat expansion RNA only in the presence of the PIN (Figure 4C). We also observed by dot blot that the ΔHNH truncation fused to PIN efficiently reduces CCUG and CAG repeat RNA levels (Figures 4D and 4E). RNA foci elimination by various forms of our system (PIN-dCas9, dCas9-GFP, PIN-ΔHNH, and WT dCas9) along with their quantification are shown in Figures S3A and S3B.

RNA-Targeting Cas9 Eliminates Pathogenic Repeat RNA Foci in Patient Cells

We next assessed the ability of RCas9 to eliminate endogenous MRE RNA in myoblasts and fibroblasts prepared from DM1 and DM2 patient biopsies (Gao et al., 2016; Xia et al., 2013), respectively. For purposes of ensuring homogeneity of expression, we constructed lentiviral vectors based upon the LentiCRISPRV2 construct (Sanjana et al., 2014), which express U6 promoter-driven sgRNAs targeting CUG^{exp} or $CCUG^{exp}$ RNAs along with elongation factor-1 short (EFS)-driven PIN-dCas9 and a puromycin resistance gene that enabled cell selection. Patient cells were infected with lentivirus carrying an sgRNA targeting the cognate repeat RNA or a non-targeting control (Table S2) and subjected to puromycin selection (4 days with 2 μg/mL). Unlike cells that are subject to exogenous repeat RNA expression, ~99% of untreated patient myoblasts and fibroblasts exhibited

characteristic nuclear RNA foci of DM1 and DM2. Importantly, while the presence of PIN-dCas9 and non-targeting sgRNA did not reduce RNA foci, PIN-dCas9 with sgRNAs targeting the respective repeat RNAs resulted in near-complete elimination of both DM1 and DM2 RNA foci (Figures 5A–5C) and reduction of CUG^{exp} RNA levels to levels observed in healthy patient cells (Figure 5D).

To assess whether the observed loss of repeat expansion RNA in DM1 patient cells could be attributed to interactions between dCas9 and genomic (DNA) repeat expansions, we constructed lentivirus carrying PIN-dCas9 and two distinct sgRNAs targeting the template DNA strand of the CTG repeat expansion, i.e., guides targeting CAG repeats ("CAG2" and "CAG3"). Both of these guides did not affect CUG repeat RNA levels by dot blot analysis (Figure 5E). To assess whether RCas9 interacted directly with *DMPK* mRNA harboring the CUG repeat expansion, we conducted RNA immunoprecipitation (RIP) PCR of dCas9 in the presence of sgRNA targeting CUG repeat RNA and normalized to pull-down with non-targeting sgRNA. We observed significant enrichment of *DMPK* RNA ($p < 0.0037$ versus an intergenic region) but not of other genes with short CTG or CAG repeats (*TCF4* and *AR*, respectively) (Figure 5F).

We also explored whether RCas9 could reduce endogenous CUG repeat RNA levels via mechanisms at the DNA level: alteration of transcription of the *DMPK* 3′ UTR that harbors the expansion and destabilization of the repeat expansion itself. To investigate the first mechanism, we conducted chromatin immunoprecipitation for RNA polymerase II (pol II) and measured its association at one site upstream and two sites downstream of the repeat expansion in DM1 patient myotubes. We observed no RCas9-induced alterations in pol II occupancy at these sites (Figure S2B and S2C), indicating that progression of elongating pol II through repeat site is not impaired in the presence of the CUG-targeting RCas9 system. In contrast, we observed that the reduction of pol II occupancy at *MYOG*, encoding the muscle-specific transcription factor myogenin, in DM1 patient cells was rescued with RCas9, suggesting reversal of the impairment of muscle differentiation in DM1 myotubes (Timchenko et al., 2001). We also observed no alteration in transcript levels of the *SIX5* gene, whose transcriptional start site is located just 400 bp downstream of *DMPK* (Figure S2D). To investigate whether RCas9 leads to contraction of the genomic repeat expansion, we conducted repeat-primed PCR. We detected no difference in the length of the expansion in patient cells between cells treated with targeting and non-targeting RCas9 systems (Figure S2E). We conclude that CUG-targeting RCas9 reduces endogenous RNA foci in patient cells independent of affecting pol II occupancy.

RNA-Targeting Cas9 Restores MBNL1 Localization and Known Splicing Defects of DM1

DM1 is characterized by the presence of CUG^{exp} RNA foci that sequester splicing factor MBNL1, resulting in widespread dysregulation of AS (alternative splicing) in muscle (Nakamori et al., 2013). We investigated whether CUG-targeting RCas9 reverses these molecular hallmarks of the disease. We first evaluated the localization of GFP-tagged MBNL1 in the presence of $(CUG)^{exp}$ RNA and the RCas9 system. We co-transfected COS-M6 cells with plasmids carrying GFP-tagged MBNL1, PIN-dCas9, and an sgRNA targeting $(CUG)^{exp}$ RNA or a non-targeting sgRNA and visualized $(CUG)^{exp}$ RNA by RNA FISH. We

observed reversal of MBNL1 co-localization with RNA foci to a diffuse nuclear pattern upon application of the RCas9 system (Figure 6A). We observed similar redistribution of endogenous MBNL1 localization evaluated by immunofluorescence with an anti-MBNL1 antibody in primary patient myoblasts infected with lentivirus encoding the RCas9 system. Our result indicates the ability of RCas9 to reverse this key disease biomarker when the repeat is present in its native context (Figures 6B and 6C).

We next set out to determine the status of established splicing biomarkers (Nakamori et al., 2013) in DM1 patient and healthy control myoblasts and myotubes in the presence of the RCas9 system. Semiquantitative RT-PCR-based splicing assays were carried out in samples subjected to RCas9 with sgRNAs targeting CUG^{exp} RNA or non-targeting controls with primers flanking known alternative exons of the Ca^{2+}-ATPase *ATP2A1* (exon 22) also known as *SERCA1*, insulin receptor *IR* (exon 11), muscle-blind-like proteins *MBNL2* (exon 6) and *MBNL1* (exon 6; see Table S5 for primer sequences and targets). In all cases, we observed that the RCas9 system targeting CUG^{exp} RNA resulted in reversal of DM1-specific AS patterns to those of healthy control cells (Figures 6D).

RCas9 Reduces the Levels of Disease-Associated Polyglutamine Protein

A host of neurodegenerative conditions are linked to the production of polyglutamine (polyQ)-containing proteins (Paulson et al., 2000) from RNAs harboring expansions of the trinucleotide repeat CAG. We therefore also assessed whether polyQ protein levels could be attenuated by RCas9 targeting CAG^{exp}. We transfected a plasmid carrying 80 CAG repeats along with various versions of Cas9 and CAG-targeting or non-targeting sgRNAs (Figure 6E) into COS-M6 cells. We observed that the levels of polyQ protein produced from this plasmid were dramatically reduced in cells co-transfected with CAG-targeting sgRNA, and either wtCas9, PIN-dCas9, or dCas9, but not dCas9 HNH construct, while no reduction was seen in the absence of the sgRNA or a non-targeting sgRNA. We conclude that RCas9 can target CAG^{exp} RNAs to reduce disease-associated polyQ protein production.

RNA-Targeting Cas9 Corrects Transcriptome-wide Splicing Defects in DM1 Patient Cells

Previous reports describe that hundreds of developmentally regulated splicing events are misregulated in DM1 tissues (Batra et al., 2014; Charizanis et al., 2012; Poulos et al., 2011). To study the genome-wide changes in AS, we used RNA sequencing (RNA-seq) to profile global splicing patterns in DM1 and control primary patient myoblasts and myotubes. We performed RNA-seq data analysis using Olego and Quantas software suites (Charizanis et al., 2012; Wu et al., 2013) and found 350 high-confidence misregulated AS events (p < 10^{-4}, false discovery rate [FDR] < 0.0015, differential index of splicing [dI] > |0.15|) in DM1 myotubes compared to control myotubes (Figure 7A; Table S3). The differential splicing index (SI or percent-spliced in [PSI] value) of a specific exon between two compared samples is the dI value. Hierarchical clustering of these top 350 AS changes showed that the control samples are expectedly distinct from the DM1 samples, but strikingly, DM1 patient myoblasts treated with lentivirus encoding the RCas9 system that target the CUG MRE RNA (CTG-DM1) grouped with the control group (Figure 7A). Similar clustering with the control group was also observed for CTG-DM1 myotubes (Figure 7B). 211 significant AS changes were found (p < 10^{-4}, FDR < 0.002, dI > |0.15|) between control and DM1

myoblasts (Table S3). Remarkably, treatment with lentivirus encoding the CUG-targeting RCas9 system reversed 327 of 350 (93%) DM1-specific AS differences in myotubes and 157 of 211 (74%) AS changes in myoblasts (Figures 7C–7E). The genome browser tracks show RNA-seq reversal of congenital myotonic dystrophy (CDM)-related AS defects in CTG-DM1 myoblasts and myotubes (Figure 7F). RT-PCR-based assays validated reversal of aberrant AS detected by our RNA-seq experiments in dystonin (*DST*), alpha-spectrin 1 (*SPTAN1*), TEA domain family member 1 (*TEAD1*), Troponin 2 (*TNNT2*), and neuron navigator 3 (*NAV3*) exons (Figure 7G). Quantification of splicing changes is shown in Figure S4. We also determined changes in gene expression after treatment with the CUG-targeting RCas9 system in DM1 myoblasts (Table S3A). Expression of 9 genes was altered (adjusted $p < 0.05$; Table S3B) and no changes in genes containing non-pathological CTG or CAG repeats were observed in the RNA-seq data (Figure S5B; Table S3B). In COS-M6 cells transfected with $(CTG)^{105}$ and the CUG-targeting RCas9 system, we observed a drop in *DMPK* transcript levels but not in *TCF4* (CTG) or *AR* (CAG) transcript levels (Figure S5A). Finally, we observed an increase in the expression of muscle differentiation markers in CTG-DM1 myotubes when compared to NT-DM1 myotubes (Figure 7H). This efficient reversal of DM1-associated splicing defects in myoblasts and increase in myotube differentiation markers (Timchenko et al., 2001) demonstrates that RCas9 can reverse hallmark molecular features of this disease (Figure 7I, uncropped splicing electropherograms depicted in Figure S6).

DISCUSSION

Pathogenic expansions in both coding and non-coding regions of DNA that produce RNAs with toxic gains-of-function are linked to a host of human diseases (La Spada and Taylor, 2010). The mechanisms of MRE-linked toxicity depend on their sequence context, size, protein binding partners, and the properties of translated products. DM-linked $C(C)UG^{exp}$ expansions (DM1-linked CTG^{exp} in the 3′ UTR of *DMPK* and DM2-linked $(CCTG)^{exp}$ in the first intron of *CNBP*) form nuclear RNA foci and sequester MBNL proteins leading to the loss of MBNL-dependent RNA processing functions. Loss of MBNL function through sequestration is directly responsible for most DM-associated phenotypes that are recapitulated in mouse models of $(CTG)^{exp}$ and *Mbnl* knockouts (Lee et al., 2013). Hexanucleotide repeats ($GGGGCC^{exp}$) associated with familial ALS are present in the first intron of the *C9orf72 gene* and have been linked to RNA-mediated pathogenesis via formation of RNA foci (DeJesus-Hernandez et al., 2011), RNA binding protein (RBP) sequestration (Goodwin and Swanson, 2014), RNA splicing alterations (Cooper-Knock et al., 2015; Prudencio et al., 2015), translation of repetitive polypeptides (Ash et al., 2013; Zu et al., 2013), and alteration of nuclear-cytoplasmic transport (Zhang et al., 2015). $(CAG)^{exp}$ is associated with Huntington's disease, SBMA, and various SCAs in the coding regions of *HTT*, *AR*, and *ATXN1*, *ATXN2*, *ATXN3*, and other loci, respectively. The resulting poly-glutamine proteins are toxic in the CNS. Importantly, elimination of the MRE RNAs associated with these conditions provides a common therapeutic principle that could address gain-of-function toxicity on both RNA and protein levels.

Motivated by the lack of a common and efficient means to engage MRE RNAs, we repurposed the CRISPR/Cas9 system to visualize and target these toxic RNA transcripts.

Recent work established the ability of nuclease-null *Streptococcus pyogenes* Cas9 (spyCas9) fused to GFP (dCas9-GFP) to bind and track RNA in living cells (Nelles et al., 2016). As MRE RNAs frequently form highly localized RNA foci, we initiated our efforts by evaluating the ability of dCas9-GFP to reveal the localization of DM1-associated $(CUG)^{exp}$ RNA (Figure 1B). We observed that only one sgRNA among the three tested sgRNAs was capable of promoting co-localization among dCas9-GFP and RNA-FISH for the repeat expansion RNA, indicating that the structure of $(CUG)^{exp}$ RNA may restrict access of sgRNAs depending on the targeted frame. Using appropriate sgRNAs, this approach also allowed imaging of $CCUG^{exp}$ (DM2) and CAG^{exp} (HD, SBMA) RNA foci (Figure 1B) that indicates the flexibility of this approach in the context of MRE RNA targeting.

We observed that higher doses of dCas9-GFP resulted in elimination of repetitive RNA foci (Figure 2). Consistent with this result, other reports involving RNA-targeting ASOs (that do not form a substrate for RNase H) and programmable RNA binding proteins (lacking nuclease activity) indicate that simple binding to CUG repeats is sufficient to promote their elimination (Wheeler et al., 2009; Zhang et al., 2014). This widely observed phenomenon is likely due to displacement of RNA binding proteins that stabilize these repetitive RNAs. We observed that dCas9-GFP binds directly to CUG repeat RNA both by gel shift and pull-down assays (Figures 3A and 3B), supporting binding-mediated displacement of sequestered RBPs and subsequent destabilization of CUG repeat RNA.

Previous work involving RNA-targeting using CRISPR/Cas9 indicated that an antisense oligonucleotide that carries a short DNA motif (the protospacer adjacent motif or PAM of the form 5′-NGG-3′) can increase the affinity of Cas9 for RNA when hybridized to a target transcript. While the PAMmer increased the efficiency of RNA pull-down by dCas9 (O'Connell et al., 2014) and improved colocalization of dCas9 and FISH signal in the context of targeting non-repetitive RNA (Nelles et al., 2016), the PAMmer was not an obligate requirement for RNA recognition in both these previous studies. Consistent with these reports, we observe RNA recognition in the absence of the PAMmer (Figure 3A). In contrast to previous reports that indicate improved RNA recognition with the PAMmer, we observed no further increase in repeat RNA elimination (Figure 2C) or pull-down efficiency (Figure 3B) in the presence of the PAMmer. This dispensability of the PAMmer in the context of repetitive RNAs may be due to the presence of multiple sgRNA binding sites in target RNAs that increase the on-rate of target recognition. As the PAMmer must be chemically synthesized and cannot be encoded for viral delivery, this system lacking a PAMmer provides permutation of the RCas9 system compatible with AAV-packaged therapeutics.

In addition to supporting AAV-mediated delivery, an ideal therapeutic embodiment of RCas9 in the context of repeat expansion disease is the efficient elimination of repetitive RNAs even at low doses of protein-RNA complexes. We find that a PIN RNA endonuclease fused to dCas9 promotes more efficient repeat RNA elimination at lower doses than the non-nucleolytic dCas9-GFP. Motivated by this efficient elimination of pathogenic RNA, we utilized this PIN-dCas9 fusion in all subsequent studies of RCas9 in patient primary cells. The importance of the PIN domain for efficient RNA elimination supports an RNA-level therapeutic mechanism. Indeed, delivery of the RCas9 system after transcriptional arrest of

the CUG^{960} RNA strongly suggests that the elimination of repeat RNA is independent of transcription of the repeat expansion (Figures 3G–3I). Similar albeit less efficient elimination of (GGGGCC)^{90} repeat RNA after transcriptional arrest indicates the importance of an RNA-level mechanism for this repeat as well (Figure S2A). This reduced efficiency of (GGGGCC)^{90} RNA elimination compared to CUG^{960} RNA could be due to slower kinetics of repeat engagement due to its high GC content and highly structured conformation (Fratta et al., 2012) or a contribution of a DNA-level mechanism to hexanucleotide repeat RNA elimination due to its high density of PAM motifs (Sternberg et al., 2014). Cumulatively, these data indicate the predominance of an RNA-level mechanism for elimination of CUG repeat expansion RNA and suggest its importance for other repeat expansions, although a DNA level mechanism cannot be completely discounted for every MRE disease.

RNA-targeting therapeutics must continuously engage pathogenic RNAs as they are transcribed, and AAV provides a promising means to generate therapeutic transgenes such as the RCas9 system for long periods (Naldini, 2015) but are limited to a packaging capacity of ~4.7 kb. Using truncated dCas9 proteins (Sternberg et al., 2015; Nishimasu et al., 2014) fused to the PIN domain, we report versions of the RCas9 system that are capable of reducing repetitive RNA levels that fall within the packaging capacity of AAV. The truncations fused to PIN termed PIN-dCas9(ΔHNH) and PIN-dCas9(ΔHNH, ΔREC2) are ~4.3 and 3.9 kb long, respectively.

Having established that RCas9 engages repeat expansion RNAs in overexpression models, we also assessed the ability of RCas9 to engage repeat expansion RNAs in patient cells. After observing efficient elimination of CUG and CCUG repeat expansion RNAs in DM1 and DM2 patient cells, respectively (Figures 5A–5D), we investigated whether this effect could occur at the level of DNA. Cumulatively, our results indicate that sgRNAs targeting both DNA strands do not alter transcriptional dynamics of CTG repeats in patient cells both in terms of CUG repeat RNA levels (Figure 5E) and ChIP-PCR for elongating pol II flanking the repeat expansion (Figures S2B and S2C). These results, as well as assessments of repeat expansion integrity (Figure S2E), indicate that RNA foci elimination in DM1 patient cells is likely not due to DNA-level transcriptional disruption or destabilization of the DNA repeats themselves.

While the elimination of repetitive RNAs provides a common recourse for many MRE-linked diseases, the downstream manifestations of these diseases are diverse. DM1 is a well-known spliceopathy and loss of MBNL proteins through sequestration by (CUG)^{exp} leads to hundreds of AS changes to resemble fetal splicing patterns. As a result, colocalization of MBNL protein with CUG repeat RNA foci is a hallmark of this disease that is efficiently reversed to resemble patterns in healthy patient cells by RCas9 (Figures 6B and 6C). This result correlates with reversal of DM1-associated splicing defects that occur in an sgRNA-dependent manner (Figures 6C and 6D). Polyglutamine diseases are linked by the presence of CAG repeat expansion translation products that are similarly reduced by the RCas9 system (Figure 6E). Overall, these data indicate the potential of RCas9 to reverse the hallmark molecular defects associated with microsatellite repeat expansion diseases.

RNA-targeting therapeutics must specifically alter target RNAs and avoid unintended alterations to the transcriptome. Transcriptome-wide measurements of RNA splicing revealed highly efficient (>93%) reversal of DM1-associated splicing pathology in patient myotubes (Figure 7E). As myotubes cultured *in vitro* constitute an approximation of developing muscle tissue, this result indicates the therapeutic potential of RCas9 to broadly reverse DM1 molecular pathology. We also assessed off-target effects in the transcriptome and observed <10 expression changes in patient myotubes among treated and untreated cells. These changes could be due to the presence of the RCas9 system or variations in the stage of differentiation among the treated and untreated cells. Indeed, expression of myotube markers were different among these groups (Figure 7H). Surprisingly the expression of *TCF4*, a gene containing small CTG repeat expansions (~20 repeats) (Du et al., 2015) and CAG repeat-containing gene *AR* (8–31 repeats) (Edwards et al., 1992) were not affected by the CUG-targeting RCas9 system by both RNA-seq and qPCR. These results were also corroborated by our RIP experiment showing that RCas9 specifically binds *DMPK* RNA (>2,700 repeats) in a sgRNA-dependent manner but not *TCF4* (off-target with ~20 repeats, Figure 5F). This highlights the specific targeting achieved with RCas9 as off-target genes with small CTG repeat expansions on both the coding (CTG) and template (CAG) DNA strands are not affected.

A major concern among DNA-mediated CRISPR-based therapeutics is the potential for permanent off-target genetic lesions. As the PIN domain fused to nuclease null Cas9 utilized here is RNA-specific, we obviate the risk of permanent off-target DNA alterations as RNA-level effects are limited to the typically short lifetime of RNA (hours to days). Further, the sgRNAs utilized to target CUG repeat expansion RNA cannot efficiently bind DNA as the encoding DNA sequence lacks a canonical PAM.

Overall, our findings indicate the potential of a new therapeutic mechanism for the CRISPR/Cas9 system among a class of neurological and neuromuscular conditions linked to microsatellite repeat expansions. While the precise pathological mechanisms among these diseases are highly diverse, they are linked by the potential of therapeutics that target the RNA products of these repetitive DNA tracts. Beyond these in vitro demonstrations of highly efficient elimination of repeat expansion RNAs, further work is required to evaluate their behavior in vivo utilizing these rationally truncated Cas9 proteins.

STAR*METHODS

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

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-*Mbnl1*	Gift from Maurice Swanson	A2764
Critical Commercial Assays		

REAGENT or RESOURCE	SOURCE	IDENTIFIER
NE-PER Nuclear and Cytoplasmic Extraction Reagents	Thermo Fisher Scientific	Cat#78833
MAXIscript SP6 in vitro transcription kit	Thermo Fisher Scientific	Cat#AM1308
TruSeq Stranded mRNA Library Prep Kit	Illumina	Cat#RS-122-2101
Deposited Data		
Data reported in this paper	N/A	GEO: GSE100943
Experimental Models: Cell Lines		
DM1 patient myoblasts	Gift from Guangbin Xia	N/A
Oligonucleotides		
FISH probes	See Table S1	N/A
PCR primers	See Table S5	N/A
Recombinant DNA		
Plasmid: psPAX2	Gift from Didier Trono	Addgene Plasmid #1
Plasmid: pCMV-VSV-G	Gift from Bob Weinberg (Stewart et al., 2003)	Addgene Plasmid #8
Plasmid: DT960	Gift from Thomas Cooper (Lee et al., 2012)	TRE-CTG960
Plasmid: LentiRCas9 (EFS-PIN-dCas9 and U6-sgRNA)	This paper	N/A (to be deposited Addgene)
Plasmid: pCDNA 3.1 PIN-dCas9	This paper	N/A (to be deposited Addgene)
Plasmid: pBluescript SKII U6-sgRNA	See Table S2 for sgRNA spacers and scaffold sequences	N/A (to be deposited Addgene)
Software and Algorithms		
OLego	Zhang lab (https://zhanglab.c2b2.columbia.edu/index.php/OLego)	N/A
Quantas	Zhang lab (https://zhanglab.c2b2.columbia.edu/index.php/Quantas_Documentation#Download)	N/A

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, Gene W. Yeo (geneyeo@ucsd.edu). Important plasmids described in this study will be deposited in the Addgene plasmid respository and available under a standard MTA.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human cell lines and primary cells and an African green monkey kidney cell line (COS-M6) were utilized in this study. Patient myoblast cell lines were derived from muscle biopsies according to a University of Florida-approved IRB protocol involving informed consent of all patients. The genders of the patient cell lines are not available. The Lenti-X HEK293T cell line is derived from human female tissue and COS-M6 from African green monkey male tissue.

METHOD DETAILS

Constructs and materials—Transient transfections in the presence of overexpressed repeat expansion RNAs were conducted with a pair of vectors encoding Cas9 and sgRNA. The dCas9-2×NLS sequence was derived from pHR-SFFV-dCas9-BFPKRAB (a gift from Stanley Qi and Jonathan Weissman; Addgene plasmid no. 46911), tagged with two SV40 NLS's on the C terminus, and fused to EGFP or PIN or placed without fusion in pCDNA 3.1 (Life Technologies) using Gibson assembly. The sgRNA plasmid carries a human U6 promoter and a modified sgRNA (Chen et al., 2013) with BbsI sites for insertion of spacer (Nelles et al., 2016). Lentivirus encoding the RCas9 system was created via replacement of wtCas9 in LentiCRISPR v2 (Sanjana et al., 2014) with PIN-dCas9 in-frame with downstream 2A-puromycin present in the parent vector using Gibson assembly. We also replaced the sgRNA scaffold with the modified scaffold mentioned above. All amplifications were conducted with ClonAmp HIFI PCR premix (Clontech).

Lentivirus production—LentiCRISPRv2-derived constructs encoding both PIN-dCas9 and a U6 promoter-driven sgRNA were co-transfected with psPAX2 (packaging plasmid) (gift from Didier Trono, Addgene plasmid #12260) and VSV-G (viral envelope) (Stewart et al., 2003) into Lenti-X HEK293T cells (Clontech) at a mass ratio of 5:4:3 using polyethyleneimine (PEI) (4:1 mass ratio PEI:DNA). Viral supernatants were collected and concentrated with Lenti-X Concentrator reagent (Clontech) according to the manufacturer's protocol (100X concentration) and aliquoted for storage.

Cell culture and lentiviral infections—COS-M6 cells were cultured in DMEM with 10% FBS and 1% penicillin/streptomycin (GIBCO) and passaged at 90%–100% confluency. Primary myoblasts were maintained in complete SKGM-2 media (Lonza) and passaged at 70% confluency. Myotube differentiation was carried out by in DMEM with 2% horse serum (GIBCO) on collagen (GIBCO) at 90% confluency.

Myoblasts and myotubes were transduced with 100X concentrated lentivirus 24h after plating in growth media. Puromycin selection (2 μg/ml) was initiated at 48h and harvested after 3 days of drug selection.

Transfections—COS-M6 cells were seeded into 4-chambered glass slides at density of $5×10^4$ cells/well for FISH and $3×10^5$ cells per well of a 6-well plate for RNA isolation. For FISH, 50ng MRE plasmid was added to 100 μL optiMEM along with 250ng of the appropriate sgRNA and 250ng of the appropriate Cas9 plasmid. For RNA isolation, 150ng of the MRE plasmid was used with 500ng each of sgRNA and Cas9 plasmids. Transfection was carried out using Lipofectamine 3000 according to the manufacturer's instructions. Cells were incubated for 24 hr at 37°C and 5% CO_2. For GGGGCC repeats, 50ng of the MRE plasmid was used along with 500ng each of the sgRNA and Cas9 plasmids (1:10) for slides.

RNA fluorescence *in situ* hybridization—Fixation was carried out for 10 min with 4% PFA after media aspiration and PBS rinse. Next, cells were rinsed 3X with PBS and incubated overnight in the 70% ethanol at 4°C. Next, ethanol was removed and cells

rehydrated for 10 min with wash buffer 40% form-amide and 2X SSC buffer. Cells were next incubated with pre-hybridization buffer (10% Dextran sulfate, 2mM ribonucleoside vanadyl complex, 2X SSC pH = 5, 50% fresh & deionized RNase-free formamide, 200 μg/mL BSA, 1 mg/ml yeast tRNA, DEPC-treated water to attain final target volume) for 15 min at 37°C (65°C for GGGGCC) in a hybridization oven. Probe was denatured at 100°C for 10 min and 500pg/μL (1 μg/μL for G4C2; final concentration) of the probe was added to cold pre-hyb buffer and immediately added to the cells. Cells were hybridized for 2 hr at 37°C (65°C for GGGGCC) in a hybridization oven. Next, cells were washed 3× for 20 min each with wash buffer. Cells were then washed 1X with PBS and slides were mounted with DAPI containing-mounting medium (ProLong Diamond Antifade Mountant with DAPI). Image quantification (Figure 4C) was conducted by counting the number of cells containing at least one RNA focus in the presence of the CRISPR/Cas9 and either NT- or C(C)TG-targeting sgRNA.

RNA isolation and RNA dot blots—RNA isolations were carried out with Trizol according to the manufacturer's protocol. RNA quality and concentrations were estimated using the Nanodrop spectrophotometer or Agilent 2200 Tapestation using the manufacturers' protocol. 2-5 μg of total RNA was re-suspended in 1mM EDTA, pH 8.0 to a final volume of 50 μL followed by addition of 30 μL 20X SSC and 20 μL 37% formaldehyde to denature the RNA. The RNA was incubated for 30 min at 60°C followed by incubation on ice. The Biorad Bio-Dot apparatus was assembled according to the manufacturer's protocol with Hybond N+ nylon membrane (GE Healthcare) on the top. The membrane was equilibrated by passing 100 μL of 20X SSC through the slots using vacuum. The denatured RNA was passed through the slots and the membrane was washed with 20X SSC. The membrane was crosslinked at the auto setting in the UV Stratalinker (Stratagene). Expresshyb (Clontech) pre-warmed at 50°C was used to pretreat the membrane for 60 min at 55°C. The probe was end-labeled with α-^{32}P ATP (Perkin Elmer) and T4-PNK (NEB) for 30 min at 37°C, filtered through a G-50 Illustra spin column (GE Healthcare), denatured at 100°C for 10 min and placed on ice. The probe was added to the prehybridization solution (1 μg/mL) being careful to not add it to the membrane directly. Hybridization was conducted overnight and membrane was washed 1X with 1X SSC, 0.1% SDS at 55°C and 2X with 0.5X SSC, 0.1% SDS at 55°C. The membrane was exposed to autoradiography film (Thermo Fisher) with an intensifying screen (Kodak) at −80°C overnight (~16 hr).

Electrophoretic mobility shift assay—1 μg of (CUG)[12] RNA (IDT) was 5′ end-labeled using γ-32P-ATP (Perkin-Elmer) and T4-PNK (NEB) in a 100 μL reaction (10 μL 10X PNK buffer, 1 μL of 1 μg/μL RNA, 8 μL T4-PNK, 2 μL or 20uCi γ-32P-ATP, 79 μL RNase free water) at 37°C for 1 hr. The probe was cleaned with G-25 probe-quant micro columns (GE) according to the manufacturer's protocol. Cellular extracts were prepared 48h after transfection of COS-M6 cells with sgRNA and Cas9 plasmids (1.5 μg each/1.5 million cells in a 10 cm cell culture dish) using NE-PER kit (Thermo) and combining nuclear and cytoplasmic extracts and supplemented with $MgCl_2$ to a final concentration of 10mM. Protein concentration was measured using the BCA assay (Biorad). For electrophoretic mobility shift assay (EMSA), increasing concentrations (0, 1.25, 2.5, 5, 10, 20, or 40 μg of total protein) of cold cellular extract was mixed with 10 ng (1 μL of labeled reaction or

~100,000 cpm) of RNA diluted 1:10 with folding buffer (20mM Tris pH 7.5, 150mM NaCl, and 10mM MgCl$_2$) per reaction on ice. Samples were incubated at 37°C for 45 min for complex formation and run on a pre-ran (10 min) native 6% TBE gel (Novex) with 0.5X TBE buffer supplemented with MgCl$_2$ to a final concentration of 10mM in cold conditions. The gel was exposed to X-ray film for 30 min.

PIN-mediated RNA cleavage assay—^{32}P-labeled (CUG)^{12} RNA and cellular extracts were prepared as described for EMSA. Increasing concentrations (10, 20, or 40 μg of total protein) of cold cellular extract was mixed with 10 ng (1 μL of labeled reaction or ~100,000 cpm) of RNA and the reaction was incubated at 37°C for 1h. 2X RNA loading dye (95% deionized formamide, 5% RNase-free water, xylene cyanol, bromophenol blue) was added and the reaction was denatured at 65°C for 5 min and immediately transferred to ice. The samples were run on a 6% TBE-urea gel and gel was exposed to an X-ray film for 30 min and developed as described above.

RNA immunoprecipitation—RNA immunoprecipitation (RIP) was carried out as described previously (Niranjanakumari et al., 2002) with modifications. DM1 patient myoblasts were transduced with lentivirus containing either NT or CTG-targeting sgRNA and PIN-dCas9-HA followed by puromycin selection as described above. Cells were fixed with 0.5% formaldehyde (v/v) in PBS. The cross-linking was stopped with 0.25M glycine (in PBS) for 5 min at RT, and cells were washed with PBS 3X for 2 min each. Fixed cells were lysed in 500 μL of RIPA buffer (50 mM Tris–Cl, pH 7.5, 1% Igepal, 0.5% sodium deoxycholate, 0.05% SDS, 1mM EDTA, 150mM NaCl) containing protease inhibitor cocktail (Roche) and DNA digestion was carried out using 10U of TurboDNase (Thermo) for 10 min at 37°C. The samples were spun at 4°C in a microcentrifuge at full speed for 20 min and supernatant was used for IP. Dynabeads protein A (Thermo) coated with anti-HA antibody were used for IP with lysate at 4°C for 2 hr. The beads with Protein-RNA complexes were washed 3 times with RIPA buffer. DNase digestion was repeated and complexes were washed with RIPA buffer again. 1 mg/ml Proteinase K (NEB) was used to digest proteins and RNA was extracted using RNA phenol (Thermo) and chloroform. cDNA preparation was done using Superscript III (Thermo) with random primers according to the manufacturer's protocol. qPCR was carried out with primers for *DMPK*, *TCF4*, *AR* and an intergenic region.

In vitro transcription and immunoprecipitation—CUG repeats were in vitro transcribed using the MAXIscript SP6 in vitro transcription kit (Thermo) and α-^{32}P UTP (Perkin Elmer). Cellular extracts were prepared 48 hr post transfection with either NT-sgRNA and Cas9-GFP or CUG-targeting sgRNA and Cas9-GFP with 1×TBS (50mM Tris-HCl pH 7.0, 150mM NaCl) and 0.1% Igepal followed by sonication in the bioruptor. Equal CPM amounts of RNA was mixed with 500 μg (protein content) of cellular extract and 1mg/ml yeast tRNA for 60 min at 37C. Dynabeads M-280 sheep anti-rabbit (Thermo) were coated with anti-GFP antibody (Abcam Ab290) and immunoprecipitation was carried out for 2 hr at RT. The beads were washed with Tris-HCl, 0.1% Igepal 5X or until the IgG control sample did not have any detectable radiation as measured by a Geiger counter. The RNA was eluted using Proteinase K in PK buffer for 30 min at 37C and was blotted on Hybond N

+ membrane using the Bio-Dot apparatus as described above. The membrane was washed 1X with 1X SSC and developed with autoradiography film (Kodak) with an intensifying screen at −80C for 2-4 hr.

RNA-seq library preparation and data processing—Library preparation and sequencing was done as described previously (Batra et al., 2016, NSMB). Total RNA was isolated using TRIzol reagent (Life Technologies) and treated with Turbo DNase (Life Technologies). Libraries were prepared with Illumina TruSeq polyA mRNA Sample Preparation reagents according to the manufacturer's protocol. All samples were sequenced on the Illumina HiSeq 4000 platform. Demultiplexed Fastq files were checked for quality and aligned to the hg19 human genome platform. Cluster 3.0 software and Java Treeview were used in combination to perform and visualize results from hierarchical gene expression clustering results. For AS analysis, we used Olego and Quantas software suite as previously described (Batra et al., 2016; Wu et al., 2013).

QUANTIFICATION AND STATISTICAL ANALYSIS

Splicing quantification and statistical analysis—RNA-seq data processing involved alignment with the Olego de-novo splice junction aligner and were quantified using Quantas. Quantas calculates exon inclusion indices (SI or splicing index equivalent to PSI values) and differential splicing measures between two groups (dI values). Fisher exact test in Quantas is used to call FDR adjusted P values. We use FDR < 0.05 and dI > |0.15| as significant AS events.

RNA dot blot quantification—RNA dot blots and western blots were quantified using the gel analysis function in FIJI (Schindelin et al., 2012).

DATA AND SOFTWARE AVAILABILITY

The accession number for the data reported in this paper is GEO: GSE100943.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, et al. Unconventional translation of C9ORF72

GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron. 2013; 77:639–646. [PubMed: 23415312]

Bañez-Coronel M, Ayhan F, Tarabochia AD, Zu T, Perez BA, Tusi SK, Pletnikova O, Borchelt DR, Ross CA, Margolis RL, et al. RAN translation in Huntington disease. Neuron. 2015; 88:667–677. [PubMed: 26590344]

Batra R, Charizanis K, Swanson MS. Partners in crime: bidirectional transcription in unstable microsatellite disease. Hum Mol Genet. 2010; 19(R1):R77–R82. [PubMed: 20368264]

Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA, Swanson MS. Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. Mol Cell. 2014; 56:311–322. [PubMed: 25263597]

Batra R, Manchanda M, Swanson MS. Global insights into alternative polyadenylation regulation. RNA Biol. 2015; 12:597–602. [PubMed: 25892335]

Batra R, Stark TJ, Clark E, Belzile JP, Wheeler EC, Yee BA, Huang H, Gelboin-Burkhart C, Huelga SC, Aigner S, et al. RNA-binding protein CPEB1 remodels host and viral RNA landscapes. Nat Struct Mol Biol. 2016; 23:1101–1110. [PubMed: 27775709]

Charizanis K, Lee KY, Batra R, Goodwin M, Zhang C, Yuan Y, Shiue L, Cline M, Scotti MM, Xia G, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. Neuron. 2012; 75:437–450. [PubMed: 22884328]

Chen B, Gilbert LA, Cimini BA, Schnitzbauer J, Zhang W, Li GW, Park J, Blackburn EH, Weissman JS, Qi LS, Huang B. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell. 2013; 155:1479–1491. [PubMed: 24360272]

Cooper-Knock J, Bury JJ, Heath PR, Wyles M, Higginbottom A, Gelsthorpe C, Highley JR, Hautbergue G, Rattray M, Kirby J, Shaw PJ. C9ORF72 GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. PLoS ONE. 2015; 10:e0127376. [PubMed: 26016851]

Daughters RS, Tuttle DL, Gao W, Ikeda Y, Moseley ML, Ebner TJ, Swanson MS, Ranum LP. RNA gain-of-function in spinocerebellar ataxia type 8. PLoS Genet. 2009; 5:e1000600. [PubMed: 19680539]

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 2011; 72:245–256. [PubMed: 21944778]

Du J, Aleff RA, Soragni E, Kalari K, Nie J, Tang X, Davila J, Kocher JP, Patel SV, Gottesfeld JM, et al. RNA toxicity and missplicing in the common eye disease fuchs endothelial corneal dystrophy. J Biol Chem. 2015; 290:5979–5990. [PubMed: 25593321]

Edwards A, Hammond HA, Jin L, Caskey CT, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics. 1992; 12:241–253. [PubMed: 1740333]

Fratta P, Mizielinska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, Isaacs AM. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci Rep. 2012; 2:1016. [PubMed: 23264878]

Gao Y, Guo X, Santostefano K, Wang Y, Reid T, Zeng D, Terada N, Ashizawa T, Xia G. Genome therapy of myotonic dystrophy type 1 iPS cells for development of autologous stem cell therapy. Mol Ther. 2016; 24:1378–1387. [PubMed: 27203440]

Geary RS, Norris D, Yu R, Bennett CF. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. Adv Drug Deliv Rev. 2015; 87:46–51. [PubMed: 25666165]

Goodwin M, Swanson MS. RNA-binding protein misregulation in microsatellite expansion disorders. Adv Exp Med Biol. 2014; 825:353–388. [PubMed: 25201111]

Holmes SE, O'Hearn EE, McInnis MG, Gorelick-Feldman DA, Kleiderlein JJ, Callahan C, Kwak NG, Ingersoll-Ashworth RG, Sherr M, Sumner AJ, et al. Expansion of a novel CAG trinucleotide repeat in the 5′ region of PPP2R2B is associated with SCA12. Nat Genet. 1999; 23:391–392. [PubMed: 10581021]

Kanadia RN, Shin J, Yuan Y, Beattie SG, Wheeler TM, Thornton CA, Swanson MS. Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. Proc Natl Acad Sci USA. 2006; 103:11748–11753. [PubMed: 16864772]

La Spada AR, Taylor JP. Repeat expansion disease: progress and puzzles in disease pathogenesis. Nat Rev Genet. 2010; 11:247–258. [PubMed: 20177426]

La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature. 1991; 352:77–79. [PubMed: 2062380]

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li HR, Jiang J, Watt AT, Chun S, Katz M, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci USA. 2013; 110:E4530–E4539. [PubMed: 24170860]

Lee JE, Bennett CF, Cooper TA. RNase H-mediated degradation of toxic RNA in myotonic dystrophy type 1. Proc Natl Acad Sci USA. 2012; 109:4221–4226. [PubMed: 22371589]

Lee KY, Li M, Manchanda M, Batra R, Charizanis K, Mohan A, Warren SA, Chamberlain CM, Finn D, Hong H, et al. Compound loss of muscleblind-like function in myotonic dystrophy. EMBO Mol Med. 2013; 5:1887–1900. [PubMed: 24293317]

Liquori CL, Ricker K, Moseley ML, Jacobsen JF, Kress W, Naylor SL, Day JW, Ranum LP. Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science. 2001; 293:864–867. [PubMed: 11486088]

Matsuura T, Yamagata T, Burgess DL, Rasmussen A, Grewal RP, Watase K, Khajavi M, McCall AE, Davis CF, Zu L, et al. Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet. 2000; 26:191–194. [PubMed: 11017075]

Nakamori M, Sobczak K, Puwanant A, Welle S, Eichinger K, Pandya S, Dekdebrun J, Heatwole CR, McDermott MP, Chen T, et al. Splicing biomarkers of disease severity in myotonic dystrophy. Ann Neurol. 2013; 74:862–872. [PubMed: 23929620]

Naldini L. Gene therapy returns to centre stage. Nature. 2015; 526:351–360. [PubMed: 26469046]

Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, Yeo GW. Programmable RNA tracking in live cells with CRISPR/Cas9. Cell. 2016; 165:488–496. [PubMed: 26997482]

Niranjanakumari S, Lasda E, Brazas R, Garcia-Blanco MA. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. Methods. 2002; 26:182–190. [PubMed: 12054895]

Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F, Nureki O. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014; 156:935–949. [PubMed: 24529477]

O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature. 2014; 516:263–266. [PubMed: 25274302]

O'Rourke JR, Swanson MS. Mechanisms of RNA-mediated disease. J Biol Chem. 2009; 284:7419–7423. [PubMed: 18957432]

Oberlé I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boué J, Bertheas MF, Mandel JL. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science. 1991; 252:1097–1102. [PubMed: 2031184]

Paulson HL, Bonini NM, Roth KA. Polyglutamine disease and neuronal cell death. Proc Natl Acad Sci USA. 2000; 97:12957–12958. [PubMed: 11058149]

Poulos MG, Batra R, Charizanis K, Swanson MS. Developments in RNA splicing and disease. Cold Spring Harb Perspect Biol. 2011; 3:a000778. [PubMed: 21084389]

Prudencio M, Belzil VV, Batra R, Ross CA, Gendron TF, Pregent LJ, Murray ME, Overstreet KK, Piazza-Johnston AE, Desaro P, et al. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat Neurosci. 2015; 18:1175–1182. [PubMed: 26192745]

Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, et al. ITALSGEN Consortium. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron. 2011; 72:257–268. [PubMed: 21944779]

Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014; 11:783–784. [PubMed: 25075903]

Schaefer KA, Wu WH, Colgan DF, Tsang SH, Bassuk AG, Mahajan VB. Unexpected mutations after CRISPR-Cas9 editing in vivo. Nat Methods. 2017; 14:547–548. [PubMed: 28557981]

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012; 9:676–682. [PubMed: 22743772]

Sheridan C. CRISPR therapeutics push into human testing. Nat Biotechnol. 2017; 35:3–5. [PubMed: 28072788]

Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature. 2014; 507:62–67. [PubMed: 24476820]

Sternberg SH, LaFrance B, Kaplan M, Doudna JA. Conformational control of DNA target cleavage by CRISPR-Cas9. Nature. 2015; 527:110–113. [PubMed: 26524520]

Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, Sabatini DM, Chen IS, Hahn WC, Sharp PA, et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA. 2003; 9:493–501. [PubMed: 12649500]

Takeshita D, Zenno S, Lee WC, Saigo K, Tanokura M. Crystal structure of the PIN domain of human telomerase-associated protein EST1A. Proteins. 2007; 68:980–989. [PubMed: 17557331]

The Huntington's Disease Collaborative Research Group. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell. 1993; 72:971–983. [PubMed: 8458085]

Timchenko NA, Iakova P, Cai ZJ, Smith JR, Timchenko LT. Molecular basis for impaired muscle differentiation in myotonic dystrophy. Mol Cell Biol. 2001; 21:6927–6938. [PubMed: 11564876]

Turner C, Hilton-Jones D. The myotonic dystrophies: diagnosis and management. J Neurol Neurosurg Psychiatry. 2010; 81:358–367. [PubMed: 20176601]

van Agtmaal EL, André LM, Willemse M, Cumming SA, van Kessel IDG, van den Broek WJAA, Gourdon G, Furling D, Mouly V, Monckton DG, et al. CRISPR/Cas9-induced (CTG,CAG)n repeat instability in the myotonic dystrophy type 1 locus: implications for therapeutic genome editing. Mol Ther. 2017; 25:24–43. [PubMed: 28129118]

Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell. 1991; 65:905–914. [PubMed: 1710175]

Wheeler TM, Sobczak K, Lueck JD, Osborne RJ, Lin X, Dirksen RT, Thornton CA. Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. Science. 2009; 325:336–339. [PubMed: 19608921]

Wilburn B, Rudnicki DD, Zhao J, Weitz TM, Cheng Y, Gu X, Greiner E, Park CS, Wang N, Sopher BL, et al. An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron. 2011; 70:427–440. [PubMed: 21555070]

Wojciechowska M, Krzyzosiak WJ. Cellular toxicity of expanded RNA repeats: focus on RNA foci. Hum Mol Genet. 2011; 20:3811–3821. [PubMed: 21729883]

Wu J, Anczuków O, Krainer AR, Zhang MQ, Zhang C. OLego: fast and sensitive mapping of spliced mRNA-Seq reads using small seeds. Nucleic Acids Res. 2013; 41:5149–5163. [PubMed: 23571760]

Xia G, Santostefano KE, Goodwin M, Liu J, Subramony SH, Swanson MS, Terada N, Ashizawa T. Generation of neural cells from DM1 induced pluripotent stem cells as cellular model for the study of central nervous system neuropathogenesis. Cell Reprogram. 2013; 15:166–177. [PubMed: 23550732]

Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, Mulley JC, Warren ST, Schlessinger D, et al. Fragile X genotype characterized by an unstable region of DNA. Science. 1991; 252:1179–1181. [PubMed: 2031189]

Yu W, Mookherjee S, Chaitankar V, Hiriyanna S, Kim JW, Brooks M, Ataeijannati Y, Sun X, Dong L, Li T, et al. Nrl knockdown by AAV-delivered CRISPR/Cas9 prevents retinal degeneration in mice. Nat Commun. 2017; 8:14716. [PubMed: 28291770]

Zhang W, Wang Y, Dong S, Choudhury R, Jin Y, Wang Z. Treatment of type 1 myotonic dystrophy by engineering site-specific RNA endonucleases that target (CUG)n repeats. Mol Ther. 2014; 22:312–320. [PubMed: 24196578]

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, Daley EL, Miller SJ, Cunningham KM, Vidensky S, et al. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature. 2015; 525:56–61. [PubMed: 26308891]

Zu T, Liu Y, Bañez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci USA. 2013; 110:E4968–E4977. [PubMed: 24248382]

Highlights

- RNA-targeting Cas9 (RCas9) supports efficient targeting of repetitive RNAs

- An RNA endonuclease fused to nuclease-null Cas9 enables an RNA-specific CRISPR system

- An RCas9 system with truncated Cas9 can be packaged in adeno-associated virus

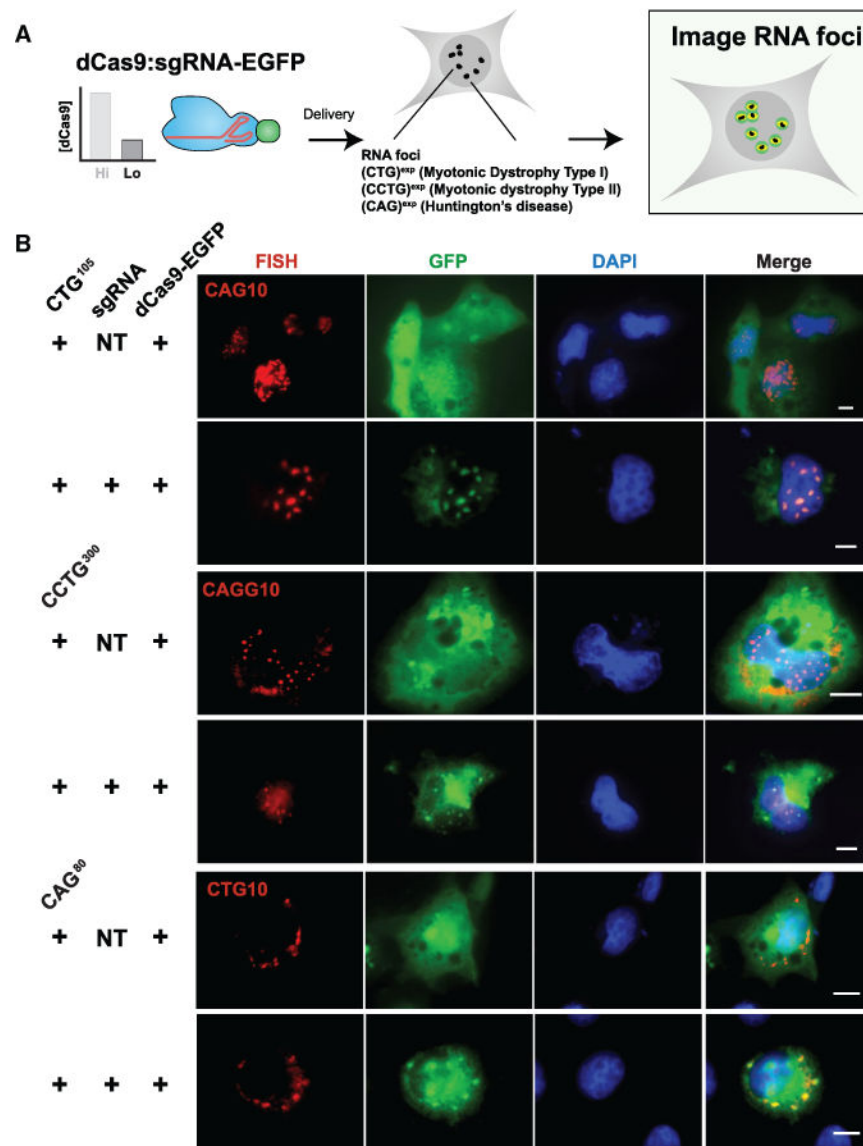- RCas9 reverses splicing defects in myotonic dystrophy type 1 patient cells

Figure 1. Visualizing Microsatellite Repeat Expansion RNAs with CRISPR/Cas9

(A) Schematic description of recognition and imaging of microsatellite repeat expansion RNA with low levels of RNA-targeting Cas9 (RCas9) fused to EGFP.

(B) RNA-FISH and RCas9 (EGFP) imaging in COS-M6 cells transfected with $(CTG)^{105}$, $(CCTG)^{300}$, or $(CAG)^{80}$. Each pair of rows features sgRNAs targeting the specified repeat (+) or a non-targeting sgRNA (NT). Scale bars are 10 μm.
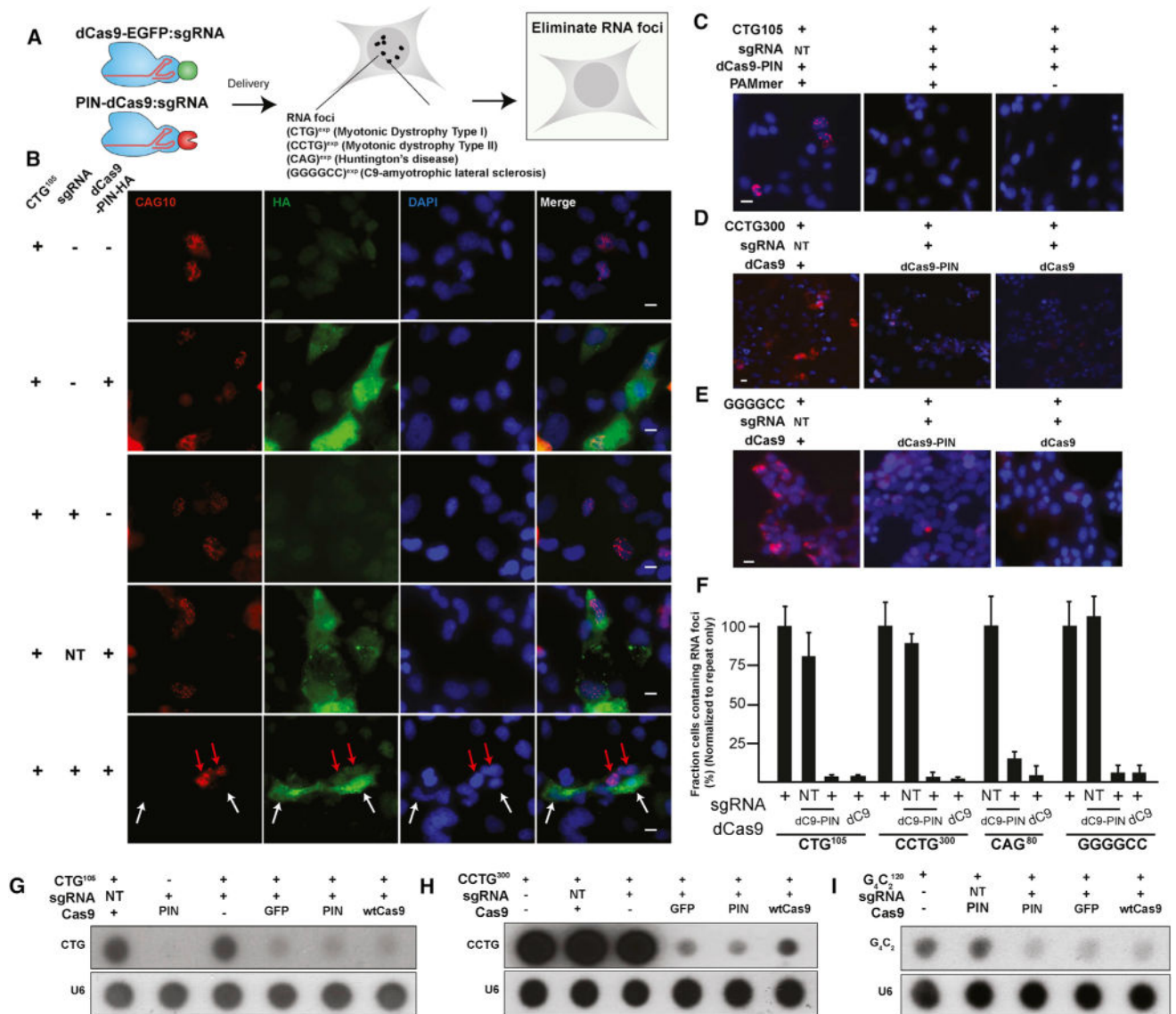
See also Figure S1 and Tables S1 and S2.

Figure 2. Degradation of Microsatellite Repeat Expansion RNA with RNA-Targeting Cas9

(A) Schematic description of elimination of microsatellite repeat expansion RNA with RNA-targeting Cas9 (RCas9) fused to EGFP or PIN domain.

(B) CUG RNA foci measured by RNA-FISH in COS-M6 cells transfected with (CTG)^{105}, either non-targeting sgRNA (NT), CUG-targeting sgRNA (+), or no sgRNA (−), and with (+) or without (−) HA-tagged PIN-dCas9. Scale bars in (B)–(E) are 10 μm.

(C) CUG RNA foci measured by RNA-FISH in COS-M6 cells transfected with (CTG)^{105} and PIN-dCas9, with either non-targeting sgRNA (NT) or CUG-targeting sgRNA (+), and with (+) or without (−) cognate PAMmer.

(D) CCUG RNA foci measured by RNA-FISH in COS-M6 cells transfected with (CCTG)^{300} and PIN-dCas9 or dCas9, with either non-targeting sgRNA (NT) or CCUG-targeting sgRNA (+).

(E) GGGGCC RNA foci measured by RNA-FISH in COS-M6 cells transfected with (GGGGCC)[120] and PIN-dCas9 or dCas9, with either non-targeting sgRNA (NT) or GGGGCC-targeting sgRNA (+).

(F) Quantification of RNA-FISH signal in COS-M6 cells transfected with various MREs and PIN-dCas9 or dCas9 with MRE-targeting (+) or non-targeting (NT) sgRNA. Cells containing at least 1 RNA focus are considered positive for MRE RNA. Measurements are normalized to the condition with the MRE-targeting sgRNA and MRE RNA but lacking dCas9. Error bars denote SDs determined from 3 biological replicates enumerating 100 transfected cells each.

(G) RNA dot blot of (CUG)[exp] levels in COS-M6 cells transfected with (CTG)[105], CTG-targeting or non-targeting (NT) sgRNA, and various forms of Cas9 (PIN-dCas9, dCas9-GFP, and wtCas9). U6 snRNA served as a loading control in (G)–(I).

(H) RNA dot blot of (CCUG)[exp] levels in COS-M6 cells transfected with (CCTG)[105], CCTG-targeting or non-targeting (NT) sgRNA, and various forms of Cas9 (PIN-dCas9, dCas9-GFP, and wtCas9).

(I) RNA dot blot of (GGGGCC)[exp] levels in COS-M6 cells transfected with (GGGGCC)[105], GGGGCC-targeting or non-targeting (NT) sgRNA, and various forms of Cas9 (PIN-dCas9, dCas9-GFP, and wtCas9).

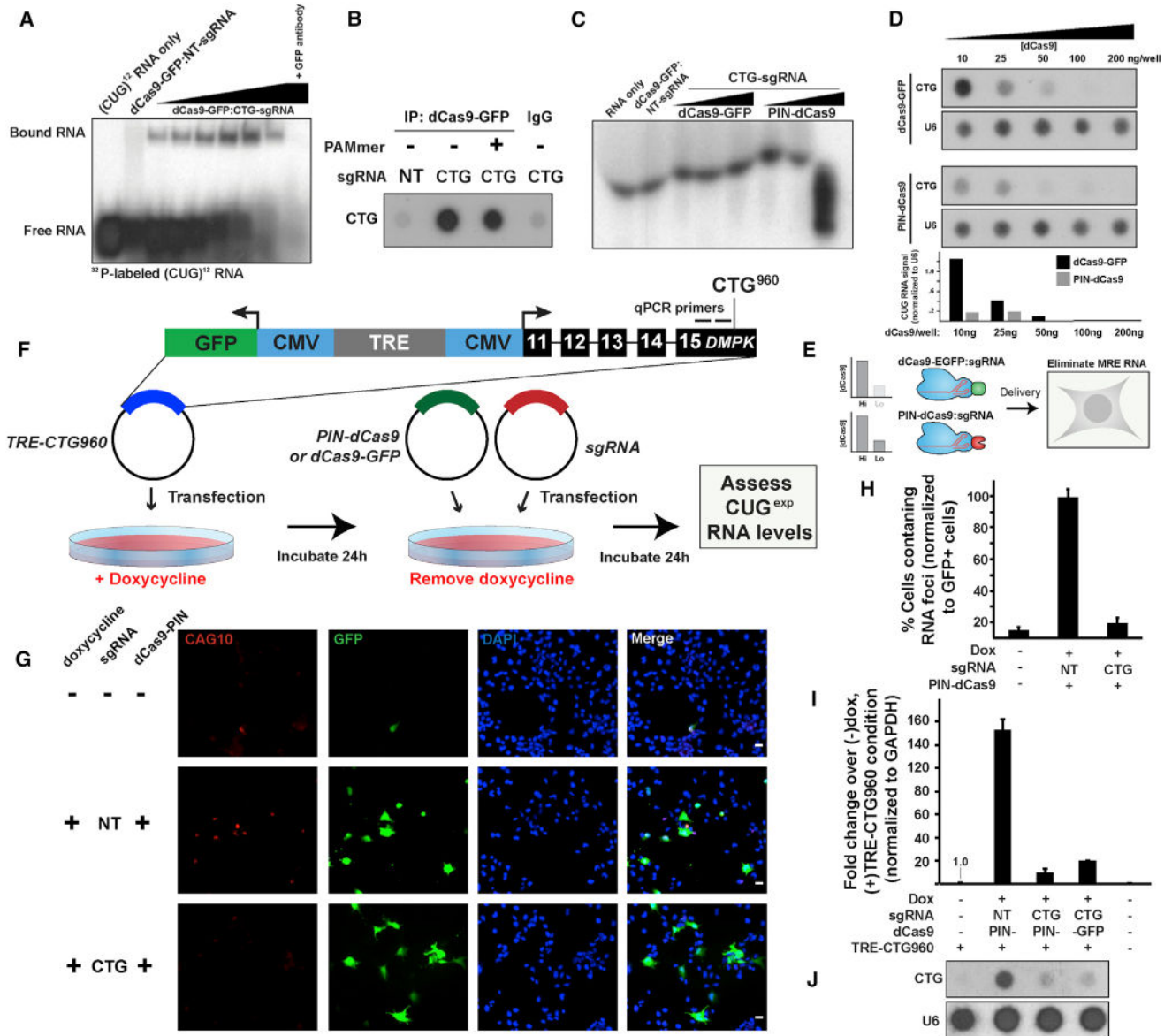See also Figures S2 and S3 and Tables S1 and S2.

Figure 3. RCas9 Binds and Targets CUG Repeat Expansion RNA for Degradation

(A) Gel shift assay of in vitro transcribed radiolabeled (CUG)^{12} RNA and increasing doses of COS-M6 cellular extract (0, 1.25, 2.5, 5, 10, 20, or 40 μg of total cellular protein) containing dCas9-GFP and sgRNAs (1.5 μg transfected plasmid per 1.5 million cells) targeting this repetitive RNA (CUG sgRNA) or a non-targeting control (NT-sgRNA) with 40 μg cell extract. The rightmost lane includes a GFP antibody (40 μg cell extract).

(B) Immunoprecipitation and dot blot quantitation of radiolabeled in vitro transcribed (CUG)^{54} after incubation with extract from cells expressing dCas9-GFP and sgRNAs targeting this repetitive RNA (CUG) or a non-targeting control (NT).

(C) In vitro cleavage assay of radiolabeled (CUG)^{12} RNA combined with increasing concentrations of extract (10, 20, 40 μg) from COS-M6 cells expressing dCas9-GFP or PIN-dCas9 and a CUG-targeting sgRNA or nontargeting (NT) sgRNA, as indicated.

(D) RNA dot blot assay of CUG^{exp} RNA levels in COS-M6 cells transfected with plasmids encoding (CTG)^{105}, and the indicated amounts of dCas9-GFP (top) or PIN-dCas9 (middle). U6 snRNA served as a loading control. Bottom: densitometric quantification.

(E) Scheme describing the relative activity of dCas9-GFP and PIN-dCas9 in the context of CUG microsatellite repeat expansion (MRE) RNA elimination as observed in (D).

(F) Scheme describing a tetracycline-inducible expression system for (CTG)^{960} repeat expansion in *DMPK* to assess whether RCas9-mediated repeat expansion RNA can be eliminated in the absence of transcription. TRE is the tetracycline responsive element.

(G) Detection of CUG^{exp} RNA foci by RNA-FISH in COS-M6 cells transfected and treated with doxycycline according to (E). Scale bars are 20 μm.

(H) Quantification of CUG^{exp}-FISH signal in COS-M6 cells transfected and treated with doxycycline according to (E). Cells containing at least 1 RNA focus were considered positive for CUG repeat RNA. Measurements were normalized to the total number of GFP-positive cells. Error bars denote SDs determined from 3 biological replicates enumerating 100 GFP-positive cells each.

(I) Quantification of CUG^{exp} containing *DMPK* transcript levels in COS-M6 cells transfected and treated with doxycycline according to (E) using RT-qPCR. *DMPK* mRNA levels were normalized to *GAPDH* and expressed as fold change over no-doxycycline condition. Error bars denote SDs determined from 3 biological replicates.

(J) RNA blot analysis quantifying CUG^{exp} containing *DMPK* transcript levels in COS-M6 cells transfected with the tetracycline-inducible expression system for (CTG)^{960} and treated with doxycycline according to (E). The legend from (I) also applies to this panel.

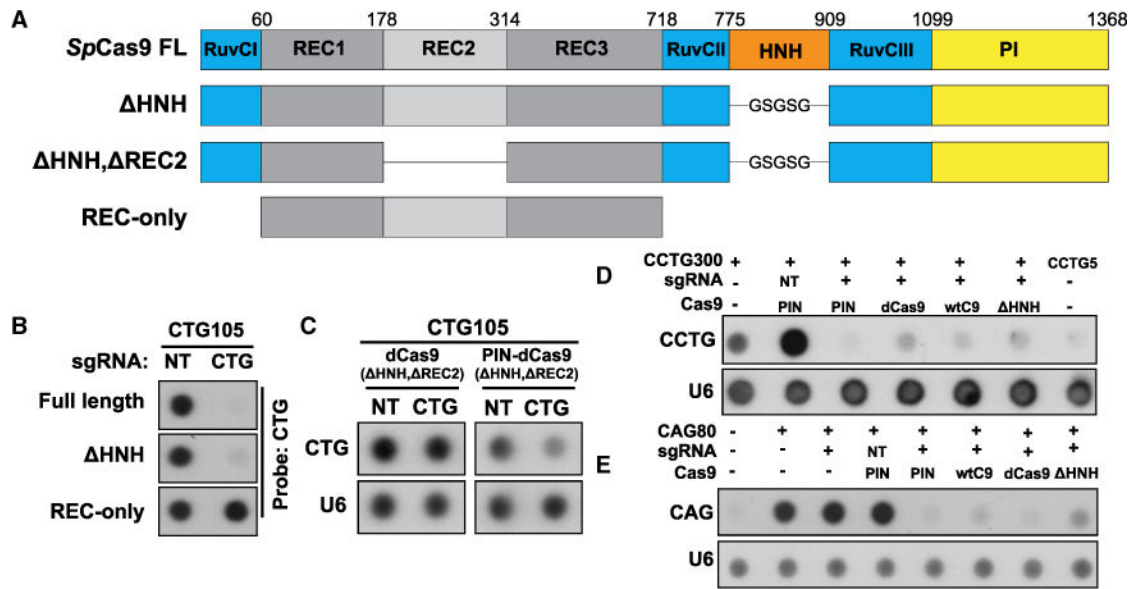See also Figure S2 and Tables S1 and S2.

Figure 4. Truncated Forms of Cas9 Maintain the Ability to Eliminate Repeat Expansion RNAs

(A) The domain structure of full-length (FL) Cas9 and three truncation constructs lacking the HNH domain (ΔHNH), the REC2 and HNH domain (ΔHNH, ΔREC2), or composed of the REC-lobe only (REC-only).

(B) RNA dot blot assay of $(CUG)^{exp}$ levels in COS-M6 cells transfected with $(CTG)^{105}$, CUG-targeting (CUG) or non-targeting (NT) sgRNA, and full-length, ΔHNH, or REC-only dCas9 fused to a PIN domain.

(C) RNA dot blot assay of $(CUG)^{exp}$ levels in COS-M6 cells transfected with $(CTG)^{105}$, CUG-targeting (CUG) or non-targeting (NT) sgRNA, and ΔHNH, ΔREC2 dCas9 truncation with and without a fused PIN RNA endonuclease domain.

(D) RNA dot blot assay of $(CCUG)^{exp}$ levels in COS-M6 cells transfected with $(CCTG)^{105}$, CCUG-targeting or non-targeting (NT) sgRNA, and ΔHNH fused to PIN compared to other forms of Cas9 (PIN-dCas9, dCas9, and wtCas9).

(E) RNA dot blot assay of $(CAG)^{exp}$ levels in COS-M6 cells transfected with $(CAG)^{105}$, CAG-targeting (CAG) or non-targeting (NT) sgRNA, and ΔHNH fused to PIN compared to other forms of Cas9 (PIN-dCas9, dCas9, and wtCas9).
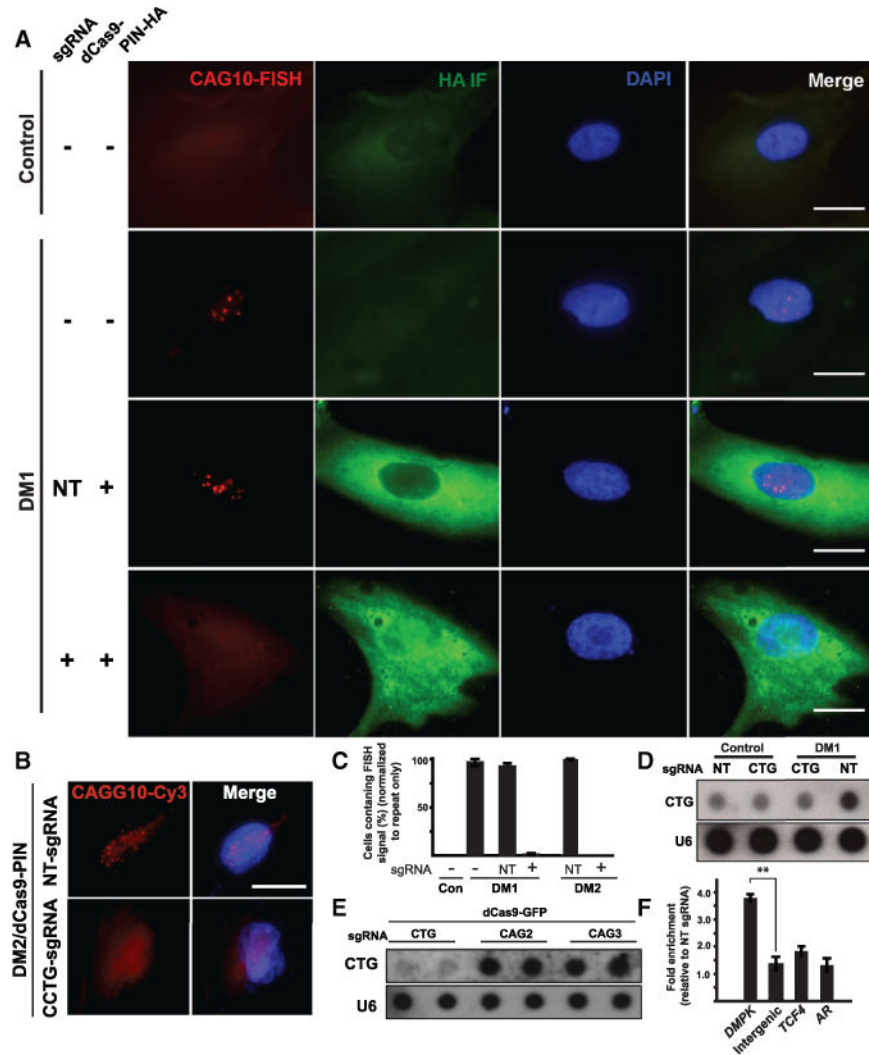
See also Table S2.

Figure 5. Degradation of CUG^{exp} RNA in DM1 Patient Myoblasts and CCUG^{exp} RNA in DM2 Patient Fibroblasts

(A) RNA-FISH with (CAG)^{10} probes recognizing (CUG)^{exp} RNA and HA immunofluorescence in control and DM1 primary myoblasts transduced with no sgRNA (−), non-targeting sgRNA (NT), or CUG-targeting sgRNA (CUG) with (+) or without (−) PIN-dCas9-HA. Scale bars in (A) and (B) are 20 μm.

(B) RNA-FISH with (CAGG)^{10} probes recognizing (CCUG)^{exp} RNA and HA immunofluorescence in control and DM2 primary fibroblasts transduced with non-targeting sgRNA (NT), or CCUG-targeting (CCUG) sgRNA with PIN-dCas9-HA.

(C) Quantification of RNA foci in patient-derived myoblast cells transduced with non-targeting (NT) or MRE-targeting sgRNAs (+) and PIN-dCas9. Cells containing at least one RNA focus are considered positive for MRE RNA. Error bars denote SDs determined from 3 biological replicates counting 100 patient cells each.

(D) Dot blot assay to detect (CUG)^{exp} in control and DM1 primary myoblasts transduced with non-targeting sgRNA (NT) or CUG-targeting sgRNA (CUG) and PIN-dCas9.

(E) RNA dot blot assay of (CUG)^{exp} RNA levels in DM1 patient myoblasts transduced with CTG-targeting sgRNA or template strand (CAG)-targeting sgRNAs, andCas9-GFP.

(F) RIP-PCR for PIN-dCas9 in DM1 patient cells. The fold enrichment for each primer pair (*DMPK*, *TCF4*, *AR*, and intergenic) was calculated for the CUG-targeting sgRNA relative to the non-targeting sgRNA.

See also Figure S2 and Tables S1 and S2.

Figure 6. MRE Elimination Reverses Downstream Molecular Phenotypes Associated with DM1 and HD

(A) RNA-FISH and EGFP fluorescence in COS-M6 cells transfected with plasmid (CTG)[105], MBNL1-EGFP, and non-targeting or CTG-targeting sgRNA. Scale bars in (A) and (B) are 20 μm.

(B) RNA-FISH and MBNL1 immunofluorescence in primary DM1 patient myoblasts transduced with PIN-dCas9 and a non-targeting (NT) sgRNA (top) or CUG-targeting sgRNA (bottom).

(C) Quantification of MBNL1 nuclear foci after RCas9 treatment with either non-targeting (NT) or CUG-targeting (CUG) sgRNA in primary DM1 patient myoblasts. Control

myoblasts from an unaffected individual are shown for comparison. Error bars denote SD determined from three biological replicates of 100 transfected cells each.

(D) RT-PCR to evaluate alternative splicing in primary DM1 patient myoblasts transduced with PIN-dCas9 and non-targeting (NT) or CUG-targeting sgRNA.

(E) Western blot for polyglutamine protein (polyQ) from extracts COS-M6 cells transfected with (+) or without (−) $(CAG)^{80}$ plasmid, and either with (+) or without (−) CAG-targeting sgRNA or non-targeting sgRNA (NT) and the indicated dCas9 constructs.

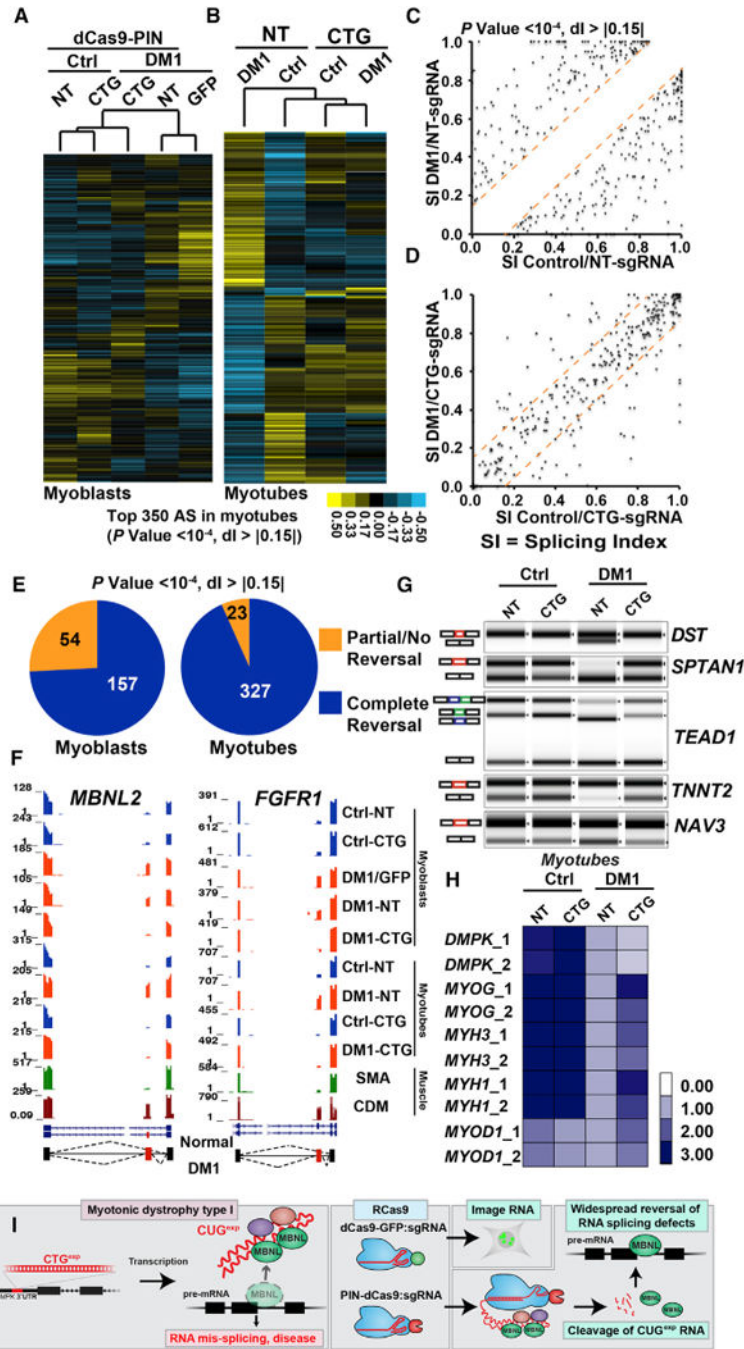See also Figures S2 and S6 and Tables S1, S2, and S5.

Figure 7. CUG-Targeting RCas9 Corrects Transcriptome-wide DM1-Related Alternative Splicing Defects

(A) Hierarchical clustered heatmap of exon inclusion indices of top 350 differential AS events in control and DM1 myoblasts.

(B) Hierarchical clustered heatmap of exon inclusion indices of top 350 differential AS events in control and DM1 myotubes.

(C) Scatterplot showing exon inclusion or splicing indices (SI) of top 350 differential AS events of control and DM1 myotubes transduced with non-targeting sgRNA (NT) and PIN-dCas9.

(D) Scatterplot showing exon inclusion or splicing indices (SI) of top 350 differential AS events of control and DM1 myotubes transduced with CUG-targeting sgRNA (CUG) and PIN-dCas9.

(E) Pie charts showing complete and partial/no reversal of top 350 DM1-related AS events after treatment with CUG-sgRNA and Cas9-PIN in myoblasts (left) and myotubes (right).

(F) RNA-seq data represented on UCSC genome-browser tracks of *MBNL2* exon 6 (left) and *FGFR1* exon 3 (right) showing DM1-related mis-splicing events in myoblasts or myotubes from unaffected individuals (ctrl) or DM1 patients (DM1), transduced with lentivirus expressing dCas9-PIN and either non-targeting (NT) or CUG-targeting sgRNA (CUG), or GFP only, as indicated.

(G) RT-PCR of candidate AS events uncovered by RNA-seq analysis.

(H) Expression levels of the indicated muscle differentiation markers evaluated by RNA-seq in myotubes differentiated from DM1 patient or healthy control myoblasts treated with non-targeting or CUG-targeting sgRNAs.

(I) Schematic description of the therapeutic mechanism of the RCas9 system in the context of myotonic dystrophy type I.

See also Figures S2, S4, and S5 and Tables S2–S5.