Genome Editing of Expanded CTG Repeats within the Human DMPK Gene Reduces Nuclear RNA Foci in the Muscle of DM1 Mice

Mirella Lo Scrudato, Karine Poulard, Célia Sourd, Stéphanie Tomé, Arnaud Klein, Guillaume Corre, Aline Huguet, Denis Furling, Geneviève Gourdon, Ana Buj-Bello

Genome editing of expanded CTG repeats within the human *DMPK* gene reduces nuclear RNA foci in muscle of DM1 mice

Mirella Lo Scrudato,[1] Karine Poulard,[1] Célia Sourd,[1] Stéphanie Tomé,[2] Arnaud F. Klein,[3] Guillaume Corre,[1] Aline Huguet,[2] Denis Furling,[3] Geneviève Gourdon,[2] and Ana Buj-Bello[1¶]

[1]Genethon, UMR_S951 Inserm, Univ Evry, Université Paris Saclay, Evry, France; [2]INSERM UMR 1163, Institut Imagine, Université Paris Descartes-Sorbonne Paris Cité Paris, France; [3]Sorbonne Université, INSERM, Association Institut de Myologie, Centre de Recherche en Myologie, F-75013 Paris, France

Correspondence should be addressed to A.B.B. (abujbello@genethon.fr)

Correspondence: Dr. Ana Buj Bello, Genethon, INSERM UMR_S951, 1 bis rue de l'Internationale, 91000 Evry, France. Tel: 33 (0) 1 69 47 28 28; Fax: 01 60 77 86 98; e-mail: abujbello@genethon.fr

Short Title: *In vivo* genome editing for DM1

ABSTRACT

Myotonic dystrophy type 1 (DM1) is caused by a CTG repeat expansion located in the 3' untranslated region of the *DMPK* gene. Expanded *DMPK* transcripts aggregate into nuclear foci and alter the function of RNA-binding proteins, leading to defects in the alternative splicing of numerous pre-mRNAs. To date, there is no curative treatment for DM1. Here, we investigated a gene editing strategy using the clustered regularly interspaced short palindromic repeat (CRISPR)-Cas9 system from *Staphylococcus aureus* (Sa) to delete the CTG repeats in the human *DMPK* locus. Co-expression of SaCas9 and selected pairs of single guide RNAs (sgRNAs) in cultured DM1 patient-derived muscle line cells carrying 2600 CTG resulted in targeted DNA deletion, ribonucleoprotein foci disappearance and correction of splicing abnormalities in various transcripts. Furthermore, a single intramuscular injection of recombinant AAV vectors expressing CRISPR-SaCas9 components in tibialis anterior muscle of DMSXL mice decreased the number of pathological RNA foci in myonuclei. These results establish the proof of concept that genome editing of a large trinucleotide expansion is feasible in muscle, and may represent a useful strategy to be further developed for the treatment of myotonic dystrophy.

Keywords: CRISPR-Cas9; DMPK; gene therapy; myotonic dystrophy; nucleotide repeat disorders

INTRODUCTION

Myotonic dystrophy type 1 (DM1) is the most common form of adult muscular dystrophy, with an estimated prevalence of 1 in 8,000 individuals. The disease is autosomal dominant and characterized by multisystemic symptoms like myotonia, muscle weakness, cardiac conduction defects, cataracts, insulin resistance and cognitive abnormalities [1]. DM1, also called Steinert disease, is caused by a CTG repeat expansion in the 3' untranslated region (UTR) of the *DMPK* gene coding for a serine/threonine kinase that is mainly expressed in smooth, skeletal and cardiac muscles [2-7].

The number of *DMPK* CTG repeats usually ranges from 5 to 37 in unaffected individuals, and from 51 to several thousands in DM1 patients [1]. The length of the expansion correlates with clinical severity and inversely to disease onset. Expanded CTG repeats are unstable in the germline, leading to the phenomenon of anticipation in members of the same family, i.e. occurrence of the disorder at progressively earlier ages in successive generations, and in somatic cells, resulting in high levels of mosaicism among different tissues, which plays a primary role in DM1 severity [8-10].

Several *in vitro* and *in vivo* studies elucidated the disease mechanism, which is mainly mediated by a toxic gain-of-function of RNA transcripts (for review see [11]). Transcription of mutated *DMPK* generates mRNAs with long CUG repeats, which accumulate in the nucleus and form stable ribonucleoprotein aggregates called foci, interfering with at least two antagonistic protein families that regulate alternative splicing, the muscleblind-like (MBNL) and CUGBP/Elav-like (CELF) protein family [11, 12]. Splicing regulators of the MBNL family, which are able to bind C/CUG sequences, are sequestered within the foci, and therefore functionally downregulated [13], and CELF proteins are upregulated through protein stabilization [14, 15]. These alterations result in

aberrant expression of embryonic splicing profiles in adult tissues; forty-two mispliced events were validated in muscle of DM1 patients, such as for the insulin receptor (IR) and cardiac troponin T (cTNT), and subsequent cellular dysfunction [16-20].

Several approaches that act at different levels of the pathological cascade, mainly targeting mutated *DMPK* transcripts, have been investigated but no efficacious treatment is currently available for DM1 patients (see review [21]). The experimental therapies that have been addressed include: i) degradation of mutated *DMPK* transcripts by small nuclear antisense RNA [22, 23], gapmer antisense oligonucleotides (ASOs) [24], an artificial RNA endonuclease [25], or deactivated Cas9 nuclease (dCas9) [26]; ii) viral vector-mediated *Mbnl1* overexpression [27]; and iii) interference with foci formation through antisense morpholino oligonucleotides (MOs), small molecules and peptides that disrupt MBNL1-C/CUG interaction [28-30]. A gapmer ASOs (IONIS-DMPKRx), based on RNase H-mediated degradation of heteroduplex ASO-DMPK RNA, was tested in a phase 1/2a clinical trial in adult patients with DM1 (https://clinicaltrials.gov/; NCT02312011), and, although well tolerated, the drug levels in muscle did not achieve the desired therapeutic benefit.

Gene editing represents an alternative strategy to correct mutations responsible for inherited disorders. The CRISPR-Cas9 system, identified in bacteria as part of the bacterial immune system [31], is a powerful tool that has been adapted for medical applications [32, 33]. Recent studies demonstrated that delivery of recombinant adeno-associated viral (rAAV) vectors expressing CRISPR-Cas9 components was able to restore dystrophin expression in muscles of Duchenne muscular dystrophy mouse models [34-36], opening new frontiers for the cure of neuromuscular disorders (see review [37]).

Here, we investigated CRISPR-Cas9 from *Staphylococcus aureus* (Sa) to identify sgRNAs capable of cutting efficiently regions flanking the *DMPK* CTG repeat. SaCas9 is a small size nuclease that fits into a rAAV vector and may have high cleavage activity in human cells [38]. We performed experiments in DM1 patient-derived immortalized myoblast cells and observed an efficient deletion of the CTG repeat tract for a pair of sgRNAs, leading to stable absence of nuclear foci and reversion of splicing abnormalities. Notably, no mutations were found in potential off-targets of each sgRNA. Next, we evaluated this CRISPR-Cas9 approach in DMSXL mice, a DM1 mouse model carrying a human *DMPK* gene with ~1200 CTG repeats under the regulation of its own promoter [39, 40]. Our results show that a single intramuscular injection of two rAAV9 vectors, expressing the nuclease SaCas9 and the most efficient pair of sgRNAs, were also able to delete the CTG repeat in muscle fibers. Consequently, a reduction in the number of myonuclei containing pathological ribonucleoprotein foci was observed. This study represents the first proof-of-concept of *in vivo* genome editing for DM1 and, with further development, provides novel perspectives for the treatment of nucleotide repeat disorders that affect muscles.

RESULTS

CRISPR-Cas9 Strategy for DM1 Disease

To selectively remove the CTG repeat expansion of the *DMPK* gene, we initially used the CRISPR-Cas9 system from *Neisseria meningitidis* (Nm) [41], as the size of NmCas9 is small and suitable for vectorization in rAAV vectors. We selected four sgRNAs that target regions flanking the CTG repeat and tested their deletion efficiency in HeLa cells (Figure S1, [42]). PCR amplification of the *DMPK* genomic region encompassing the sgRNA target sites resulted in very weak intensity bands of the deleted fragments compared to the undeleted PCR products, suggesting a low cutting efficiency for NmCas9 (Figure S1C).

Based on these results, we pursued this study by testing the CRISPR-Cas9 system from *Staphylococcus aureus* (Sa) [43], which also encodes a small size nuclease. We restricted the target region to the portion of the *DMPK* 3' UTR between the stop codon of the gene and the polyadenylation signal, to avoid interference with the coding sequence and mRNA maturation. First, we selected 15 sgRNAs targeting this region and used HeLa cells for screening. We quantified the frequency of indels (insertions and deletions) by PCR amplification and sequencing of the respective sgRNA genomic targets, followed by TIDE analysis (https://tide.nki.nl/), a method based on the recovery of indels spectrum from the sequencing chromatogram to quantify the proportion of templated editing events, including point mutations [44]. The sgRNAs binding the region upstream the CTG repeat (Figure 1A; Table S1; sgRNA 1, 4, 7 and 8) induced indels at high frequencies, with values between 42% and 47%. However, downstream sgRNAs resulted in scattered indel values that ranged from 1% to 48.3%. For sgRNA 12, 13 and 17, we generated two versions of the protospacer, with either 21 (sgRNA

12A, 13A, 17A) or 24 (sgRNA 12B, 13B, 17B) nucleotides. Interestingly, the cutting efficiency of the shorter sgRNAs 12A and 13A was higher than the longer forms.

In order to test the ability of sgRNA couples to delete the CTG repeat, we generated constructs expressing SaCas9 and two sgRNAs in tandem, targeting upstream and downstream regions of the repeat. Based on the best single cutting efficiency (Table S1), we selected 4 sgRNAs targeting each side of the CTG repeat, and assorted them in 16 couples (Figure 1A, sgRNA targets in black). We tested these constructs in HeLa cells by transfection and performed genomic PCR analysis using primers F and R, which anneal regions upstream and downstream the more distant sgRNAs 1 and 23 (Figure 1A; Table S4). PCR fragments with deleted CTG repeats were observed for all sgRNA couples (Figure 1B, bands of ~0.2-0.8 kb, black arrow); only two couples, 1-13A and 1-23, showed deleted PCR products of weaker intensity. Notably, the intensity of the undeleted bands compared to the edited fragments was negligible (Figure 1B, band of ~0.9 kb, white arrow), indicating that the selected sgRNA couples drive efficient deletion of the CTG repeat flanking regions.

Genome Editing of *DMPK* CTG Repeats in Human DM1 muscle cells

Next, we aimed to test the ability of CRISPR-SaCas9 to delete a pathogenic CTG repeat expansion. We chose as DM1 *in vitro* model, an immortalized myoblast cell line derived from a patient carrying 2600 CTG repeats in the *DMPK* gene [45]. This cell line reproduces the most important cellular hallmarks of the disease, in particular, presence of nuclear foci and splicing defects in various transcripts.

As these myoblasts were hardly transfected (about 30% for a control GFP plasmid), we delivered CRISPR-SaCas9 by lentiviral vectors, which are known to infect with high efficiency cultured

muscle cells. In particular, we designed a dual vector system to test various combinations of sgRNAs with SaCas9 (Figure 2A). Taking into consideration the following criteria for each sgRNA: 1) individual frequency of indels at the target site, 2) number of predicted off-targets and 3) distance of the target site to the CTG repeat extremities, we selected couples 4-23, 4-12A, 8-12A and 8-23 for *in vitro* studies (Figure 1A and 1B; Table S1). DM1 patient-derived muscle line cells were transduced with increasing equal MOI (multiplicity of infection) of both SaCas9 and sgRNAs lentiviral vectors, and deletions of the CTG repeat region were analyzed by genomic PCR from the bulk cell population (Figure 2B). All four sgRNA couples were able to delete the targeted region, although 4-23, 4-12A, and 8-23 appeared more efficacious as the deleted band was visible at low MOI (Figure 2B; faint band at MOI=5). The amount of edited PCR fragments increased proportionally to the MOI tested. To notice, the expanded allele was not PCR amplified due to the size of the 2600 CTG repeat (corresponding to a >8 kb fragment). Thus, we estimated the percentage of CTG repeat deletion by comparing the intensity of edited products, originating from the two alleles, to that of undeleted bands from the normal allele (Figure 2C; % DEL), and sgRNA couple 4-23 appeared as the most efficient one in deleting the targeted genomic region (Figure 2B and 2C).

We then monitored the presence of nuclear foci in lentiviral-transduced DM1 myoblasts by fluorescent in situ hybridization (FISH) with a (CAG)7 probe. CRISPR-SaCas9-mediated deletion of the expanded CTG repeat region resulted in a reduction and/or completely disappearance of nuclear foci in treated cells. In particular, the percentage of cultured DM1 cells without foci was higher in presence of sgRNA couple 4-23 and reached 19.2 and 27.5% at MOI 50 and 100 versus 4.4% and 6.2% for cells treated only with one of the two lentiviral vectors (Figure 2D). The other couples of sgRNA resulted in similar (sgRNA 4-12A) or lower (sgRNA

8-23 and 8-13A) percentage of cells without foci (Figure S2). Overall, these data demonstrate the ability of the selected sgRNAs to delete CTG repeats from the 3' UTR of the *DMPK* gene, and likely from the pathogenic expanded allele.

To better analyze genome editing events in DM1 cells, we isolated and characterized myoblast clones transduced with MOI 50 of lentiviral vectors expressing SaCas9 and $sgRNA_{4\text{-}23}$-GFP. We analyzed 50 GFP positive clones by FISH and identified 5 clones negative for the presence of nuclear foci, which were also positive for the expression of SaCas9 (Figure 3A, DM1-Delta). DM1 clones expressing only SaCas9 (DM1-Cas9) or sgRNAs 4-23 (DM1-sgRNA), and non-transduced cells were used as controls. We selected several DM1-Delta clones (10, 3, 17 and 22) for further analysis, and confirmed by genomic PCR the deletion of the CTG repeat in the *DMPK* locus (Figure 3B). For clones 10 and 3, only one PCR fragment was amplified, suggesting a biallelic deletion, whereas for clones 17 and 22, undeleted (upper band) and deleted (lower band) PCR products were observed. The intermediate bands observed in the gel for clones 17 and 22 are heteroduplexes of PCR amplicons with and without CTG repeat deletions, as shown by denaturation and renaturation of a mixture of PCR amplicons (Figure S4). Sequencing of the shorter PCR fragments from the four clones (Figure 3B, band of ~0.4-0.5 kb) confirmed the complete deletion of the CTG repeat for clones 10, 17 and 22, and a partial deletion with 8 remaining nucleotides of the CTG repeat and the upstream 99 nucleotides of the genomic sequence for clone 3 (Figure 3C and 3D). To notice the cutting and joining position was not always at position between nucleotide N3 and N4 upstream the PAM sequence, as a variable number of nucleotides upstream the expected cutting sites was deleted. In addition, longer PCR products of DM1-Delta clones 17 and 22 (Figure 3B, upper band of ~1 kb) revealed microdeletions of variable length, including the PAM sequence, in the target sites but intact

unexpanded CTG repeat (Figure 3D). These results suggest that non-synchronized cuts occurred at each target site of the normal *DMPK* allele, which were repaired by non-homologous end-joining (NHEJ), and the expanded allele was properly edited. Control DM1-sgRNA clones, still containing nuclear foci, did not show any indels at the target sites (Figure 3D). Interestingly, we did not find off-target indels in DM1-Delta clones (see Supplemental Information).

In order to have a clear evidence of the deletion of the expanded CTG repeat in the *DMPK* gene, we performed southern blot analysis on *Eco*RI digested genomic DNA, which was hybridized with a probe annealing a 1.4 kb region of the *DMPK* 3' UTR (Figure 3E). In DM1 cells, the presence of 2600 CTG repeats in the expanded allele (CTG_{exp}), and 13 CTG in the normal allele (ctg), resulted in two fragments of 17.4 and 8.6 kb, respectively, compared to 9.6 and 8.6 kb bands in immortalized myoblasts from a control individual (Ctrl) containing 5 and 14 CTG repeats. In all three DM1-Delta clones analyzed, the genomic excision of the expanded CTG repeat region was revealed by the absence of the 17.4 kb *Eco*RI band in the southern blot (Figure 3F and 3G). Moreover, because of the presence of a 1 kb *Alu* polymorphism in the CTG_{exp} allele but not in the normal allele of this DM1 cell line [46], it was possible to distinguish the edited fragments originating from the expanded (ΔCTG_{exp}, 9.1 kb) and normal (Δctg, 8.1 kb) alleles. Thus, southern blot analysis confirmed a biallelic deletion in clone 10 ($\Delta CTG_{exp}/\Delta$ctg; 9.1 kb and 8.1 kb bands), and a monoallelic deletion in clones 17 and 22 (ΔCTG_{exp}/ctg; 9.1 kb and 8.6 kb bands). Altogether, these results show that CRISPR-SaCas9 is able to excise expanded CTG repeats in the *DMPK* 3' UTR of human DM1 cells. To notice, in clones 17 and 22 the two *Eco*RI fragments corresponding to deleted expanded CTG (ΔCTG_{exp}) and undeleted unexpanded CTG (ctg) appeared as a unique diffused band instead two well-separated bands because they differ only ~ 0.5 kb.

Correction of Alternative Splicing Defects in DM1 Edited Muscle Cells

Sequestration of MBNL splicing factors in nuclear foci of DM1 cells leads to alterations in the alternative splicing of numerous pre-mRNAs, some of which are reproduced in differentiated muscle cells in culture [45]. We therefore assessed the splicing pattern of LIM domain binding 3 (*LDB3)* exon 11, Sarco/endoplasmic reticulum calcium ATPase 1 (*ATP2A1*) exon 22, Muscleblind like splicing regulator 1 (*MBNL1*) exon 7, Duchenne muscular dystrophy (*DMD*) exon 78, insulin receptor (*IR*) exon 11 and Bridging integrator 1 (*BIN1*) exon 11 in edited DM1-Delta clones compared to DM1, DM1-sgRNA, and control (Ctrl) cells (Figure 4). RT-PCR analyses revealed that the splicing profile of these transcripts in DM1-Delta myotubes was significantly different than in untreated DM1 cells, and in general comparable to control myotubes, with some variations for *MBNL1* and *BIN1* splicing in clone 10. Therefore, CRISPR-SaCas9-mediated excision of the expanded *DMPK* CTG repeat region resulted in the correction of the splicing defects of DM1 cells. Interestingly, the levels of *DMPK* mRNA after CTG deletion in DM1-Delta clones were similar or even higher than in untreated DM1 cells (Figure S5A), suggesting that the excision of the CTG repeat in the 3'UTR did not affect the stability of this transcript.

Deletion of Expanded CTG Repeats in Muscle Fibers of DMSXL Mice

To assess whether the CRISPR-SaCas9 system was able to delete the CTG repeats *in vivo*, we performed experiments in DMSXL transgenic mice, a mouse model of myotonic dystrophy type 1 that carries a copy of the human *DMPK* gene with ~1200 CTG repeats in the 3' UTR, and exhibits features of the disease, such as presence of nuclear RNA foci in muscles [47]. We

generated serotype 9 rAAV vectors (rAAV9), which are known to transduce efficiently myofibers [48], containing expression cassettes for either SaCas9 under the muscle-specific SPc5-12 promoter or the couple of sgRNAs 4-23 under the U6 promoter (Figure 5A).

First, we evaluated the feasibility of this approach by intramuscular delivery of the vectors in heterozygous DMSXL mice at various doses and ages, as these mice do not display the high mortality rate observed in homozygous mice. A combined dose of 0.6×10^{11} viral genomes (vg) and 1×10^{11} vg was injected into the left tibialis anterior (TA) muscle of HTZ mice at 3 and 6 weeks of age, respectively (the injected dose contained a mixture of equal vg of each vector per muscle). The contralateral TA muscle was injected with equivalent volume of PBS as control. The effect of vector administration was analyzed 4 weeks post-injection, and hematoxylin and eosin (H&E) staining of TA sections did not reveal signs of muscle degeneration (only few and small foci of inflammatory infiltrates were observed in muscles of 2 out 7 mice of the older group of treated animals, data not shown). To evaluate CRISPR-SaCas9 activity in muscles, we PCR amplified the targeted genomic region from the two groups of animals. PCR amplicons corresponding to the edited fragment (0.4 kb) were detected in TA muscles co-injected with rAAV9-SaCas9 and rAAV9-sgRNA$_{4\text{-}23}$, but not in the contralateral PBS-injected muscles (Figure S3A). Moreover, a specific PCR protocol was used to amplify the genomic region containing the undeleted CTG repeats (4.5 kb), which was detected only in PBS-injected TA (Figure S3B).

Based on these results, we evaluated whether administration of SaCas9- and sgRNA$_{4\text{-}23}$-expressing vectors could also delete the CTG repeat expansion in muscles of homozygous (HMZ) DMSXL mice, and correct pathological signs of the disease. For that purpose, rAAV9 vectors were delivered intramuscularly at a dose of 1×10^{11} total vg in the left TA muscle of 5-9

week-old homozygous mutant mice, and an equal volume of PBS was injected in the contralateral muscle. Four weeks after injection, immunofluorescence analysis showed, as expected, the localization of SaCas9 within nuclei of muscle fibers (Figure 5B). The eGFP-K reporter that localizes to the membrane of myonuclei was used to visualize indirectly sgRNA expressing nuclei. In AAV-treated muscles, 21% and 76% of myonuclei were positive for SaCas9 and GFP, respectively, and 18% were positive for both. The relatively low percentage of SaCas9 positive myonuclei was similar by using anti-HA tag and anti-SaCas9 antibodies (data not shown). Since the same dose of each vector was injected in muscles, these results suggest that the sensitivity of antibodies against the HA tag or SaCas9 by immunohistology were quite low. H&E staining of muscle cross-sections from HMZ mice indicated the absence of significant damages in muscle tissue (only small foci of inflammatory infiltration in 1 out of 10 homozygous, and 1 in 5 wild type mice, data not shown). PCR amplification of the *DMPK* 3' UTR region resulted in small fragments corresponding to deleted CTG repeats, which was also confirmed by Sanger sequencing (Figure 5C and 5D). Interestingly, in muscle the target site was cut at nucleotide N3 upstream the PAM sequence, contrary to more variable cutting positions observed in edited DM1 cells in culture (Figure 5E and 3D). In order to further investigate the cutting pattern of SaCas9-sgRNA$_{4\text{-}23}$ in DMSXL skeletal muscle, we performed deep sequencing analysis of PCR amplicons from 8 DMSXL HMZ mice, and compared the results to that obtained from the bulk population of cultured DM1 myoblasts transduced with lentiviral vectors expressing the same CRISPR-Cas9 components, and appropriate controls (Figures 6A-C). PCR amplicons were encompassing the CTG repeat deletion (DEL) and the sgRNA$_4$ or sgRNA$_{23}$ target sites (sgRNA4 and sgRNA23) depending on whether the repeat deletion had occurred or not. We found that the percentage of reads with indels was lower in skeletal muscle compared to

cultured cells, 9.6% versus 44.7% in DEL, 7.6% versus 30.8% for sgRNA4 and 6.3% versus 17.3% in sgRNA23, respectively. These results suggest that CRISPR-SaCas9-mediated genome editing results in low levels of indels in skeletal muscle.

Finally, we analyzed the effect of genomic CTG repeat deletion in *DMPK* mRNA levels and found that transcript levels did not change in tibialis anterior muscle treated with SaCas9-sgRNA$_{4\text{-}23}$ compared to the contralateral uninjected muscle (Figure S5B).

Reduction of nuclear foci in genome edited DM1 skeletal muscle

Next, we assessed the effect of CRISPR-SaCas9-mediated deletion of expanded CTG repeats in RNA foci of myonuclei from DMSXL mice. For that purpose, we analyzed tibialis anterior muscle cross-sections by FISH using the (CAG)7 probe and laminin immunostaining (α-LMN) to clearly delimit the basal lamina of muscle fibers and distinguish between nuclei located inside from outside myofibers (Figure 7A). Analysis of confocal images showed a statistically significant reduction of myonuclei containing foci in rAAV9-treated TA muscle. The percentage of nuclei with foci located in muscle fibers decreased by 24.17% four weeks after co-delivery of rAAV9-SaCas9 and rAAV9-sgRNA$_{4\text{-}23}$ compared to PBS-injected contralateral muscles (Figure 7B; $P < 0001$; N=10 mice). The number of myonuclei per fiber, calculated as ratio between total number of myonuclei and myofibers, was similar between treated and untreated TA muscles (Figure 7C; WT-PBS: 0.94±0.22; WT-AAV: 1.02±0.07; HMZ-PBS: 0.85±0.17; HMZ-AAV: 0.75±0.15), indicating that the reduction of myonuclei containing foci in the group of rAAV9-treated muscles resulted from the excision of expanded CTG repeats rather than changes in the total number of nuclei within myofibers.

In conclusion, we have shown that local administration of rAAV9 vectors expressing components of the CRISPR-Cas9 system can excise long CTG repeats *in vivo* and reduce pathological RNA foci within myonuclei of a mouse model of myotonic dystrophy type 1.

DISCUSSION

The possibility of editing the genome has opened new perspectives for the treatment of inherited diseases. Here, we investigated whether the CRISPR-Cas9 system is efficient in correcting the genetic defect of myotonic dystrophy type 1 both *in vitro* and *in vivo*. Our study demonstrates that intramuscular administration of rAAV vectors that express Cas9 and selected sgRNAs in DM1 mice can excise the expanded CTG repeats in the human *DMPK* 3' UTR and rescue pathological signs of the disease, establishing the first proof of concept that *in vivo* gene editing with active nucleases is feasible for DM1.

We used CRISPR-Cas9 from *Staphylococcus aureus* to target the flanking regions of the CTG repeat tract and performed experiments in cultured cells to select the best sgRNA candidates for *in vivo* studies. We found that the genomic region downstream the CTG repeats was more difficult to cut than the upstream region. Among eight tested sgRNAs (plus three variants) only three of them, sgRNA 12A, 13A and 23, showed individual high indels frequencies (>30%) in cells. The region downstream the CTG repeats could be less accessible to the sgRNA-nuclease complex due to the presence of DNA secondary structures that may interfere with PAM recognition, sgRNA-DNA heteroduplex formation and/or Cas9 activity, as the conformation of the chromatin appears to influence CRISPR-Cas9-mediated genome editing, being more efficient in euchromatic than heterochromatic DNA regions [49]. We selected 4 pairs of sgRNAs from 15 designed sgRNAs that target the *DMPK* 3' UTR to test gene editing in DM1 patient-derived immortalized myoblasts *in vitro*, and they were able to delete the targeted region with high efficiency, reduce nuclear RNA foci and revert the aberrant splicing of several transcripts. Other studies reported recently a similar approach by using CRISPR-Cas9 from *Staphylococcus pyogenes* (Sp) in cell cultures [50-52], however, the use of SaCas9 could be more advantageous for a

translational perspective as it is a smaller endonuclease that can be easily packaged into recombinant AAV vectors for *in vivo* studies. This would allow the inclusion of larger regulatory sequences in the expression cassette, which might be of interest for driving tissue-specificity of the nuclease activity, and/or sequences for sgRNA expression as all-in-one vectors.

As alternative strategies, Gao and colleagues inserted a polyA signal upstream the *DMPK* CTG repeats in human DM1 induced pluripotent stem (iPS) cells by TALEN to prevent the transcription of toxic mutant transcripts, which resulted in foci disappearance and reversion of aberrant splicing in DM1 differentiated neural stem cells and cardiomyocytes [53]. Genome editing has also been applied to contract and therefore shorten CTG trinucleotide repeats in cellular models by inducing either double-strand DNA breaks within the repeat tract with meganucleases, zinc-finger nucleases (ZNF) and transcription activator-like effector (TALE) nucleases (TALEN) [54, 55], or single-strand DNA breaks with a CRISPR-Cas9 D10A nickase in a GFP-based chromosomal reporter system [56]. Finally, two other recent approaches exploited deactivated forms of Cas9, which do not cut DNA, to mediate inhibition of *DMPK* gene transcription [57] and degradation of toxic *DMPK* transcripts [26]. The relevance of these various approaches for clinical translation and therapeutic intervention in patients with myotonic dystrophy type 1 remains to be assessed.

So far, all *DMPK* gene editing studies with active nucleases were performed in cellular models or cells derived from DM1 patients. Here, we report that rAAV vectors expressing SaCas9 under the SPc5-12 muscle-specific promoter and U6-driven sgRNA 4 and 23 can excise the expanded repeat tract in skeletal muscle of DMSXL mice, which carry a human *DMPK* transgene with ~1,200 CTG repeats [39, 40]. The feasibility of genome editing in postnatal muscle was previously reported in mdx mouse models of Duchenne muscular dystrophy, a disease caused by mutations

in the large *DMD* gene [34-36]. In these mice, CRISPR-Cas9-mediated NHEJ was able to remove the mutated exon 23 from the dystrophin gene with low cutting efficiency, which ranged from ~2 % to 8 %, but resulted in the expression of a functional truncated dystrophin protein and significant increase in muscle fiber size and strength. In the context of a disease with a dominant form of inheritance and large nucleotide expansions, *in vivo* gene editing might be more challenging as therapeutic approach. Dual rAAV vectors expressing CRISPR-Cas9 components appeared more efficient than "all-in-one" single vector approaches in mice [36, 58]. Therefore, in this study we administrated locally dual rAAV9 vectors into the tibialis anterior muscle of homozygous DMSXL mice when the pathology is already present. Four weeks after a single injection of 1.0×10^{11} total vg of vectors (0.5×10^{10} vg of each vector), SaCas9 and the sgRNA reporter GFP were expressed in the skeletal muscle with no or minimal inflammatory infiltrates, and resulted in the deletion of the expanded *DMPK* CTG repeat region. Deep sequencing of PCR products from the targeted genomic region revealed that the percentage of reads with indels was rather low (6% to 9%) in skeletal muscle of DMSXL HMZ mice compared to cultured DM1 cells (17% to 44%), indicating that differences in DNA repair mechanisms between postmitotic myofibers and mitotic myoblasts exist. In addition, no indels were found in 20 potential off-target sites of the selected sgRNAs. However, it has to be noted that a more comprehensive method, such as whole genome or exome sequencing, should be used in order to detect indels in other parts of the genome or other genomic modifications. Importantly, rAAV9-mediated CRISPR-SaCas9 delivery resulted in ~25% reduction in the number of myonuclei containing RNA foci, indicating that genome editing was efficacious in ameliorating a major histopathological hallmark of myotonic dystrophy. The splicing pattern of several transcripts (*Ir*, *Ttn*, *Mbnl1* and *Mbnl2*) were found normal in tibialis anterior muscle of DMSXL mice 4 weeks

after injection of CRISPR-SaCas9 vectors, however these results are inconclusive as no significant differences were observed in the relative expression of the spliced variants of these genes in TA muscle between untreated mutant and WT mice (data not shown), reflecting the very mild splicing phenotype of DMSXL mice. Gene editing had also no effect at the level of TA muscle weight and strength in homozygous mutant mice under these conditions. This is in contrast with other gene editing studies in muscle, in particular for Duchenne muscular dystrophy, where an amelioration was observed at the level of muscle histology and/or function in animal models [34-36, 58, 59], and reflects probably the challenge of reducing toxic RNA levels in a gain of function disease like DM1 instead of increasing the amount of a functional protein in a recessive disease like DMD. The threshold of foci reduction in tissues and the age of treatment for DM1 phenotype correction remain to be elucidated. Interestingly, this gene editing approach could be used not only for muscle but also for other tissues affected in the disease by changing the promoter specificity of the Cas9 transgene and evaluating various routes of administration.

In conclusion, we have established the proof-of-concept that CRISPR-SaCas9-mediated genome editing can efficiently delete the pathological CTG expansion from the human *DMPK* gene *in vivo* in skeletal muscle, and demonstrated that this approach can reduce RNA foci accumulation in myonuclei, which is a major pathological sign of the disease. Altogether, with further development, our study supports CRISPR-Cas9 based genome editing as a potential therapeutic approach for myotonic dystrophy type 1.

MATERIALS AND METHODS

Plasmids and Design of sgRNAs

The list of main constructs and primers employed in this study are reported in Tables S2 and S4.

Plasmids were constructed with traditional cloning strategies using inserts PCR amplified, or synthetically synthetized, or sub-cloned upon enzymatic digestion of other existing constructs. Plasmid encoding for *S. aureus* Cas9 derives from plasmid pX601-AAV-CMV::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA (MLS42, Addgene plasmid # 61591; [43]). EFS promoter was PCR amplified from a plasmid containing EF1-alpha promoter with primers F-XhoI-MreI-EFS (MLS63) and R-XmaI-NruI-EFS (MLS64) and cloned into *Xho*I/*Age*I site of promoterless pX601-AAV-::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA to obtain pAAV-EFS::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BsaI-sgRNA (MLS43).

The second cassette for sgRNA (U6::BbsI-sgRNA) was synthetically synthesized (GeneCust) using the same sequence of the existing cassette U6::BsaI-sgRNA of plasmid MLS42 but exchanging the sgRNA protospacer cloning site from *Bsa*I into *Bbs*I. Then, the *Bbs*I cassette was cloned into *Acc65*I site of plasmid MLS43, upstream and in tandem the first sgRNA cassette, to obtain the construct pAAV-EFS::NLS-SaCas9-NLS-3xHA-bGHpA;U6::BbsI-sgRNA;U6::BsaI-sgRNA (MLS47).

Sa sgRNA protospacers, with n ID number, were synthesized as couple of oligonucleotides forward and reverse and *in vitro* annealed prior their cloning into the restriction sites *Bbs*I or *Bsa*I of plasmid MLS47, to obtain derivative plasmids pAAV-EFS::NLS-SaCas9-NLS-3xHA-bGHpA;U6::n-DMPK-sgRNA;U6::BsaI-sgRNA and pAAV-EFS::NLS-SaCas9-NLS-3xHA-bGHpA;U6::n(up)-sgRNA;U6::n(dw)-sgRNA_DMPK, this last with sgRNAs targeting upstream (up) and downstream (dw) the CTG repeat (see construct MLS93 as example).

SaCas9 target sequences within the *DMPK* 3' UTR were screened by the programs CasBLASTR (http://www.casblastr.org/) and CRISPOR (http://tefor.net/crispor). The PAM sequence NNGRRT was used for the screening, with R= A or G (AYYCNN in the non-coding strand, with Y= T or C).

For each sgRNA protospacer, the number of potential off-targets was calculated by the program CasOFFinder (http://www.rgenome.net/cas-offinder/) based on the human genome "Homo sapiens (GRCh38/hg38) - Human (02 April 2014 Updated). The selection of sgRNA protospacers was done taking into consideration the respective number of potential off-targets and their target position within the *DMPK* 3' UTR region. Targets having potential off-targets carrying 3 or less mismatches were excluded. The length of the Sa sgRNA protospacer is between 21 and 24 nt. Whenever the protospacer did not start with a G, this nucleotide was added to the 5' of the sequence to optimize the U6-driven transcription (Table S1).

Cell Culture and Transfection Experiments

HeLa cells were cultured in Dulbecco's modified Eagle medium (DMEM) with high glucose and GlutaMAX (Invitrogen), supplemented with 10% Fetal Bovin Serum (FBS, Invitrogen). Immortalized control (C25-Cl48, abbreviated Ctrl) and DM1 (DM11-Cl5) myoblasts were cultivated either in Skeletal Muscle Cell Growth Medium (Promocell) supplemented with 15% FBS, or in DMEM mixed to 199 medium (1:4 ratio; Life Technologies) and supplemented with 20% FBS, 25 μg/ml fetuin, 0.5 ng/ml bFGF, 5 ng/ml EGF, 5 μg/ml insulin and 0.2 μg/ml dexamethasone (Sigma-Aldrich). Differentiation of myoblasts into myotubes was induced in confluent cells by replacing the growth medium with differentiation medium (DMEM

supplemented with 10 μg/ml insulin) for 5-6 days. Standard temperature of 37°C and 5% CO_2 were used to grow and maintain cells in culture.

Cells were seeded the day before transfection in 6 or 12 well plates and transfected at 70-80% of confluence. Transfection reagent FuGENE HD (FuGENE-DNA ratio 3:1; Promega) was used to transfect HeLa cells. Cells were harvested by centrifugation 2-3 days post transfection and cellular pellet was kept at – 80°C until genomic DNA extraction.

Lentiviral Vectors and Transduction Experiments

Lentiviral vectors were constructed by cloning inserts U6::n(up)-sgRNA;U6::n(dw)-sgRNA_DMPK into the *Xho*I/*Eco*RV site of a pCCL plasmid [pCC-hPGK.GFP (MLS87); gift from Dr. Mario Amendola] to obtain pCCL-U6::n(up)-sgRNA;U6::n(dw)-sgRNA_DMPK-hPGK.GFP (see construct MLS100 as example). The CMV promoter, derived from plasmid MLS42, was cloned into the *Xho*I/*Age*I site of promoterless pCCL-GFP (MLS87 without hPGK promoter) to obtain pCCL-CMV-GFP (MLS107). The construction of the lentiviral vector pCCL-CMV-SaCas9 (MLS110) was done by cloning SaCas9 PCR insert [primers F-AgeI-SaCas9 (MLS142) and R-SalI-SaCas9 (MLS143); plasmid MLS42 as template] into *Sal*I/*Age*I site of pCCL-CMV (MLS107 without GFP).

Lentiviral vectors were produced by calcium phosphate transient transfection of 293T cells as previously described [60]. Vector titers [vector genome per ml (vg/ml)] were determined by quantitative PCR (qPCR) on genomic DNA of infected HCT116 cells (0.32 $x10^9$ vg/ml for Cas9 and 1-1.7 $x10^9$ vg/ml for sgRNAs; virus production and titration by Genethon Vector Core and Quality Control Services, respectively).

For transduction studies, DM1 myoblasts were seeded the day before in 12 well plates and infected at 70% of confluence. Growth medium was removed before transduction and replaced with a minimal volume (400 µl/dish) of transduction medium [skeletal muscle basal medium (Promocell) or DMEM, supplemented with 10% FBS and 4 µg/ml polybrene]. Vectors were added directly to the transduction medium and cells were incubated for 5-6 hours before adding full growth medium. At day 1 post-transduction, cells were transferred to 6 well plates and kept in culture for two total passages before to 1) collect and freeze them for gDNA extraction, 2) fix them for fluorescence in situ hybridization and immunofluorescence analyses.

rAAV Vectors and Animal Experimentation

Recombinant adeno-associated virus (rAAV) vectors for SaCas9 and sgRNA couple 4-23 have been constructed by using pAAV plasmids [Genethon plasmid bank]. The small synthetic promoter SPc5-12, driving high expression of the transgene in muscle [61], was selected for the regulatory cassette of SaCas9, and human U6 promoter for both sgRNAs cassettes. SaCas9 was PCR amplified with primers F-PmeI-SaCas9 (MLS146) and R-NotI-SaCas9_3xHE (MLS147) and using plasmid MLS42 as template. Gel-purified insert SaCas9 was cloned into *Pme*I/*Not*I site of AAV plasmid pC512-Int-smSVpolyA (MLS1) in order to obtain pAAV-SPc5-12-SaCas9 (MLS118). pAAV-Des-eGFP-KASH-U6::4-23-sgRNA_DMPK (MLS123) was obtained by cloning PCR insert U6::4-23-sgRNA_DMPK [primers F-MCS-before-U6SasgRNA (MLS163) and R-PmlI-EndSasgRNA-up (MLS166); plasmid MLS93 as template] into *Afl*II/*Mss*I site of pAAV-Des-eGFP-KASH (MLS23/MLS27). For production of rAAV9 vectors, the *cis*-acting plasmids expressing either SaCas9 or sgRNA, a trans-complementing rep-cap9 plasmid and an

adenovirus helper plasmid were co-transfected into HEK293 cells. Vector particles were purified and titrated as previously described [62].

The transgenic DMSXL mouse line carrying a 45 kb expanded human DMPK genomic fragment was used in this study [47]. Care and manipulation of mice were performed in accordance with national and European legislations on animal experimentation and approved by the institutional ethical committee. The genotype of the animals was assessed by PCR as previously described [63]. The intramuscular injection of rAAV9 vectors was performed in ketamine/xylazine anesthetized WT, heterozygous and homozygous DMSXL mice at 3 to 9 weeks of age depending on the experiments. Equal amounts of the two vectors (1:1) were injected into the left tibialis anterior muscle (0.6×10^{11} total vg/TA at 3 weeks and 1×10^{11} total vg/TA in older mice; PBS was injected into the right TA as control. Four weeks post-injection, TA muscles were collected and frozen in liquid nitrogen for DNA and RNA extractions, or fixed in 4% PFA for FISH/immunofluorescence analyses.

Genomic DNA Extraction and PCR

Genomic DNA was extracted from HeLa cells and immortalized myoblasts either with GeneJet Genomic DNA purification Kit (Thermo Fisher Scientific) or with QIAmp DNA Micro and Mini Kit (QIAGEN), according to manufacturer's instruction. PCR was performed with Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen) in presence of 150 ng of gDNA as template; in particular to amplify *DMPK* 3' UTR, the PCR master mix was supplemented with 10% DMSO; primers MLS14 and MLS15 were used for the experiments shown in Figure 1B, 2B and S1C, and primers MLS14 and MLS17 for the other experiments (see primer list in Table S4). The PCR in Figure S3B was performed in presence of 2x DNA polymerase buffer, 38 cycles and

5min extension time in order to amplify the CTG repeat expansion present in DMSXL muscles. PCR products were separated by electrophoresis in a 1.5-2% agarose gel containing GelRed DNA stain. PCR products obtained by gel extraction or purification (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel) were sequenced by Sanger DNA sequencing (Beckman Coulter Genomics and Genewiz).

FISH and Immunofluorescence

Cultured cell analyses. FISH experiments were performed with the probe Cy3-labeled 2'OMe (CAG)7 (Sigma-Aldrich; 1:1000 dilution, 100 μM stock) as described by Taneja KL [64]. Briefly, cells cultivated in chamber slides (Corning) were washed in phosphate-buffered saline (PBS) and fixed in 4% PFA. After fixation, cells were washed in PBS and stored in 70% ethanol at 4°C for at least 30 min. Cells were hydrated in PBS and incubated with the Cy3-(CAG)7 in hybridization buffer (40% formamide, 2x saline-sodium-citrate (SSC), 0.2% BSA). Cells were then washed before adding mounting solution containing DAPI (SouthernBiotech) and kept at 4°C. When FISH was coupled to immunofluorescence (IF), after hybridization, microscopy slides were washed several times before permeabilization in PBS/0.25% TritonX-100. SaCas9 was detected by antibodies directed against the HA tag epitope located at the C-terminus of the protein. Purified mouse monoclonal anti-HA tag (Covance) was used as primary antibody at dilution 1/400 in 5% BSA and incubated for 1 hour and 30 min at RT. Goat anti-mouse 633 secondary antibody (Themo Fisher Scientific) was used at dilution 1/1000 in 5% BSA and incubated for 1 hour at RT.

Muscle analyses. For muscle analysis, TA muscles were collected and immediately fixed in 4% PFA, incubated in 15% sucrose solution and frozen in ice-cold isopentane. Frozen muscles were

cut in section of 8 μm and subject to FISH coupled to IF or to simply IF. FISH experiments were done with the same probe used for experiments in DM1 cells (Cy3-labeled 2'OMe; Sigma-Aldrich) and protocol previously described [65]. Briefly, slides were pretreated in a boiled Target Retrieval Solution (DAKO), and then washed several times, before a 5 min incubation in PBS/2% ice-cold acetone. After permeabilization, microscopy slides were incubated first in 2x SSC buffer/30% formamide for 10 min at RT, and then in hybridization buffer (2x SSC, 30% formamide, 0.02% BSA, 2 mM vanadyl ribonucleoside complex, 66 μg/ml yeast tRNA, Cy3-(CAG)7 probe 1/150 dilution from 100 μM stock) for 2 hours at 37°C. After wash, slides were subject to IF and incubated with primary antibodies overnight at 4°C. In particular, rabbit polyclonal antibodies anti-laminin (Dako, Z0097) were used at 1/1000 dilution, rabbit polyclonal antibodies anti-GFP (Abcam, Ab6556) at 1/2000 and monoclonal antibodies anti-HA tag (Covance) at 1/100. Muscle sections were incubated with secondary antibodies at 1/1000 dilution: Alexa Fluor 647 donkey anti-rabbit IgG (H+L) (Jackson ImmunoResearch) was used for laminin detection, Alexa Fluor 488 goat anti-rabbit IgG (H+L) (Invitrogen) for GFP, and Alexa Fluor 594 F(ab')2 goat anti-mouse IgG (H+L) (Invitrogen) for HA tag (Cas9). Slides were mounted with an antibleaching solution containing DAPI (Prolong Diamond Antifade Mountant with Dapi, Molecular probes) for microscopy image acquisition.

Confocal microscopy. FISH-IF images were captured with a spectral confocal LEICA SP8 scanning microscope (Leica Microsystems, Germany). We employed the following laser excitation wavelengths: 405 nm for DAPI, 552 nm for Cy3-FISH, 635 nm for laminin, 488 nm for GFP and 552 nm for HA-Cas9. Z-stack images were obtained from the stacking of 23 serial images with 0.45 μm interval. For each TA muscle section, a total of 16 Z-stack images were acquired and analysed by Leica Application Suite X software. Counting was performed manually

and nuclei outside the muscle fibers were excluded from the analysis. Images were processed either with Adobe Photoshop and/or with ImageJ software.

Southern Blot Analysis

Genomic DNA was extracted from control (Ctrl) and DM1 immortalized cell lines, and derivative clones as described above. Approximately 5 μg of gDNA was digested with *Eco*RI restriction enzyme overnight at 37°C. Digested DNA was resolved on a 0.7% agarose gel for ~16 hours at 50V. After migration, agarose gel was incubated for 1 hour in 1M NaOH solution, to denaturate the DNA, and then for 2 hours in neutralization buffer (1M Tris, 3M NaCl pH 8.5). Genomic DNA fragments were transferred from gel to Genescreen Plus Hybridization membrane (Perkin Elmer) via capillary action in 6x SSC buffer, and cross-linked to membrane using the Stratalinker UV crosslinker. DNA was hybridized with 2 x 10^6 cpm/ml of 1.4 kb *Bam*HI probe (B1.4) covering the region of *DMPK* CTG repeat [39]. Probe was pre-labeled with High Prime DNA labeling Kit (Sigma) and hybridization was performed at 68°C overnight in PerfectHyb Plus Hybridization buffer (Sigma) containing 50 μg/ml Human Cot-1 DNA (Thermo Fisher Scientific). Signal was revealed by using Phosphorimager.

RT-PCR for Alternative Splicing

Cells and tissues were homogenized in TRIzol reagent (Life Technologies) using a FastPrep apparatus, and total RNA extraction was done according to manufacturer's instructions. RT-PCR was performed as previously described [45]. Briefly, 1 μg of total RNA was reverse transcribed by M-MLV reverse transcriptase (Life Technologies) and 1 μl of cDNA preparation was used for the PCR (KAPA2G Fast ReadyMix, Sigma) with primers listed in Table S4. PCR products were

separated by 2% agarose gel electrophoresis and visualized with GelRed DNA stain upon UV exposition. Optical density of each PCR band was quantified using ImageJ software and percentage of exon inclusion was calculated as [exon inclusion band/ (summa exon inclusion + exclusion bands)] x 100.

qRT-PCR for *DMPK* mRNA analysis

Total RNA was extracted from cells lysates and homogenized muscle tissues in TRIzol reagent (Life Technologies) according to manufacturer's instructions. RNA was subject to DNase treatment to remove genomic DNA contamination (Ambio DNA-free^{TM} DNA Removal Kit, Life Technologies) and reverse-transcribed using random hexamers and RevertAid H minus Reverse Transcriptase (Fermentas) (100 and 500 ng of total RNA from cells and muscle tissues, respectively). Quantitative PCR was performed in a LyghtCycler 480 system (Roche) by using 4 µl of 1/10 diluted cDNA, SybrGreen mix (ThermoScientific) and primers listed in Table S4.

Deep sequencing

Investigation of indels types and distributions was conducted by Illumina deep sequencing of PCR amplicons generated from DM1 cells treated with MOI 100 of lentiviral vectors expressing SaCas9 and sgRNA4-23, and from DMSXL TA muscles intramuscularly injected at 5-6 weeks of age with rAAV9 vectors expressing SaCas9 and sgRNA4-23. Untreated DM1 cells and TA muscle injected with PBS were used as controls. PCR amplicons of around 300 bp were generated by nested PCR: first PCR reaction was performed with a set of primers specific for the targets (F1-DMPK-3UTR and R2-DMPK-3UTR for the region containing the CTG repeat deletion; F1-DMPK-3UTR and R-DMPK bef CTG for the region surrounding the target of sgRNA 4; F-DMPK-149up-sgRNA23 and R2-DMPK-3UTR for the region surrounding the

target of sgRNA 23). Amplicons generated from the first PCR served as template for a second reaction, which was performed with a second set of primers annealing downstream the first set and containing Illumina adaptors at the 5' end of the sequence (see Table S4). Row deep sequencing reads from Illumina PE sequencing (IGATech) were merged using PEAR (v0.9.6, using : -n 50 –v 20 –q 20 –t 20) [66], and NEXTERA adaptor trimmed using cutadapt (v1.18, using –e 0.2, -o 10 –q 20 –m 50) [67]. Merged sequences were then submitted to CRISPRESSO2 for INDELS discovery (using --ignore-substitutions --q 20) [68].

Statistical Analysis

All data are represented as mean ± SD. Statistical analyses were performed with two-tailed Student-t test, paired for untreated and treated TA muscles from mice of the same genotype, or unpaired and equal variance for TA muscles from mice of different genotype. Analysis of the RT-PCR data on myoblasts cell lines was also done with two-tailed Student-t test. Differences were considered to be statistically different at P *< 0.05; **<0.01; ***<0.001.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and four tables.

AUTHOR CONTRIBUTIONS

A.B.B. and M.L.S conceptualized the study. M.L.S, K.P., C.S., S.T., A.K., G.C. and A.H. performed and/or analyzed the experiments. A.B.B., D.F. and G.G. supervised the project. M.L.S. and A.B.B. wrote the manuscript, which was revised and approved by all authors.

CONFLICTS OF INTEREST

A.B.B. and M.L.S. are inventors of two patent applications related to this work. The remaining authors K.P., C.S., S.T., A.K., G.C., A.H., D.F. and G.G. declare no competing financial interests.

ACKNOWLEDGEMENTS

We are very grateful to the platforms of Genethon for their excellent technical expertise and contribution to this work, in particular the “Bioexperimentation and Functional Evaluation” team, the “Histology Core Facility” and Dr. Edith Renaud-Gabardos for help in mouse experiments, the “Vector Core Facility” for virus production, the “Imaging and Cytometry Core Facility” for technical support, and Genopole Research (Evry, France) for the purchase of the imaging equipment. We also thank Dr. Jean-Paul Concordet (Museum National d’Histoire Naturelle, INSERM U1154, Paris) for providing plasmids Nm Cas9 and sgRNA, Dr. Mario Amendola (Genethon, UMR_S951 Inserm) for providing lentiviral plasmid backbone, and the

platform for immortalization of human cells from the Institut de Myologie (Paris, France). This work was supported by the "*Association Française contre les Myopathies*" (AFM-Telethon) and M.L.S. was partially supported by the grant ANR-16-CE18-0012.

REFERENCES

1. Udd, B. and Krahe, R. (2012). The myotonic dystrophies: molecular, clinical, and therapeutic challenges. Lancet Neurol. *11*, 891-905.
2. Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., *et al.* (1992). Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. Nature *355*, 547-548.
3. Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., *et al.* (1992). Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. Cell *68*, 799-808.
4. Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., *et al.* (1992). Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. Nature *355*, 545-546.
5. Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., *et al.* (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. Science *255*, 1253-1255.
6. Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsilfidis, C., *et al.* (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. Nature *355*, 548-551.
7. Kaliman, P. and Llagostera, E. (2008). Myotonic dystrophy protein kinase (DMPK) and its role in the pathogenesis of myotonic dystrophy 1. Cell Signal. *20*, 1935-1941.
8. Fu, Y.H., Pizzuti, A., Fenwick, R.G., Jr., King, J., Rajnarayan, S., Dunne, P.W., *et al.* (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. Science *255*, 1256-1258.
9. De Temmerman, N., Sermon, K., Seneca, S., De Rycke, M., Hilven, P., Lissens, W., *et al.* (2004). Intergenerational instability of the expanded CTG repeat in the DMPK gene: studies in human gametes and preimplantation embryos. Am. J. Hum. Genet. *75*, 325-329.
10. Morales, F., Couto, J.M., Higham, C.F., Hogg, G., Cuenca, P., Braida, C., *et al.* (2012). Somatic instability of the expanded CTG triplet repeat in myotonic dystrophy type 1 is a heritable quantitative trait and modifier of disease severity. Hum. Mol. Genet. *21*, 3558-3567.
11. Gomes-Pereira, M., Cooper, T.A. and Gourdon, G. (2011). Myotonic dystrophy mouse models: towards rational therapy development. Trends Mol. Med. *17*, 506-517.
12. Taneja, K.L., McCurrach, M., Schalling, M., Housman, D. and Singer, R.H. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. J. Cell Biol. *128*, 995-1002.
13. Fardaei, M., Larkin, K., Brook, J.D. and Hamshere, M.G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. Nucleic Acids Res. *29*, 2766-2771.
14. 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.
15. 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.
16. Kanadia, R.N., Johnstone, K.A., Mankodi, A., Lungu, C., Thornton, C.A., Esson, D., *et al.* (2003). A muscleblind knockout model for myotonic dystrophy. Science *302*, 1978-1980.
17. 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.
18. Lin, X., Miller, J.W., Mankodi, A., Kanadia, R.N., Yuan, Y., Moxley, R.T., *et al.* (2006). Failure of MBNL1-dependent post-natal splicing transitions in myotonic dystrophy. Hum. Mol. Genet. *15*, 2087-2097.
19. Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., *et al.* (2013). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature *498*, 241-245.

20. Nakamori, M., Sobczak, K., Puwanant, A., Welle, S., Eichinger, K., Pandya, S., *et al.* (2013). Splicing biomarkers of disease severity in myotonic dystrophy. Ann. Neurol. *74*, 862-872.
21. Klein, A.F., Dastidar, S., Furling, D. and Chuah, M.K. (2015). Therapeutic Approaches for Dominant Muscle Diseases: Highlight on Myotonic Dystrophy. Curr. Gene Ther. *15*, 329-337.
22. Furling, D., Doucet, G., Langlois, M.A., Timchenko, L., Belanger, E., Cossette, L., *et al.* (2003). Viral vector producing antisense RNA restores myotonic dystrophy myoblast functions. Gene Ther. *10*, 795-802.
23. Francois, V., Klein, A.F., Beley, C., Jollet, A., Lemercier, C., Garcia, L., *et al.* (2011). Selective silencing of mutated mRNAs in DM1 by using modified hU7-snRNAs. Nat. Struct. Mol. Biol. *18*, 85-87.
24. Wheeler, T.M., Leger, A.J., Pandey, S.K., MacLeod, A.R., Nakamori, M., Cheng, S.H., *et al.* (2012). Targeting nuclear RNA for in vivo correction of myotonic dystrophy. Nature *488*, 111-115.
25. Zhang, W., Wang, Y., Dong, S., Choudhury, R., Jin, Y. and Wang, Z. (2014). Treatment of type 1 myotonic dystrophy by engineering site-specific RNA endonucleases that target (CUG)(n) repeats. Mol. Ther. *22*, 312-320.
26. Batra, R., Nelles, D.A., Pirie, E., Blue, S.M., Marina, R.J., Wang, H., *et al.* (2017). Elimination of Toxic Microsatellite Repeat Expansion RNA by RNA-Targeting Cas9. Cell *170*, 899-912 e810.
27. Kanadia, R.N., Shin, J., Yuan, Y., Beattie, S.G., Wheeler, T.M., Thornton, C.A., *et al.* (2006). Reversal of RNA missplicing and myotonia after muscleblind overexpression in a mouse poly(CUG) model for myotonic dystrophy. Proc. Natl. Acad. Sci. U S A *103*, 11748-11753.
28. Wheeler, T.M., Sobczak, K., Lueck, J.D., Osborne, R.J., Lin, X., Dirksen, R.T., *et al.* (2009). Reversal of RNA dominance by displacement of protein sequestered on triplet repeat RNA. Science *325*, 336-339.
29. Warf, M.B., Nakamori, M., Matthys, C.M., Thornton, C.A. and Berglund, J.A. (2009). Pentamidine reverses the splicing defects associated with myotonic dystrophy. Proc. Natl. Acad. Sci. U S A *106*, 18551-18556.
30. Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E. and Artero, R.D. (2011). In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. Proc. Natl. Acad. Sci. U S A *108*, 11866-11871.
31. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.
32. Cornu, T.I., Mussolino, C. and Cathomen, T. (2017). Refining strategies to translate genome editing to the clinic. Nat. Med. *23*, 415-423.
33. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., *et al.* (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.
34. Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., *et al.* (2016). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400-403.
35. Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., *et al.* (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403-407.
36. Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., *et al.* (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407-411.
37. Nelson, C.E., Robinson-Hamm, J.N. and Gersbach, C.A. (2017). Genome engineering: a new approach to gene therapy for neuromuscular disorders. Nat. Rev. Neurol. *13*, 647-661.
38. Xie, H., Tang, L., He, X., Liu, X., Zhou, C., Liu, J., *et al.* (2018). SaCas9 Requires 5'-NNGRRT-3' PAM for Sufficient Cleavage and Possesses Higher Cleavage Activity than SpCas9 or FnCpf1 in Human Cells. Biotechnol. J. *13*, e1700561.

39. Gourdon, G., Radvanyi, F., Lia, A.S., Duros, C., Blanche, M., Abitbol, M., *et al.* (1997). Moderate intergenerational and somatic instability of a 55-CTG repeat in transgenic mice. Nat. Genet. *15*, 190-192.
40. Gomes-Pereira, M., Foiry, L., Nicole, A., Huguet, A., Junien, C., Munnich, A., *et al.* (2007). CTG trinucleotide repeat "big jumps": large expansions, small mice. PLoS Genet. *preprint*, e52.
41. Zhang, Y., Heidrich, N., Ampattu, B.J., Gunderson, C.W., Seifert, H.S., Schoen, C., *et al.* (2013). Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis. Mol. Cell *50*, 488-503.
42. Lo Scrudato, M., Martin, S., Gourdon, G., Furling, D. and Buj-Bello, A. (2016). Genome editing for nucleotide repeat disorders: towards a new therapeutic approach for Myotonic Dystrophy type 1. Mol. Ther. *24*, S129-S130.
43. Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., *et al.* (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186-191.
44. Brinkman, E.K., Chen, T., Amendola, M. and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res. *42*, e168.
45. Arandel, L., Polay Espinoza, M., Matloka, M., Bazinet, A., De Dea Diniz, D., Naouar, N., *et al.* (2017). Immortalized human myotonic dystrophy muscle cell lines to assess therapeutic compounds. Dis. Model. Mech. *10*, 487-497.
46. Imbert, G., Kretz, C., Johnson, K. and Mandel, J.L. (1993). Origin of the expansion mutation in myotonic dystrophy. Nat. Genet. *4*, 72-76.
47. Huguet, A., Medja, F., Nicole, A., Vignaud, A., Guiraud-Dogan, C., Ferry, A., *et al.* (2012). Molecular, physiological, and motor performance defects in DMSXL mice carrying >1,000 CTG repeats from the human DM1 locus. PLoS Genet. *8*, e1003043.
48. Buj-Bello, A., Fougerousse, F., Schwab, Y., Messaddeq, N., Spehner, D., Pierson, C.R., *et al.* (2008). AAV-mediated intramuscular delivery of myotubularin corrects the myotubular myopathy phenotype in targeted murine muscle and suggests a function in plasma membrane homeostasis. Hum. Mol. Genet. *17*, 2132-2143.
49. Jensen, K.T., Floe, L., Petersen, T.S., Huang, J., Xu, F., Bolund, L., *et al.* (2017). Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. FEBS Lett. *591*, 1892-1901.
50. van Agtmaal, E.L., Andre, L.M., Willemse, M., Cumming, S.A., van Kessel, I.D.G., van den Broek, W., *et al.* (2017). CRISPR/Cas9-Induced (CTGCAG)n Repeat Instability in the Myotonic Dystrophy Type 1 Locus: Implications for Therapeutic Genome Editing. Mol. Ther. *25*, 24-43.
51. Provenzano, C., Cappella, M., Valaperta, R., Cardani, R., Meola, G., Martelli, F., *et al.* (2017). CRISPR/Cas9-Mediated Deletion of CTG Expansions Recovers Normal Phenotype in Myogenic Cells Derived from Myotonic Dystrophy 1 Patients. Mol. Ther. Nucleic Acids *9*, 337-348.
52. Dastidar, S., Ardui, S., Singh, K., Majumdar, D., Nair, N., Fu, Y., *et al.* (2018). Efficient CRISPR/Cas9-mediated editing of trinucleotide repeat expansion in myotonic dystrophy patient-derived iPS and myogenic cells. Nucleic Acids Res. *46*, 8275–8298.
53. Gao, Y., Guo, X., Santostefano, K., Wang, Y., Reid, T., Zeng, D., *et al.* (2016). Genome Therapy of Myotonic Dystrophy Type 1 iPS Cells for Development of Autologous Stem Cell Therapy. Mol. Ther. *24*, 1378-1387.
54. Richard, G.F. (2015). Shortening trinucleotide repeats using highly specific endonucleases: a possible approach to gene therapy? Trends Genet. *31*, 177-186.
55. Mosbach, V., Poggi, L., Viterbo, D., Charpentier, M. and Richard, G.F. (2018). TALEN-Induced Double-Strand Break Repair of CTG Trinucleotide Repeats. Cell Rep. *22*, 2146-2159.
56. Cinesi, C., Aeschbach, L., Yang, B. and Dion, V. (2016). Contracting CAG/CTG repeats using the CRISPR-Cas9 nickase. Nat. Commun. *7*, 13272.
57. Pinto, B.S., Saxena, T., Oliveira, R., Mendez-Gomez, H.R., Cleary, J.D., Denes, L.T., *et al.* (2017). Impeding Transcription of Expanded Microsatellite Repeats by Deactivated Cas9. Mol. Cell *68*, 479-490 e475.

58. Bengtsson, N.E., Hall, J.K., Odom, G.L., Phelps, M.P., Andrus, C.R., Hawkins, R.D., *et al.* (2017). Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. Nat. Commun. *8*, 14454.
59. Amoasii, L., Hildyard, J.C.W., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., *et al.* (2018). Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. Science.
60. Cantore, A., Ranzani, M., Bartholomae, C.C., Volpin, M., Valle, P.D., Sanvito, F., *et al.* (2015). Liver-directed lentiviral gene therapy in a dog model of hemophilia B. Sci. Transl. Med. *7*, 277ra228.
61. Li, X., Eastman, E.M., Schwartz, R.J. and Draghia-Akli, R. (1999). Synthetic muscle promoters: activities exceeding naturally occurring regulatory sequences. Nat. Biotechnol. *17*, 241-245.
62. Daniele, N., Moal, C., Julien, L., Marinello, M., Jamet, T., Martin, S., *et al.* (2018). Intravenous Administration of a MTMR2-Encoding AAV Vector Ameliorates the Phenotype of Myotubular Myopathy in Mice. J. Neuropathol. Exp. Neurol. *77*, 282-295.
63. Seznec, H., Lia-Baldini, A.S., Duros, C., Fouquet, C., Lacroix, C., Hofmann-Radvanyi, H., *et al.* (2000). Transgenic mice carrying large human genomic sequences with expanded CTG repeat mimic closely the DM CTG repeat intergenerational and somatic instability. Hum. Mol. Genet. *9*, 1185-1194.
64. Taneja (1998). Localization of trinucleotide repeat sequences in Myotonic Distrophy cells using a single fluorochrome-labeled PNA probe. BioTechniques *24*, 472-476.
65. Seznec, H., Agbulut, O., Sergeant, N., Savouret, C., Ghestem, A., Tabti, N., *et al.* (2001). Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. Hum. Mol. Genet. *10*, 2717-2726.
66. Zhang, J., Kobert, K., Flouri, T. and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614-620.
67. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10.
68. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., *et al.* (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224-226.

FIGURE LEGENDS

Figure 1. CRISPR-SaCas9 Genome Editing to Target *DMPK* CTG Repeats. (A) Scheme of selected Sa sgRNA target sites located in the 3' UTR of the *DMPK* gene between the stop codon (stop) and the polyadenylation signal (pA), flanking the CTG repeats [(CTG)n]; sgRNAs in black resulted in higher percentage of indels as assessed by TIDE (Table S1). (B) Genomic PCR analysis showing the deletion of the region flanking the CTG repeats in HeLa cells. Cells were transfected with plasmids expressing SaCas9 and the indicated sgRNA couples (combinations of sgRNAs downstream the CTG repeat region 12A, 12B, 13A, or 23, and sgRNAs upstream 1, 4, 7, or 8). (NT): non-transfected cells; (-): SaCas9 expressing plasmid without sgRNA. White and black arrows indicate undeleted and deleted PCR amplicons, respectively.

Figure 2. CRISPR-SaCas9 Lentiviral Vectors Delete CTG Repeats in DM1 patient-derived muscle line cells. (A) Scheme of lentiviral vector constructs containing SaCas9 and sgRNA sequences. LTR: long terminal repeat; CMV: cytomegalovirus promoter; NLS: nuclear localization signal; HA: human influenza hemagglutinin epitope; pA: polyadenylation signal; U6: human U6 small nuclear RNA (snRNA) gene promoter; sgRNAup: sgRNA sequence targeting regions upstream the CTG repeats; sgRNAdw: sgRNA sequence targeting regions downstream the CTG repeats; hPGK: human phosphoglycerate kinase gene promoter; GFP: enhanced green fluorescent protein. (B) PCR amplicons of the genomic region containing the CTG repeats in DM1 myoblasts transduced with increasing MOI of lentiviral vectors expressing SaCas9 and the indicated sgRNA couples. Triangle in gradient colors reflects the MOI, from 5 in white to 100 in black. A total MOI of 5, 10, 20, 50 and 100 was used for the two vectors at 1:1 ratio. White and black arrows indicate undeleted and deleted PCR amplicons, respectively.

Genomic DNA from non-transduced cells (NT) and cells transduced with only one lentiviral vector (SaCas9 or sgRNA, MOI 50) was used as controls. (C) Percentage of *DMPK* CTG repeats deletion (% DEL) quantified from agarose gel images showed in panel B. (D) Percentage of DM1 myoblasts without nuclear foci visualized by FISH images after treatment with indicated MOI of lentiviral vectors SaCas9 and sgRNA couple 4-23. Histograms show average values from three independent biological replicates ± SD. Statistical analysis by two-tailed Student *t*-test. *: $P < 0.05$; ns: not significant.

Figure 3. Deletion of Expanded CTG Repeats and Foci Disappearance in DM1 Myoblasts Treated with CRISPR-SaCas9. DM1 myoblast clones were isolated from the bulk population after transduction with lentiviral vectors; isolated clones were analyzed for the presence of nuclear foci (A) and presence of *DMPK* CTG repeats (B to F). (A) FISH-IF images of a representative DM1 myoblast clone without foci (DM1-Delta clone 22); DM1 clones non-transduced (DM1) or transduced with MOI 50 of a lentiviral vector expressing SaCas9 (DM1-Cas9) or $sgRNA_{4\text{-}23}$ only (DM1-sgRNA) were used as control. SaCas9 (α-HA) is shown in red, GFP in green, RNA foci in yellow [(CAG)7], and nuclei in blue (DAPI). Scale bar: 10 μm. (B) PCR analysis of *DMPK* 3' UTR in DM1-Delta clones 10, 3, 17 and 22 amplified with primers F1-DMPK-3UTR and R2-DMPK-3UTR, annealing regions surrounding the cutting sites of sgRNA 4 and 23. Ctrl: control myoblasts; DM1: non-transduced DM1 clone; DM1-sgRNA: DM1 clone transduced with only lentiviral vector expressing $sgRNA_{4\text{-}23}$. PCR amplicons with undeleted (a; white arrowhead) and deleted (b; black arrowhead) CTG repeats are shown. (C) Sequencing chromatograms of deleted PCR products of panel B from DM1-Delta clones 10, 3, 17 and 22, showing resection of the CTG repeats and resulting DNA end-joining. (D) Sequence

alignment of deleted and undeleted PCR products of panel B of ~1 kb (a) and ~0.4 kb (b), respectively. Indels and/or deletions of the CTG repeats are indicated by dashes; nucleotide substitutions are in red. Target sequences of sgRNA 4 and 23 are indicated in blue and pink, respectively; PAM sequences are underlined; the CTG repeats region [(CTG)n] is highlighted in black. (E) Schematic representation of the *DMPK* gene and exon 15, indicating the relative positions of *Eco*RI cutting sites, the 1 kb *Alu* polymorphism (*Alu*) and the annealing region of probe B1.4. (F) Southern blot showing the genomic deletion of 2600 CTG repeats in DM1-Delta clones. Genomic DNA was digested with *Eco*RI and hybridized with probe B1.4 showed in panel E. Bands corresponding to the two alleles can be distinguished by size because the *Alu* insertion is 1 kb. (G) Number of CTG repeats [(CTG)n] in each allele of the *DMPK* gene, with (1) and without (2) the *Alu* insertion, and expected size of *Eco*RI bands, in control (Ctrl), DM1, DM1-Delta and DM1-sgRNA clones. ctg: non expanded CTG repeats; CTG_{exp}: expanded CTG repeats; Δ: deletion of the CTG repeats; white and blue arrows: allele with and without the *Alu* insertion.

Figure 4. Reversion of Splicing Abnormalities in DM1 patient-derived muscle cells by CRISPR-SaCas9 Deletion of Expanded CTG Repeats. Splicing profiles and quantification of *LDB3* exon 11, *ATP2A1* exon 22, *MBNL1* exon 7, *DMD* exon 78, *IR* exon 11 and *BIN1* exon 11 containing transcripts in differentiated myoblasts from DM1-Delta clones 10, 3, 17 and 22 compared to control (Ctrl), DM1, and DM1-sgRNA clones. Graphs show average values from independent biological replicates ± SD (N=6 for Ctrl and DM1, N=3 for the other samples). Statistical analysis by two-tailed Student *t*-test. *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$; ns: not significant; < DL: below detection limit.

Figure 5. r**AAV-mediated CRISPR-SaCas9 Delivery in Muscle of DMSXL Mice Results in Deletion of Expanded *DMPK* CTG Repeats.** (A) Schematic representation of rAAV9 constructs: the expression of SaCas9 and sgRNAs 4-23 is under the control of the SPc5-12 and U6 promoters, respectively. The sgRNA construct contains a eGFP-Kash reporter under the Desmin promoter (Desm). ITR: Inverted Terminal Repeat; SPc5-12: synthetic muscle-specific promoter; Int: intron; NLS: nuclear localization signal; HA: human influenza hemagglutinin epitope; pA: polyadenylation signal; eGFP-K: enhanced Green Fluorescent Protein fused to Kash peptide; U6: human U6 small nuclear RNA (snRNA) gene promoter. (B) Representative immunofluorescence images of tibialis anterior muscle cross sections from homozygous (HMZ) DMSXL mice 4 weeks after intramuscular injection of CRISPR-SaCas9 rAAV9 vectors, showing expression of SaCas9 (α-HA, in red) and sgRNAs (eGFP-K, in green) within muscle fibers. DAPI was used for nuclear staining. 21% and 76% of myonuclei were positive for SaCas9 and GFP, respectively, and 18% were positive for both. A higher magnification of myofibers from the upper panels (white boxes I., II. and III. in merged image) is shown in panels below. Scale bar: 50 μm in upper panels, 25 μm in panels I., II. and III. (C) Genomic PCR of *DMPK* 3' UTR from nine homozygous DMSXL mice (HMZ 1 to 9) 4 weeks after injection of PBS (-) in the left TA muscle and rAAV9 vectors expressing SaCas9 and sgRNAs 4-23 (+) in the contralateral TA. The amplified band of ~0.4 kb corresponds to the edited PCR amplicons with deletion of 1200 CTG repeats. (D) Sequence of deleted PCR products showing the end-joining site (black arrowhead) of sgRNA targets 4 (blue) and 23 (pink) after double stranded breaks. (E) Alignment of unedited (DMSXL) and edited (Δ) *DMPK* 3' UTR sequences from genomic DNA showing the sharp cutting position at nucleotide N3 upstream the PAM of sgRNA 4 and 23.This

representative sequence was obtained after Sanger sequencing of PCR products from TA muscles injected with rAAV-SaCas9 and sgRNA 4-23 (number of TA analyzed is equal to 9).

Figure 6. Indels examination in DM1 patient cells and DMSXL mice after treatment with CRISPR-Cas9. Genomic deep-sequencing of the DMPK 3'-UTR region with (DEL, primers F1-R1) or without CTG repeat deletion (sgRNA4, primers F1-R2; sgRNA23, primers F2-R1). Indels analysis was performed by alignment with the sequence resulting from a cut between nucleotides N3 and N4 at targets 4 and 23 for PCR amplicons with CTG repeat deletion, or alternatively with the respective unmodified genomic sequence. (A) Percentage of reads with indels in bulk population of DM1 cells (DM1 bulk) and in TA muscle of DMSXL mice treated with CRISPR-Cas9 (+). Untreated DM1 cells and TA muscle injected with PBS were used as negative controls (-). (B-C) Indels distribution upstream and downstream the expected cutting sites (0) in PCR amplicons with (DEL) and without CTG repeat deletion (sgRNA4 and sgRNA23) generated from gDNA of treated DM1 cells and of representative TA muscles containing the lowest and the highest percentage of reads with indels.

Figure 7. CRISPR-SaCas9 Expression in DM1 Muscle Decreases Nuclear Foci. (A) Representative confocal images of a TA muscle section stained with antibodies against laminin (α-LMN, red), FISH [(CAG)7, yellow] and DAPI (blue) (upper panels) from HMZ DMSXL mice. TA muscle section from WT animal shows the FISH background. A higher magnification of myofibers from the upper panels (white box in merged image) is shown in panels below. Arrows indicate myonuclei with nuclear foci. Scale bar: 50 μm in upper panels, 25 μm in lower panels. (B) Percentage of myonuclei containing foci in TA muscle fibers from DMSXL mice 4

weeks after PBS (64.83±9.25%) or rAAV9-SaCas9 + rAAV9-sgRNA$_{4\text{-}23}$ (49.25±8.42%) intramuscular injection. Data are represented as means ± SD (N=10 HMZ mice). Statistical analysis with two-tailed Student's *t* test. ***: P < 0.001. (C) Total number of myonuclei per fiber in TA of wild-type (WT) and homozygous (HMZ) mice 4 weeks after injection of either PBS or rAAV9-SaCas9 + rAAV9-sgRNA$_{4\text{-}23}$ vectors. Data are represented as means ± SD (N=3 for WT mice; N=10 for HMZ mice). Statistical analysis with two-tailed Student's *t* test. ns: not significant.

Figure 1.

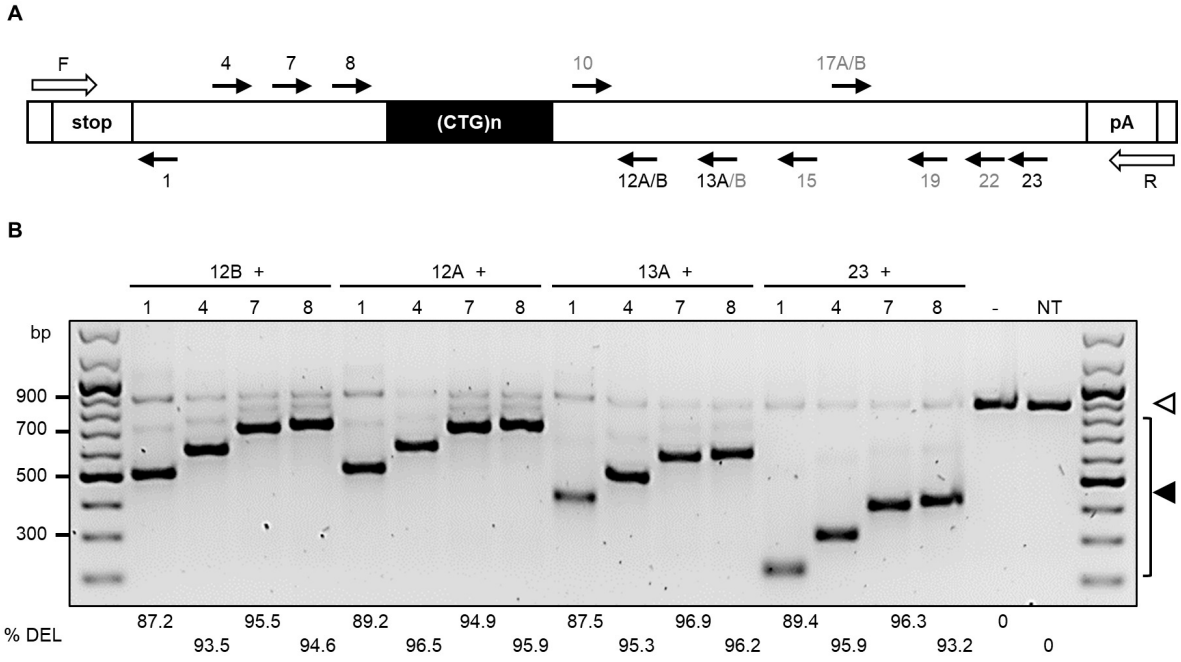

A
F
4
7
8
10
17A/B
stop
(CTG)n
pA
1
12A/B
13A/B
15
19
22
23
R
B
12B +
12A +
13A +
23 +
1 4 7 8 1 4 7 8 1 4 7 8 1 4 7 8 - NT
bp
900
700
500
300
% DEL
87.2 95.5 89.2 94.9 87.5 96.9 89.4 96.3 0
93.5 94.6 96.5 95.9 95.3 96.2 95.9 93.2 0

Figure 2.

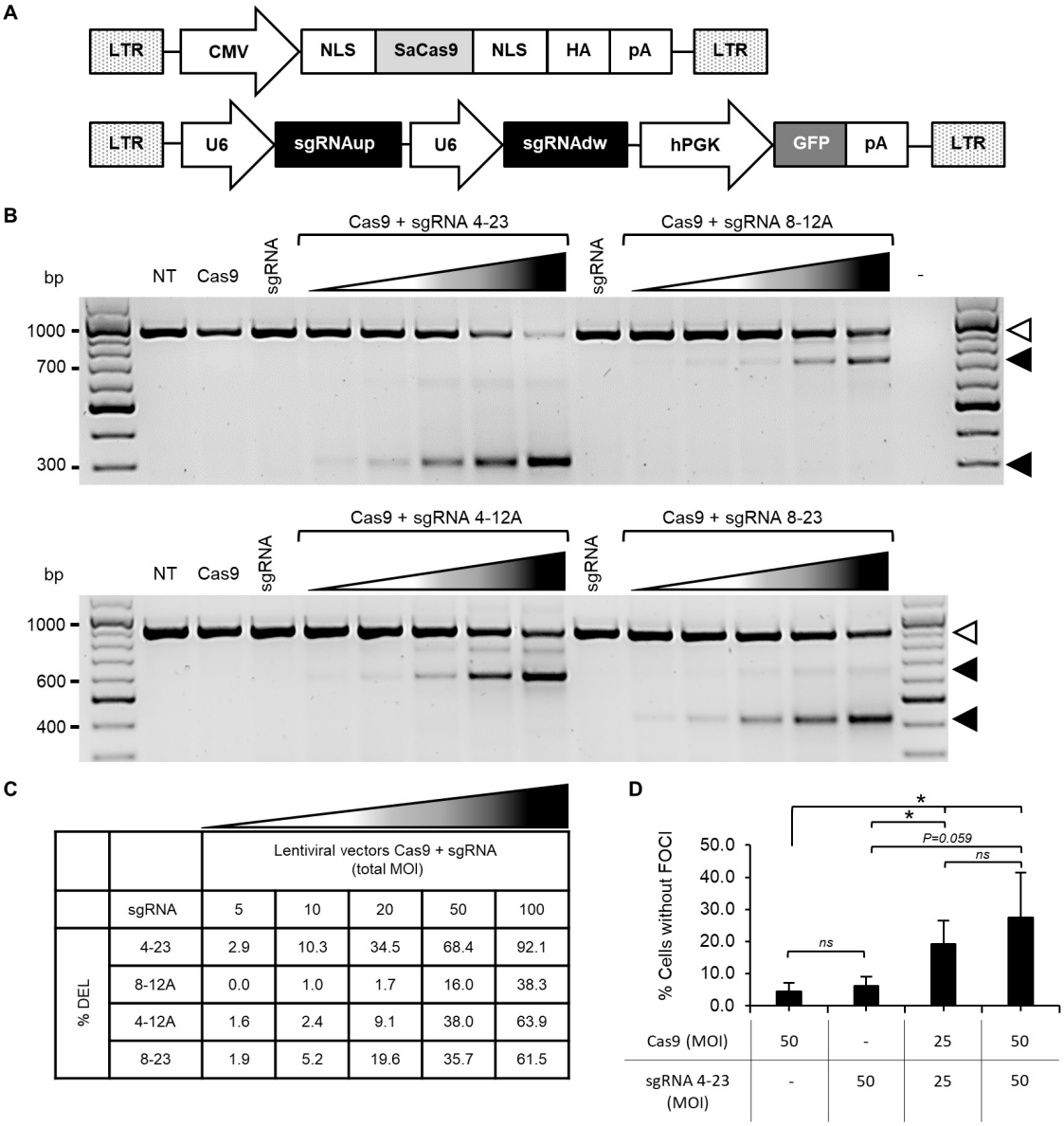
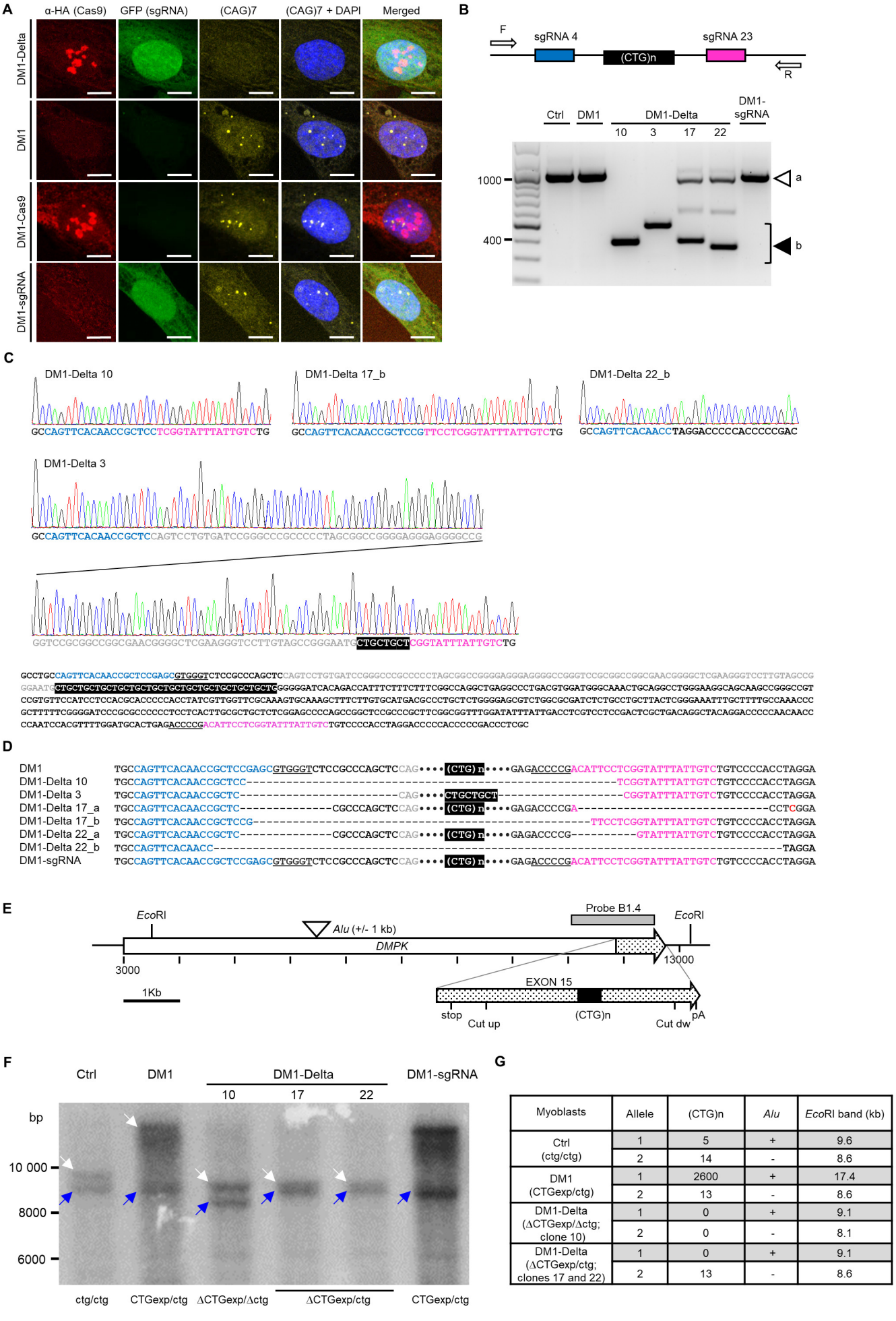

A
LTR
CMV
NLS
SaCas9
NLS
HA
pA
LTR
LTR
U6
sgRNAup
U6
sgRNAdw
hPGK
GFP
pA
LTR
B
bp
NT
Cas9
sgRNA
Cas9 + sgRNA 4-23
sgRNA
Cas9 + sgRNA 8-12A
-
1000
700
300
bp
NT
Cas9
sgRNA
Cas9 + sgRNA 4-12A
sgRNA
Cas9 + sgRNA 8-23
1000
600
400
C
Lentiviral vectors Cas9 + sgRNA (total MOI)
sgRNA
5
10
20
50
100
% DEL
4-23
2.9
10.3
34.5
68.4
92.1
8-12A
0.0
1.0
1.7
16.0
38.3
4-12A
1.6
2.4
9.1
38.0
63.9
8-23
1.9
5.2
19.6
35.7
61.5
D
% Cells without FOCI
50.0
40.0
30.0
20.0
10.0
0.0
*
*
P=0.059
ns
ns
Cas9 (MOI)
50
-
25
50
sgRNA 4-23 (MOI)
-
50
25
50

Figure 3.



D

```
DM1           TGCCAGTTCACAACCGCTCCGAGCGTGGGTCTCCGCCCAGCTCCAG••••(CTG)n••••GAGACCCCGACATTCCTCGGTATTTATTGTCTGTCCCCACCTAGGA
DM1-Delta 10  TGCCAGTTCACAACCGCTCC----------------------------------------------------------TCGGTATTTATTGTCTGTCCCCACCTAGGA
DM1-Delta 3   TGCCAGTTCACAACCGCTC------------------------CAG••••CTGCTGCT----------------CGGTATTTATTGTCTGTCCCCACCTAGGA
DM1-Delta 17_a TGCCAGTTCACAACCGCTC--------------CGCCCAGCTCCAG••••(CTG)n••••GAGACCCCGA----------------------------CCTCGGA
DM1-Delta 17_b TGCCAGTTCACAACCGCTCCG-------------------------------------------------------TTCCTCGGTATTTATTGTCTGTCCCCACCTAGGA
DM1-Delta 22_a TGCCAGTTCACAACCGCTC--------------CGCCCAGCTCCAG••••(CTG)n••••GAGACCCCG----------GTATTTATTGTCTGTCCCCACCTAGGA
DM1-Delta 22_b TGCCAGTTCACAACC-----------------------------------------------------------------------------------TAGGA
DM1-sgRNA     TGCCAGTTCACAACCGCTCCGAGCGTGGGTCTCCGCCCAGCTCCAG••••(CTG)n••••GAGACCCCGACATTCCTCGGTATTTATTGTCTGTCCCCACCTAGGA
```


G

Myoblasts	Allele	(CTG)n	*Alu*	*Eco*RI band (kb)
Ctrl (ctg/ctg)	1	5	+	9.6
	2	14	-	8.6
DM1 (CTGexp/ctg)	1	2600	+	17.4
	2	13	-	8.6
DM1-Delta (ΔCTGexp/Δctg; clone 10)	1	0	+	9.1
	2	0	-	8.1
DM1-Delta (ΔCTGexp/ctg; clones 17 and 22)	1	0	+	9.1
	2	13	-	8.6

Figure 4.

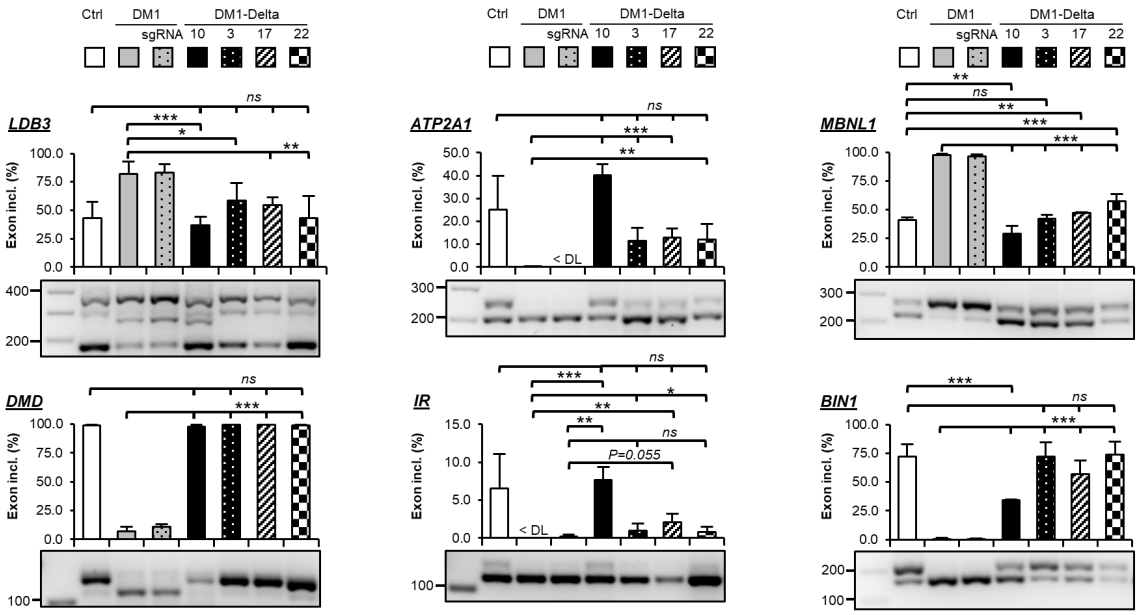

Ctrl
DM1
DM1-Delta
sgRNA 10 3 17 22
LDB3
ATP2A1
MBNL1
DMD
IR
BIN1
Exon incl. (%)
< DL
P=0.055
ns

Figure 5.

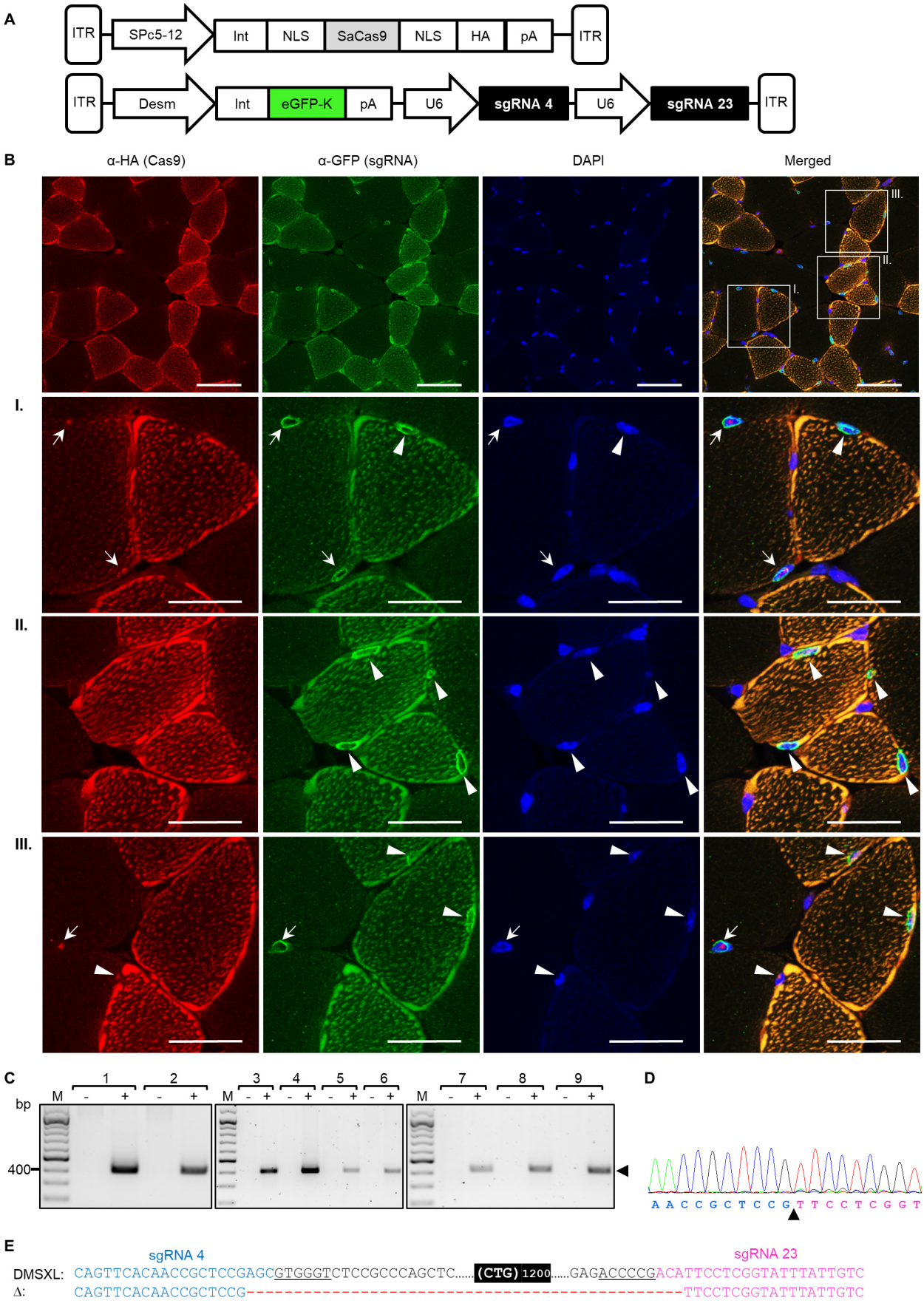
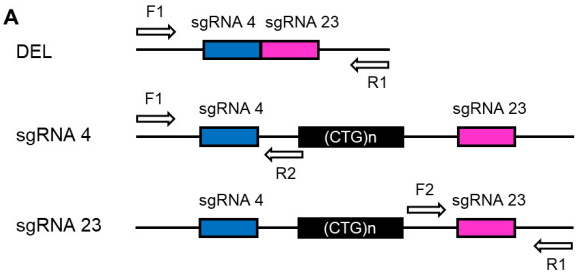
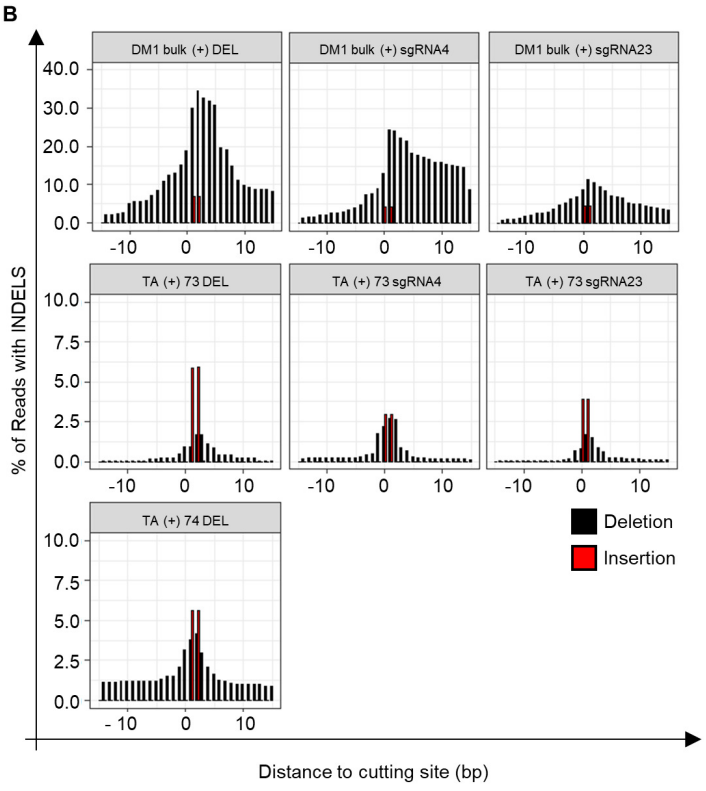

A
ITR
SPc5-12
Int
NLS
SaCas9
NLS
HA
pA
ITR
ITR
Desm
Int
eGFP-K
pA
U6
sgRNA 4
U6
sgRNA 23
ITR
B
α-HA (Cas9)
α-GFP (sgRNA)
DAPI
Merged
I.
II.
III.
C
bp
400
M
1
2
3
4
5
6
7
8
9
D
A A C C G C T C C G T T C C T C G G T
E
sgRNA 4
sgRNA 23
DMSXL: CAGTTCACAACCGCTCCGAGCGTGGGTCTCCGCCCAGCTC.....(CTG)1200.....GAGACCCCGACATTCCTCGGTATTTATTGTC
Δ: CAGTTCACAACCGCTCCG----------TTCCTCGGTATTTATTGTC

Figure 6.



Sequence DMPK 3'-UTR from DM1 patient cells	Treatment	Reads with Indels		Total Reads
PCR amplicon	Lenti CRISPR-Cas9	N	%	N
DEL carrying CTG deletion (primers F1-R1)	+	5525	44.7	12366.0
sgRNA 4 without CTG deletion and covering sgRNA 4 target (primers F1-R2)	_	1	0.0	10901.0
	+	3648	30.8	11844.0
sgRNA 23 without CTG deletion and covering sgRNA 23 target (primers F2-R1)	_	1	0.0	15954.0
	+	1494	17.3	8634.0

Sequence DMPK 3'-UTR from TA muscle of DMSXL mice		Treatment	Reads with indels		Total Reads
PCR amplicon	Mice N	AAV-CRISPR-Cas9	N	%	N
DEL carrying CTG deletion (primers F1-R1)	64	+	808.0	9.1	8906.0
	67		924	8.8	10457.0
	69		993	10.2	9775.0
	71		1066	8.7	12253.0
	72		958	9.1	10519.0
	70		1058	10.7	9920.0
	73		1092	8.5	12810.0
	74		1239	11.8	10505.0
	Average N=8 mice		8138.0	9.6	85145.0
sgRNA 4 without CTG deletion and covering sgRNA 4 target (primers F1-R2)	73	-	6	0.0	15064.0
		+	1320	7.6	17324.0
sgRNA 23 without CTG deletion and covering sgRNA 23 target (primers F2-R1)		-	2	0.0	11725.0
		+	924	6.3	14688.0



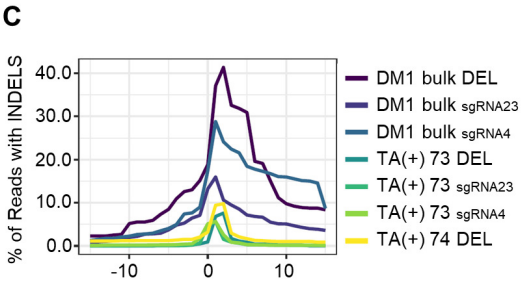
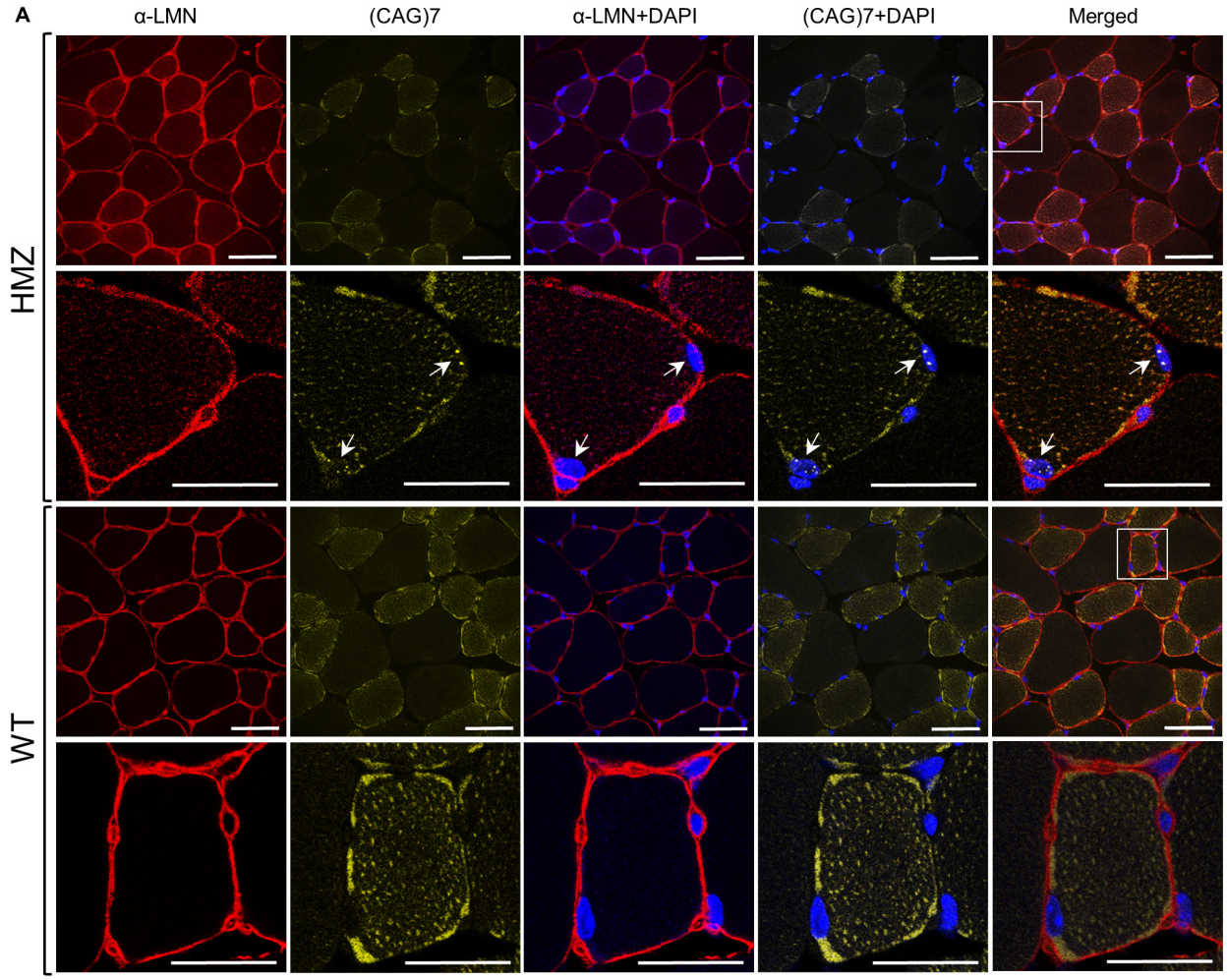
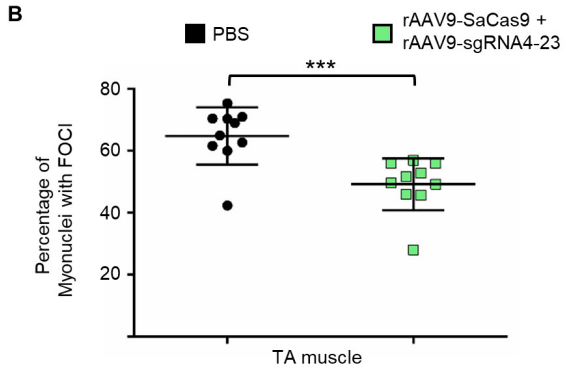
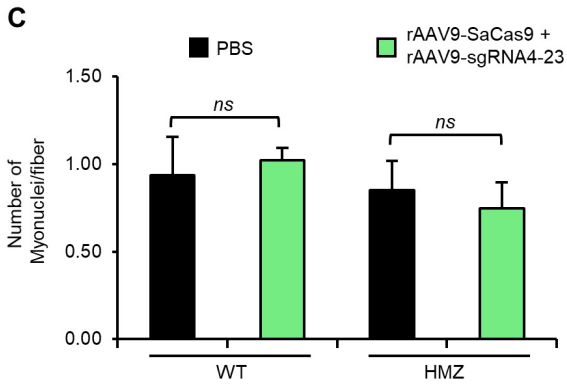


Figure 7.


A
α-LMN
(CAG)7
α-LMN+DAPI
(CAG)7+DAPI
Merged
HMZ
WT


B
PBS
rAAV9-SaCas9 +
rAAV9-sgRNA4-23
***
Percentage of
Myonuclei with FOCI
80
60
40
20
TA muscle


C
PBS
rAAV9-SaCas9 +
rAAV9-sgRNA4-23
ns
ns
Number of
Myonuclei/fiber
1.50
1.00
0.50
0.00
WT
HMZ