

**MUSCLE SELECTIVE HYBRID REGULATORY COMBINATIONS AND
METHODS OF USE THEREOF FOR THE TREATMENT OF MYOTONIC
DYSTROPHY TYPE 1**

RELATED APPLICATIONS

[0001] This application claims priority to, and benefit of, U.S. Provisional Application No. 63/621,332, filed on January 16, 2024 and U.S. Provisional Application No. 63/622,795, filed on January 19, 2024 the contents of each of which are incorporated by reference in their entireties.

FIELD OF THE DISCLOSURE

[0002] The disclosure is directed to molecular biology, gene therapy, and compositions and methods for modifying expression and activity of RNA molecules.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0003] The contents of the electronic sequence listing (ASGT-005_001WO_SeqList_ST26.xml; Size: 197,847 bytes; and Date of Creation: January 14, 2025) are herein incorporated by reference in its entirety.

BACKGROUND

[0004] There are long-felt but unmet needs in the art for providing RNA-targeting systems which provide effective gene therapies. In particular, this disclosure provides compositions and methods for specifically targeting and reducing/eliminating/blocking toxic RNAs expressed from repetitive tracts in the microsatellite repeat expansion (MRE) disease known as myotonic dystrophy type 1 (DM1). DM1 is a multisystemic, autosomal-dominant inherited disorder caused by CTG MREs in the 3' untranslated region of the DMPK gene. As for all MRE diseases, available treatments address symptoms of DM1 but do not target its underlying etiology. Without wishing to be bound by theory, elimination of MREs in DNA with genome editing could eliminate the pathogenic MREs causing DM1. However, generation of DNA breaks near repeats activates the repair machinery whose activity is linked to expansion growth and may cause further mutation of the repeats and/or may fail to distinguish the pathogenic repeats from the normal repeats which possess regulatory roles in transcription. Other potential DM1 therapeutics have been evaluated, such as antisense

oligonucleotides, shRNAs and small molecules. These approaches suffer from issues including, but not limited to, frequent redosing, poor penetration of affected tissues, lack of direct engagement with repeats, toxicity and off-target effects. In an effort to overcome these issues, Cas9-based RNA-targeting systems (RCas9) have been shown to be capable of specifically targeting toxic CUG repeat RNA and providing long-term repair of the disease phenotypes associated with DM1 in adult-onset myotonic dystrophy in mice. Non-Cas9 RNA binding systems (Cas13d, PUF and PUMBY-based) have been developed as gene therapies for the treatment of DM1 and are described in International Patent Application Publication No. WO 2022/119979 which is incorporated herein by reference in its entirety).

[0005] However, improved compositions will allow for more potent yet biased (*i.e.* muscle-selective) expression levels in skeletal muscle tissue over cardiac muscle expression.

Accordingly, novel promoter combinations are disclosed herein for fine-tuning muscle specific gene therapies. In a non-limiting example, the novel promoter combinations disclosed herein can be used in AAV-based PUF compositions comprising such promoter combinations for treating DM1. These AAV-based PUF compositions are also provided herein.

SUMMARY

[0006] The present disclosure provides recombinant adeno-associate viral (rAAV) vectors comprising: a) first inverted terminal repeat (ITR) sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 32; b) an MYLPF enhancer sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 1; c) a CK6 promoter sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 8; d) an SV40 intron sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 35; e) a nucleic acid sequence encoding a PUF polypeptide, wherein the PUF polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 135; f) a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 38; g) an SV40 polyA sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 39; and h) a second ITR sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 34.

[0007] In some aspects, the rAAV vectors of the present disclosure comprise: a) a first inverted terminal repeat (ITR) sequence comprising a nucleic acid sequence at least 99%

identical to SEQ ID NO: 32; b) an MYLPF enhancer sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 1; c) a CK6 promoter sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 8; d) an SV40 intron sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 35; e) a nucleic acid sequence encoding a PUF polypeptide, wherein the PUF polypeptide comprises an amino acid sequence at least 99% identical to SEQ ID NO: 135; f) a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 38; g) an SV40 polyA sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 39; and h) a second ITR sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 34.

[0008] In some aspects of the rAAV vectors of the present disclosure, the nucleic acid sequence encoding the PUF polypeptide sequence comprises a nucleic acid sequence at least 95% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 45. In some aspects of the rAAV vectors of the present disclosure, the nucleic acid sequence encoding the PUF polypeptide sequence comprises a nucleic acid sequence at least 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 45.

[0009] In some aspects, the rAAV vectors comprise a nucleic acid sequence at least 95%, preferably at least 99%, identical to the nucleic acid sequence set forth in SEQ ID NO: 47, SEQ ID NO: 139, or SEQ ID NO: 150.

[0010] In some aspects of the rAAV vectors of the present disclosure comprise an AAV capsid protein. In some aspects, the AAV capsid protein is an AAV1 capsid protein, an AAV2 capsid protein, an AAV4 capsid protein, an AAV5 capsid protein, an AAV6 capsid protein, an AAV7 capsid protein, an AAV8 capsid protein, an AAV9 capsid protein, an AAV10 capsid protein, an AAV11 capsid protein, an AAV12 capsid protein, an AAV13 capsid protein, an AAVPHP.B capsid protein, an AAVrh74 capsid protein, a myoAAV capsid protein or an AAVrh.10 capsid protein. In some aspects, the AAV capsid protein is a myoAAV capsid protein. In some aspects, the AAV capsid protein comprises an amino acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 54, preferably wherein the AAV capsid protein is encoded by a nucleic acid sequence that is at least 95%, preferably at least 99%, identical to SEQ ID NO: 53.

[0011] The present disclosure provides nucleic acid molecules comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises: a) a myosin light chain, phosphorylatable, fast skeletal muscle gene (MYLPF) enhancer sequence

or a desmin enhancer sequence, wherein the MYLPF enhancer sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, and SEQ ID NO: 4, wherein the desmin enhancer sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical any one of SEQ ID NO: 12 and SEQ ID NO: 14; and b) a creatine kinase (CK) promoter sequence or a desmin promoter sequence, wherein the CK promoter sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, or SEQ ID NO: 17, wherein the desmin promoter sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 16 or SEQ ID NO: 22.

[0012] In some aspects, the hybrid regulatory sequences of the present disclosure comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 26, or SEQ ID NO: 27.

[0013] In some aspects, the hybrid regulatory sequences of the present disclosure further comprise an intron sequence. In some aspects, the intron sequence is an SV40 (simian virus 40) intron sequence or an SIE (short intronic enhancer). In some aspects, the intron sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, or SEQ ID NO: 35.

[0014] In some aspects, the skeletal muscle selective hybrid regulatory sequences of the present disclosure comprise: a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 1; a CK6 promoter sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 8;

[0015] an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7 and SEQ ID NO: 35; and a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11.

[0016] In some aspects, the nucleic acid molecules of the present disclosure comprise a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 136.

[0017] In some aspects, skeletal muscle selective hybrid regulatory sequences of the present disclosure comprise: a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 2; a CK1 promoter sequence

comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 9;

[0018] an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7; and a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11.

[0019] In some aspects, the nucleic acid molecules of the present disclosure comprise a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 137.

[0020] In some aspects, the skeletal muscle selective hybrid regulatory sequences of the present disclosure comprise: a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 2; a CK7 promoter sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 10; an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7; and a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11.

[0021] In some aspects, the nucleic acid molecules of the present disclosure comprise a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 138.

[0022] In some aspects, the hybrid regulatory sequences of the present disclosure are in operable linkage with a nucleic acid sequence encoding a PUF polypeptide. In some aspects, the PUF polypeptide binds a CUG-repeat RNA sequence. In some aspects, the CUG-repeat RNA sequence is a *DMPK* CUG-repeat sequence.

[0023] In some aspects, the PUF polypeptides of the present disclosure comprise an amino acid sequence at least 95%, preferably at least 99%, identical to the amino acid sequence of any one of SEQ ID NO:133, SEQ ID NO:134, and SEQ ID NO:135. In some aspects, the PUF polypeptides of the present disclosure are encoded by a nucleic acid sequence that is at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 44, SEQ ID NO: 45, and SEQ ID NO: 46.

[0024] The present disclosure provides rAAV vectors comprising any of the nucleic acid molecules of the present disclosure.

[0025] The present disclosure provides pharmaceutical composition comprising any of the rAAV vectors of the present disclosure or any of the nucleic acid molecules of the present disclosure.

[0026] The present disclosure provides a method of treating myotonic dystrophy type 1 (DM1) in a subject, the method comprising administering to the subject a pharmaceutical composition of the present disclosure.

[0027] The present disclosure provides the pharmaceutical compositions described herein for the treatment of DM1 in a subject.

[0028] In some aspects, a subject has CUG microsatellite repeat expansion (MRE) in a DMPK RNA sequence.

[0029] In some aspects, the pharmaceutical compositions of the present disclosure are administered to the subject intravenously or intramuscularly, preferably wherein the pharmaceutical composition is administered intravenously.

[0030] The disclosure provides compositions and methods for treating myotonic dystrophy type 1 (DM1).

[0031] Disclosed herein is a nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence, and a CK6 (muscle creatine kinase 6) enhancer-promoter sequence.

[0032] In one embodiment, the hybrid regulatory sequence further comprises an intron sequence. In one embodiment, the intron sequence is an SV40 (simian virus 40) intron sequence or an SIE (short intronic enhancer) sequence.

[0033] In one embodiment, the hybrid regulatory sequence is in operable linkage with a PUF(CUG) sequence. In one embodiment, the PUF(CUG) sequence set forth in SEQ ID NO: 44, SEQ ID NO: 45, or SEQ ID NO: 46.

[0034] In one embodiment, the hybrid regulatory sequence and PUF(CUG) sequence is comprised within an AAV vector. In additional embodiments, the AAV vector is AAV9 or a AAV9 variant. In one embodiment the AAV9 variant is a myoAAV4 vector. In another embodiment, the AAV is a LBV201 capsid set forth in SEQ ID NO: 43.

[0035] In one embodiment, a cell comprises the AAV vector.

[0036] Also disclosed herein is a composition for treating DM1 comprising the hybrid regulator sequence in operable linkage with a PUF(CUG) sequence. In one embodiment, is a composition for treating DM1 comprising a nucleic acid sequence set forth in SEQ ID NO: 47.

[0037] Also disclosed herein is a method of treating myotonic dystrophy type 1 (DM1) in a mammal comprising administering a composition or AAV vector disclosed herein to a toxic

target CUG microsatellite repeat expansion (MRE) RNA sequence in tissues of the mammal whereby the level of expression of the toxic target RNA is reduced. In one embodiment, the composition or AAV vector is administered to the mammal subject intravenously or intramuscularly. In one embodiment, the reduced level of expression of the toxic target RNA thereby ameliorates symptoms of DM1 in the mammal. In one embodiment, the level of expression of the toxic target RNA is reduced compared to the reduction in the level of expression of untreated toxic target CUG RNA. In one embodiment, the level of reduction is between 1-fold and 20-fold.

[0038] Also disclosed herein is a nucleic acid molecule comprising a muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a promoter sequence, an enhancer sequence, and an intron sequence, wherein the enhancer sequence is a MYLPP (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence, and wherein the promoter sequence is muscle creatine kinase (CK) or desmin. In one embodiment, the CK promoter is selected from the group consisting of CK1, CK6, and CK7. In one embodiment, the Desmin promoter is selected from the group consisting of full-length Desmin and modified Desmin. In one embodiment, the intron is a SV40 intron or an SIE intron. In one embodiment, the muscle selective hybrid regulatory sequence is cardiac muscle selective. In one embodiment, the muscle selective hybrid regulatory sequence is skeletal muscle selective.

[0039] Also disclosed herein is an AAV capsid comprising LBV201 of SEQ ID NO: 54.

[0040] Any of the above aspects, or any aspect described herein can be combined with any other aspect.

[0041] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In the Specification, the singular forms also include the plural unless the context clearly dictates otherwise; as examples, the terms “a,” “an,” and “the” are understood to be singular or plural and the term “or” is understood to be inclusive. By way of example, “an element” means one or more element. Throughout the specification the word “comprising,” or variations such as “comprises” or “comprising,” will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%,

0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from the context, all numerical values provided herein are modified by the term “about.”

[0042] Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. The references cited herein are not admitted to be prior art to the claimed invention. In the case of conflict, the present Specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description and claim.

BRIEF DESCRIPTION OF THE DRAWINGS

[0043] FIG. 1 provides a table showing a series of hybrid regulatory sequences comprising human Desmin (hDesmin) promoters in AAV-based constructs driving PUF(CUG) in human DM1 patient-derived. A04298 is used as the control and is an ssAAV comprising the standard full-length (FL) hDesmin promoter driving Lambda. A02239 is an ssAAV comprising the standard FL hDesmin promoter driving PUF(CUG). A04949, A04959, A04546 comprise a modified (Mod) hDesmin promoter combined with a small intronic enhancer (SIE) driving PUF(CUG). A04949 is an scAAV but is otherwise the same as A04959. A04546 is the same as A04959 except that it comprises a myc tag on the PUF. These AAV9 viruses were used to transduce DM1 patient-derived cardiomyocytes that had 1150 CUG repeats and an MOI of 1e6 was applied. The cardiomyocytes were collected 7 days after transduction.

[0044] FIG. 2 shows a bar graph of PUF(CUG) RNA expression from the Desmin-based promoter series shown in Fig. 1 as measured by ddPCR. The constructs with the hybrid regulatory sequences comprised of the Mod hDesmin-SIE promoters exhibited an increase of expression by 5 to 16-fold over the hDesmin FL PUF(CUG) construct A02239.

[0045] FIG. 3 shows a bar graph of PUF protein expression from the Desmin-based promoter series shown in Fig. 1 as measured by Western blot. The hDesmin FL promoter drove minimal expression of PUF while the hybrid regulatory sequences comprised of the Mod hDesmin-SIE promoters drove robust expression. The higher level in the scAAV A104949 is expected because the onset of protein expression is more rapid from scAAV than

ssAAV. The PUF protein assay does not distinguish between PUF(CUG) and PUF-Lambda which is why there is a signal in the control A04298 sample.

[0046] FIG. 4A-4B provides electrophoresis images and bar graphs measuring enhanced correction of splicing markers (**FIG. 4A** shows markers: DMD Exon 78, LDB3 Exon 11, and CRTC2 Exon 13; **FIG. 4B** shows markers MBNL Exon 5, and cTNNT2 Exon 5) from the hybrid regulatory sequences comprised of the Mod hDesmin-SIE promoters shown in Fig. 1. RNA CUG repeats sequester MBNL proteins which result in abnormal splicing. The PUF protein competes and displaces MBNL from the CUG repeats, and the free MBNL is then able to properly modulate splicing. The RNA splicing from the human patient-derived cardiomyocytes was analyzed across 5 different transcripts. The upper panels show the splicing pattern for the given transcript and the bottom is the quantification. The hybrid regulatory sequences comprised of the Mod hDesmin-SIE promoters exhibit a much more potent splicing correction compared to the hDesmin FL PUF(CUG) construct A02239.

[0047] FIG. 5A-5B. FIG. 5A provides a table of the constructs A02239 and A04959. **FIG. 5B** shows the in vivo results of greater relative enhancement in cardiac muscle (7.7x) and in the gastrocnemius (3.2x) from the AAV9-based A04959 construct comprising the Mod hDesmin-SIE regulatory sequences compared to the A02239 construct comprising the FL hDesmin promoter. The constructs were intravenously injected into HSALR mice and PUF RNA expression was measured at 4 weeks post injection by ddPCR.

[0048] FIG. 6A-6B. FIG 6A provides a table of the constructs A02239 and A05288 which is an AAV9 construct comprising the hybrid regulatory sequences MYLPF-CK6-SV40intron driving PUF(CUG). **FIG. 6B** provides a bar graph of in vivo results showing greater skeletal and less cardiac expression with the A05288 construct comprising the hybrid regulatory sequences compared to the FL hDesmin construct. The constructs were intravenously injected into HSALR mice and PUF RNA expression was measured at 4 weeks post injection by ddPCR. AAV9 A05288 increased expression in the gastrocnemius and decreased expression 5x in the heart compared to the A02239 comprising the standard FL hDesmin promoter, a desirable expression profile for treating DM1.

[0049] FIG. 7A-7B. FIG. 7A provides a table of the construct A05288. **FIG. 7B** provides a bar graph showing expression in vivo continues to increase out to 8 weeks for the hybrid regulatory combination MYLPF-CK6-SV40intron. The construct was injected into HSALR mice and PUF RNA expression was measured at 8 weeks post injection by ddPCR. This data

is compared to the 4 weeks post injection expression data indicating that at 8 weeks post injection there was more PUF expression in both gastrocnemius and the heart.

[0050] FIG. 8 shows that construct A05288 is safe and well tolerated in vivo as there was no observed change in body weight compared to the vehicle control.

[0051] FIG. 9 shows correction of myotonia by AAV9 A05288. Functional myotonic dystrophy can be measured in mice by EMG studies of myotonia. In the gastrocnemius of the HSALR mice there is a high frequency of myotonic runs after stimulation consistent with the disease phenotype. The treatment with A05288 significantly decreased the frequency of myotonic runs indicating a functional correction of the disease.

[0052] FIG. 10 provides bar graphs indicating that the AAV9 A05288 construct results in a significant reduction in the CUG repeat containing RNA from the HSA transgene in the HSALR mouse model. RNA was harvested from the gastrocnemius and quadricep muscles of mice intravenously injected with AAV9 A05288 at both 4- and 9- weeks post injection. There was a significant reduction in the CUG repeat containing RNA at both time points with a slightly greater reduction at 8 weeks correlating with a reduction of the pathological RNA transcript in a DM1 patient.

[0053] FIG. 11 provides a bar graph measuring the in vivo results of % inclusion of Atp2a1 exon 22. RNA was harvested from the gastrocnemius and quadricep muscles of HSALR mice intravenously injected with AAV9 A05288 at 8 weeks post injection. The splicing of Atp2a1 exon 22 was analyzed because it is a reliable marker in the mouse model. The results show that AAV9 A05288 corrected splicing in both skeletal muscles (Gc and Quad).

[0054] FIG. 12A-12C provides bar graphs measuring PUF RNA expression from the AAV9 A05288 construct and a novel capsid, LBV201, A05288 construct in heart (FIG. 12A) skeletal muscle (FIG. 12B-12C) of HSALR mice by ddPCR. The constructs were injected into the mice at 1.2×10^{14} vg/kg. The novel LBV201 capsid resulted in higher PUF expression in both the cardiac and skeletal muscle tissues at 4 weeks post injection compared to AAV9 A05288 at 8 weeks. This demonstrates LBV201 is a more effective capsid for transducing muscle than AAV9 and that the capsid context can be utilized to predictably increase transgene expression from the novel combinations of the hybrid regulatory sequences disclosed herein.

[0055] FIG. 13A-13B shows in vivo A05288 PUF protein expression can be further increased by the use of a novel capsid such as LBV201. The constructs were intravenously injected into HSALR mice at 1.2×10^{14} vg/kg and PUF protein was measured by Western blot. The novel LBV201 capsid A05288 construct resulted in higher PUF expression in the

gastrocnemius (**FIG. 13A**) and quadriceps (**FIG. 13B**) 4 weeks post injection compared to AAV9 A05288 construct at 8 weeks. This further demonstrates LBV201 is a more effective capsid for transducing muscle than AAV9 and underscoring that the capsid context can be utilized to predictably increase transgene expression from the novel combinations of the hybrid regulatory sequences disclosed herein.

[0056] FIG. 14A-14B provides bar graphs showing greater in vivo molecular reduction of disease-causing transcript by the novel capsid LBV201 A05288 construct. RNA was harvested from the gastrocnemius (**FIG. 14A**) and quadriceps (**FIG. 14B**) muscles of mice intravenously injected with AAV9 A05288 and LBV201 A05288. The novel LBV201 capsid resulted in greater relative reduction of the pathological CUG repeat-containing transcript.

[0057] FIG. 15 provides a bar graph demonstrating more potent in vivo splicing correction by the novel capsid LBV201 A05288. RNA was harvested from the gastrocnemius and quadriceps muscles of mice intravenously injected with AAV9 A05288 and LBV201 A05288. The splicing of Atp2a1 exon 22 was analyzed because it is a reliable readout of splicing deficits in the HSALR model. LBV201 A05288 had a splicing correction at 4 weeks that was equivalent to AAV9 A05288 at 8 weeks post treatment.

[0058] FIG. 16 provides a gel image of showing transduction of human myotubes. Protein was harvested from transduced human skeletal myotubes at 7 days post transduction. While PUF protein from AAV9 A02239 was only detected from the 5e5 vg/cell dose there was robust PUF protein from the AAV9 A05288 and LBV201 A05288 constructs at the 5x lower dose of 1e5 vg/cell indicating that the novel capsid LBV201 can effectively transduce human myotubes.

[0059] FIG. 17A-17B. FIG. 17A provides a table showing a MyoAAV4A based-A05288 construct. **FIG. 17B** provides a bar graph showing that the construct potently reduces the pathological CUG repeat RNA in the HSALR mouse model. RNA was harvested from the gastrocnemius, quadriceps, tibialis anterior, and triceps muscles of mice intravenously injected with MyoAAV4A A05288 4 weeks post injection at a dose of 8E13 vg/kg. There was a significant reduction in CUG repeat containing RNA in all the skeletal muscles analyzed.

[0060] FIG. 18A-18B. FIG. 18A provides a table showing the MyoAAV4A A05288 construct. **FIG. 18B** provides a bar graph showing MyoAAV4A A05288 corrects splicing in the HSALR mouse model. RNA was harvested from the gastrocnemius, quadriceps, tibialis anterior, and triceps muscles of mice intravenously injected with MyoAAV4A A05288 4 weeks post injection at a dose of 8E13 vg/kg. The splicing of Atp2a1 exon 22 was analyzed

because it is a reliable readout of splicing deficits in the HSALR model. There was significant splicing correction in all muscles except the triceps which had a high baseline (non-pathologic) level Atp2a1 exon 22 inclusion.

[0061] FIG. 19 provides a bar graph showing in vivo PUF expression from low dose MyoAAV4A A05288 is greater than high dose AAV9 A05288. PUF expression at the protein level from AAV9 and MyoAAV4A mediated delivery of A05288 was compared. PUF protein expression was much greater in the MyoAAV4A at 4 weeks than AAV9 at 8 weeks. MyoAAV4A A05288 expression continued to increase from 4 to 8 weeks indicating good durability of high expression.

[0062] FIG. 20 provides a composite splicing index showing robust correction of DM1-associated alternative splicing of 19 exons with A05288 treatment in various capsids as the composite index neared wildtype (WT) levels. RNA-Seq was performed on HAS-L4 skeletal muscle after intravenous treatment with AAV9, MyoAAV4A, or LBV201 capsids packaged with MYLPF-CK6-SV40intron-PUF(CUG) construct A05288. Alternative splicing events were evaluated with Olego (open source, from Zhang Lab, Columbia University) and a composite score of splicing index (exon inclusion fraction) was determined for 19 known DM1-associated missplicing events. (Tanner et al. M.K. et al. 2021. Targeted splice sequencing reveals RNA toxicity and therapeutic response in myotonic dystrophy. Nucleic Acids Res. 49(4): 2240-2254.)

[0063] FIG. 21A-21B. FIG. 21A provides a dosage table of the A05288 constructs: DM35, DM36, DM39 and MN13. FIG. 21B provides a heatmap showing splicing patterns of skeletal muscle treated with intravenous A05288 (per the dosage table of FIG. 12A) appeared similar to wildtype. RNA-Seq was performed on HAS-L4 skeletal muscle after intravenous treatment with AAV9, MyoAAV4A, or LBV201 capsids packaged with MYLPF-CK6-SV40intron-PUF(CUG) construct A05288. Alternative splicing events were evaluated with Olego (open source, from Zhang Lab, Columbia University) and a composite score of splicing index (exon inclusion fraction) was determined for known DM1-associated missplicing events. (Tanner et al. M.K. et al. 2021. Targeted splice sequencing reveals RNA toxicity and therapeutic response in myotonic dystrophy. Nucleic Acids Res. 49(4): 2240-2254.)

DETAILED DESCRIPTION

[0064] The disclosure provides, *inter alia*, RNA-targeting gene therapy compositions and methods for treating myotonic dystrophy type 1 (DM1). The disclosure also provides novel

promoter-enhancer-intron combinations which drive selective and preferential expression in certain muscle cells over other types of muscle cells.

[0065] DM1 is a multisystemic, autosomal-dominant inherited disorder caused by CTG microsatellite repeat expansions (MREs) in the 3' untranslated region of the *DMPK* gene. RNA transcripts containing the CUG repeat expansions sequester muscleblind-like (MBNL) proteins which are the regulators of the alternative splicing switch from fetal to adult isoforms. Adult DM1 patients experience debilitating myotonia and progressive weakness in skeletal muscle while infants born with DM1 (congenital DM or CDM) are hypotonic and display respiratory insufficiency. The *DMPK* gene encodes a protein called myotonic dystrophy protein kinase which is believed to play a role in muscle, heart, and brain cells. The protein may be involved in communication within cells. It also appears to regulate the production and function of important structures inside muscle cells by interacting with other proteins. For example, myotonic dystrophy protein kinase has been shown to inhibit part of a muscle protein called myosin phosphatase. Myosin phosphatase is an enzyme that plays a role in muscle tensing (contraction) and relaxation.

[0066] One region of the *DMPK* gene contains a segment of three DNA building blocks (nucleotides) that is repeated multiple times. This sequence, which is written as CTG, is called a triplet or trinucleotide repeat. In most unaffected people, the number of CTG repeats in this gene ranges from 5 to 34. In DM1 patients, there is a CTG repeat expansion which increases the size of the CTG repeat in the *DMPK* gene. DM1 is classified as either adult-onset or as congenital forms that are distinguished by the size of the expanded CTG tract. Repeats in such CTG repeat expansions can range from about 50 to about 1,000 CTG repeats in most cells and in certain cell types, such as muscle cells, the number of repeats are typically greater. Indeed, the size of the trinucleotide repeat expansion is associated with the severity of signs and symptoms of DM1. Classic features such as muscle weakness and wasting beginning in adulthood and correlate with about 100 to about 1,000 CTG repeats per cell. The more severe congenital form of DM1 tends to correlate with over 1,000 CTG repeats per cell. The mild form of DM1 typically ranges from about 50 to about 150 CTG repeats per cell. RNA produced from the *DMPK* gene, including pre-mRNA and mRNA, thus comprises CUG repeat sequences corresponding to the repeats observed within the *DMPK* gene. Further, pre-mRNA and mRNA transcripts of the *DMPK* gene in subjects having DM1 will have the expanded number of CUG repeats as set forth above. As would be appreciated by the skilled artisan, the number or the approximate number of CUG repeats in a given pre-

mRNA or mRNA transcript can be represented as CUGⁿ, wherein the superscript *n* represents the number of repeats. The repetitive RNAs produced by the DMPK locus form nuclear RNA foci that sequester RNA binding proteins such as MBNL1 (Muscleblind Like Splicing Regulator 1) and divert them from their homeostatic RNA processing activities. Loss of MBNL1 function is linked to hundreds of alternative splicing defects and respiratory insufficiency which contribute in varying degrees to patient mortality. Targeting and eliminating (or blocking) CUG repeats is a therapeutic strategy for DM1.

[0067] The RNA-targeting gene therapy compositions disclosed herein provide efficacious binding and therefore blocking of toxic CUG repeats in methods of treating DM1.

[0068] In some embodiments, gene therapy compositions of the present disclosure comprise at least one PUF polypeptide capable of binding a target RNA sequence comprising CUG repeats. Without wishing to be bound by theory, these PUF polypeptides can bind expanded CUG repeat (CUG^{exp}) RNA directly and block MBNL sequestration to preserve near normal free MBNL levels and function that will reverse DM1 disease phenotypes such as splicing dysfunction, myotonia and others. In some embodiments, compositions suitable for blocking CUG-repeat RNA bind a CUG-repeat containing RNA and prevent translation of the CUG-repeat RNA. In some embodiments, this blocked translation results in reduced protein expression from CUG-repeat containing RNA sequences. The PUF polypeptides of the present disclosure that are capable of binding to a target RNA sequence comprising CUG repeats are referred to herein as “PUF(CUG)”, “PUF(CUG) polypeptides”, “PUF(CUG) RNA-targeting constructs”, and “PUF(CUG)-repeat targeting system”.

[0069] As described herein, CTG microsatellite expansions in the non-coding 3' untranslated region of *DMPK* cause DM1. Expanded CUG (CUG^{exp}) repeats in *DMPK* mRNA directly sequester MBNL proteins resulting in loss of function. MBNL loss of function is directly responsible for alternative splicing defects and clinical manifestations observed in DM1. The PUF(CUG) polypeptides of the present disclosure bind CUG^{exp} RNA directly and block MBNL sequestration to preserve near normal free MBNL levels and function leading to the reversal of DM1 disease phenotypes such as splicing dysfunction, myotonia and others. The following PUF(CUG) RNA-targeting constructs are exemplary embodiments for blocking CUG repeat DMPK RNA sequences.

[0070] In some embodiments, the PUF(CUG)-repeat targeting system targets expanded CUG repeats (CUG^{exp}), wherein the CUG repeats are CUG⁵⁰ repeats or more. In some embodiments, the CUG repeats are CUG¹⁰⁰ repeats or more. In some embodiments, the CUG

repeats are CUG⁵⁰⁰ repeats or more. In some embodiments, the CUG repeats are CUG⁹⁶⁰ repeats. In some embodiments, the CUG¹⁰⁰⁰ repeats are 1,000 CUG repeats or more. To be clear, CUG⁵⁰ or CUG¹⁰⁰ or CUG⁹⁶⁰ or CUG¹⁰⁰⁰ refers to 50 CUG repeats or 100 CUG repeats or 960 CUG repeats or 1,000 CUG repeats, respectively, in a CUG repeat containing gene. Any other number or range of CUG repeats are possible, including 50, 55, 60, 65, 70, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 90, 95, 100, 105, 110, 115, 120, 150, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 960, 1,000, 2,000, 3,000, 4,000, or 5,000 CUG repeats, or any other number of CUG repeats in between.

[0071] In any of the preceding or subsequent RNA-targeting compositions for treating DM1, any particular construct element (e.g., linker, promoter, signal sequence, etc.) described in the context of a specific RNA-targeting composition, can be substituted for another of the same element type (e.g., linker, promoter, signal sequence, etc.). In some embodiments, any particular construct element can be omitted or removed (such as a tag sequence). In other words, the exemplary combinations of elements in any particular gene therapy composition described herein is not intended to be limiting.

[0072] *Preferential and Selective Tissue-specific Hybrid Regulatory Sequences*

[0073] Disclosed herein are gene regulatory sequences including enhancers, core promoters, and intronic sequences which can be modularly assembled to form hybrid regulatory sequences to achieve controlled expression of nucleic acid sequences in cardiac and skeletal muscle. Hybrid regulatory sequences of the disclosure can be used, for example, in vector-based delivery platforms including rAAV vectors. the Modularly assembling various regulatory components to form hybrid regulatory sequences can fine-tune expression of a target nucleic acid or transgene in a target tissue. Hybrid regulatory sequences utilizing the regulatory sequences disclosed herein for tissue selective expression of a transgene or RNA-targeting composition disclosed herein can be combined with various AAV capsids as well as various routes of administration to yield a more potent and efficacious therapeutic composition..

[0074] The disclosure provides gene therapy compositions comprising potent muscle-specific hybrid regulatory sequences.

[0075] Disclosed herein are compositions comprising nucleic acid molecules and vectors comprising hybrid muscle-specific regulatory sequences.

[0076] Hybrid regulatory sequences can include, without limitation, the combination of two or more regulatory sequences such as promoter(s), enhancer(s), and/or intronic sequence(s).

In some embodiments, the two or more regulatory sequences are heterologous (i.e. derived from different genetic loci).

[0077] A “promoter” is a control sequence that is a region of a polynucleotide sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. In some embodiments, a promoter may be comprised of two or more genetic elements. In some embodiments, the two or more genetic elements are heterologous (i.e. derived from different genetic loci). In some embodiments, a core promoter sequence is a minimal region of a promoter which is derived from a corresponding standard promoter sequence which retains the capability of recruiting RNA Polymerase II for RNA transcription.

[0078] An “intron” or “small intronic enhancer” or “SIE” or “modified intron” or “modified intron sequence” as disclosed herein is derived from an intron present in an abundantly expressed gene and modified via truncation and/or mutation in such a manner so as increase the ability of the modified intron to enhance and/or sustain transcription and nuclear export of a transcript. In one embodiment, a modified intron is derived from an intron of a gene which is abundantly expressed and/or the intron is known to be a key regulator of the expression of the corresponding gene of the intron. In another embodiment, an intron sequence or modified intron sequence is derived from SV40 (Simian Virus 40), MVM (Minute Virus of Mice), EF1 alpha, or mouse creatine kinase. In one embodiment, a modified intron comprises mutations of start and/or alternative start codons.

[0079] In some embodiments, the hybrid regulatory sequences of the present disclosure can comprise at least one SIE sequence. In some embodiments, an SIE sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 5. In some embodiments, an SIE sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 6.

[0080] In some embodiments, a hybrid regulatory sequence of the disclosure can comprise at least one SV40 intron sequence. In some embodiments, an SV40 intron sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 7. In some embodiments, an SV40 intron sequence comprises,

consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 35.

[0081] An “enhancer sequence” as disclosed herein is an enhancer region derived from enhancers known to aid in the recruitment of RNA Polymerase II. In some embodiments, an enhancer sequence is a regulatory element that increases the expression of a target sequence.

[0082] A “modified enhancer region” or “modified enhancer sequence” as disclosed herein is an enhancer region derived from enhancers known to aid in the recruitment of RNA Polymerase II and altered via truncation and/or mutation in such a manner so as to improve recruitment of RNA Polymerase II as compared to an unmodified enhancer region.

[0083] A hybrid regulatory sequence disclosed herein is operably linked to a nucleic acid sequence of interest (NOI), transgene, or therapeutic cargo. Within the context of a recombinant expression vector, the terminology “operably linked” is intended to mean that the hybrid regulatory sequence is linked to an NOI, transgene, or therapeutic cargo in a manner permitting expression of the nucleotide sequence in, for example, a host cell when the vector is introduced into (or in contact with) the host cell. In one embodiment, the NOI/therapeutic cargo is a PUF polypeptide sequence.

[0084] Hybrid regulatory sequences of the disclosure can comprise at least one kozak sequence. In some embodiments, the kozak sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 11. In some embodiments, the kozak sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 18. In some embodiments, the kozak sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 30. Hybrid regulatory sequences of the present disclosure can comprise at least one hMYPLF enhancer sequence. In some embodiments, an hMYPLF enhancer sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to any one of SEQ ID NOs: 1-3.

[0085] Hybrid regulatory sequences of the present disclosure can comprise at least one desmin enhancer sequence. In some embodiments, a desmin enhance sequence comprises,

consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 12 or SEQ ID NO: 14.

[0086] Hybrid regulatory sequences of the present disclosure can comprise at least one desmin proximal sequence. In some embodiments, a desmin proximal sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 13 or SEQ ID NO: 15.

[0087] Hybrid regulatory sequences of the present disclosure can comprise at least one desmin sequence. In some embodiments, a desmin sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 16.

[0088] Hybrid regulatory sequences of the present disclosure can comprise at least one human desmin reporter sequence. In some embodiments, a desmin reporter sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 22.

[0089] The hybrid regulatory sequences disclosed herein are comprised of the regulatory sequences set forth in Table 1A used alone or in various combinations. Hybrid regulatory sequences of the disclosure can comprise other regulatory sequences known in the art in combination with the sequences set forth in Table 1A. Hybrid regulatory sequences are set forth in Table 1B. The orientation of regulatory sequences within a nucleic acid construct can be in any orientation from 5' to 3' in an expression construct, the components including without limitation, 1) a promoter sequence, or promoter region thereof, 2) a promoter/enhancer combination sequence, 3) an intronic sequence, and/or 3) an enhancer sequence. In some embodiments, the orientation of a hybrid regulatory sequence from 5' to 3' in an expression construct is 1) an enhancer sequence, 2) a promoter sequence, and 3) an intron sequence. In some embodiments, the orientation of the hybrid regulatory sequence from 5' to 3' in an expression construct is 1) an enhancer sequence, 2) a promoter/enhancer sequence, and 3) an intron sequence.

Table 1A: Regulatory Sequences:

Regulatory sequences	Sequences (SEQ ID NO)
----------------------	-----------------------

hMYPLF enhancer (Longer form)	TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATG TCCTGTCAATTCCCACTCAGTGTCTGATCTCCTTCTCCTCACC TTTCCCATCTCCCGTTTGACCCAGCTTCCTGAGCTCTCCTCCC ATTCCCCTTTTTGGAGTCCTCCTCCTCTCCCAGAACCCAGTAA TAAGTGGGCTCCTCCCTGGCCTGGACCCCCGTGGTAACCCTA TAAGGCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGT GTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGGTGGCTCCTT CTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCC AGAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAG GAGGGGGGACAGGGCAGGGGGAGGCATCTTCTCAGGACAT CGGGTCCTAGAGG (SEQ ID NO: 1)
hMYPLF enhancer - 111 bp (Short form 1)	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGC AGCTGCTAAGTTTAGGGTGGCTCCTTCT (SEQ ID NO: 2)
hMYPLF enhancer - 60 bp (Short form 2)	GCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCTAA GTTTAGGGTGGCTCCTTCT (SEQ ID NO: 3)
hMYPLF enhancer - 51 bp (Short form 3)	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAG (SEQ ID NO: 4)
SIE	Caggtaaagttggctgtaacaaagttgaatttgagttgataagtagtactgtctgccgccagattctaaaaa taaccacctggacagcagcaccgaaggtcttggcgggtccctgttctcttccctggcttcagctca ccgtgggtccattcataggtatttagtatgctgtatatattttcaggt (SEQ ID NO: 5)
SIE	aggtaaagttggctgtaacaaagttgaatttgagttgataagtagtactgtctgccgccagattctaaaaat aaccacctggacagcagcaccgaaggtcttggcgggtccctgttctcttccctggcttcagctcac cgtgggtccattcataggtatttagtatgctgtatatattttcaggt (SEQ ID NO: 6)
SV40 intron	AAGTTTAGTCTTTTTGTCTTTTATTTACAGGTCCCAGATCTGGT GGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGCTTCTAGGT (SEQ ID NO: 7)
SV40 intron	CAGGTAAGTTTAGTCTTTTTGTCTTTTATTTACAGGTCCCAGAT CTGGTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGCTTC TAGGT (SEQ ID NO: 35)
CK6 enhancer/ promoter (mutated)	ccactacgggtctaggtgccccatgaaggaggcaaggcctggggacacccgagatgcctggttat aattaaccccaacacgtgctcccccccccccccaacacgtgctgagcctgagcgggtacccca ccccggtgcttgggtcttaggtctgtacacatggaggagaagctgctctaaaaataacacctgtcc ctggtggatGGGCCCaaggctgtgggggactgagggcaggctgtaacaggcttgggggcca gggttatacgtgctgggactccaaagttactgttccatgttccggcggaagggccagctgtccc ccgccagctagactcagcacttagtttaggaaccagttagcaagtcagcccttggggcagccatac aaggccatggggctgggcaagctgcacgcctgggtcgggggtgggcacgggtcccggggaacga gctgaaagctcatctgcttcaggggccccctccctggggacagccccctctgggtagtcacacctgt aggctccttatataaccaggggcacagggtgctgcccCCGGGTCACCACC (SEQ ID NO: 8)
CK1 enhancer/ promoter	CCACTACGGGTCTAGGCTGCCCATGTAAGGAGGCAAGGCCTG GGGACACCCGAGATGCCTGGTTATAATTAACCCAGACATGTG GCTGccccccccccAACACCTGCTGCCTGAGCCTCACccccccccggt gcctgggtcttaggtctgtacacatggaggagaagctgctctaaaaataacctgtccctggtg atGGGCCCaaggctgtgggggactgagggcaggctgtaacaggcttgggggccaagggtta tacgtgctgggactccaaagttactgttccatgttccggcggaagggccagctgtccccgcca gctagactcagcacttagtttaggaaccagttagcaagtcagcccttggggcagccatacaaggcc

	atggggctgggcaagctgcacgcctgggtccggggtgggcacggtgcccgggcaacgagctgaa agctcatctgctctcaggggccccctcctggggacagcccccttggtatgtacacctgtaggtctc ctctatataaccaggggacaggggtgcccCCGGGTCACCACC (SEQ ID NO: 9)
CK7 enhancer/ promoter w/ 3' sequences	Ccactacgggtctaggtgccccatgtaaggaggcaaggcctggggacaccgagatgcctggtat aattaaccagacatgtggtgcccccccccccccaacacctgtgcctctaaaaataacctgtccct ggtggatcccctgcatgcgaagatcttgaacaaggctgtgggggactgagggcaggctgtaacag gcttggggggccagggttatactgctgggactccaaagtattactgtccatgttcccggaag ggccagctgtcccccgccagctagactcagcacttagtttaggaaccagtgagcaagtacgcccttg gggcagcccatacaaggccatggggctgggcaagctgcacgcctgggtccggggtgggcacggt gcccgggcaacgagctgaaagctcatctgctctcaggggccccctcctggggacagccccctctgg ctagtcacacctgtaggctcctctatataaccaggggacaggggtgcctcattctaccaccac ctccacagcaccag (SEQ ID NO: 10)
Ck8e promoter	TGCCCATGTAAGGAGGCAAGGCCTGGGGACACCCGAGATGC CTGGTTATAATTAACCCAGACATGTGGCTGCCCCCCCCCCCC AACACCTGCTGCCTCTAAAAATAACCCTGCATGCCATGTTCC CGGCGAAGGGCCAGCTGTCCCCGCCAGCTAGACTCAGCACT TAGTTTAGGAACCAGTGAGCAAGTCAGCCCTTGGGGCAGCCC ATACAAGGCCATGGGGCTGGGCAAGCTGCACGCCTGGGTCC GGGTGGGCACGGTGCCCGGGCAACGAGCTGAAAGCTCATC TGCTCTCAGGGGCCCCCTCCCTGGGGACAGCCCCCTCTGGCTA GTCACACCCTGTAGGCTCCTCTATATAACCCAGGGGCACAGG GGCTGCCCTCATTCTACCACCACCTCCACAGCACAGACAGAC ACTCAGGAGCCAGCCAGC (SEQ ID NO: 17)
Desmin enhancer (full length)	ACCCATGCCTCCTCAGGTACCCCTGCCCCACAGCTCCTCT CCTGTGCCTTGTTTCCCAGCCATGCGTTCTCCTCTATAAATAC CCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAG ATGGTTGGGTGACATGCGGCTCCTGACAAAACACAAACCCC TGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATC AGGGAGGGGGCGGGGGACCCAGGGGGCAGGAGCCACACAA AGTCTGTGCGGGGGTGGGAGCGCACATAGCAATTGGAACT GAAAGCTTATCAGACCCTTTCTGGAAATCAGCCCACTGTTTA TAAACTTGAGGCCCCACCCTCGA (SEQ ID NO: 12)
Desmin proximal (full length)	GATAACCAGGGCTGAAAGAGGCCCCGCCTGGGGGCTGCAGAC ATGCTTGCTGCCTGCCCTGGCGAAGGATTGGCAGGCTTGCCC GTCACAGGACCCCCGCTGGCTGACTCAGGGGCGCAGGCCTCT TGCGGGGGAGCTGGCCTCCCCGCCCCACGGCCACGGGCCGC CCTTTCCTGGCAGGACAGCGGGATCTTGACGCTGTCAGGGGA GGGGAGGCGGGGGCTGATGTCAGGAGGGATAAATAAGTGC CGACGGCTGGGGGGCCCTGTCTCCCTCGCCGCATCCACTCTC CGGCCGGCCGCTGCCCGCCGCTCCTCCGTGCGCCCGCCAG CCTCGCCCCGCGCCGTC (SEQ ID NO: 13)
Desmin enhancer (truncated)	CCCACAGCTCCTCTCCTGTGCCTTGTTTCCCAGCCATGCGTTC TCCTCTATAAATACCCGCTCTGGTATTGGGGTTGGCAGCTGT TGCTGCCAGGGAGATGGTTGGGTGACATGCGGCTCCTGACA AAACACAAACCCCTGGTGTGTGTGGGCGTGGGTGGTGTGAGT AGGGGGATGAATCAGGGAGGGGGCGGGGGACCCAGGG(SEQ ID NO: 14)

Desmin proximal (truncated)	AGGGGCGCAGGCCTCTTGCGGGGGAGCTGGCCTCCCCGCCCC CACGGCCACGGGCCGCCCTTTCTTGGCAGGACAGCGGGATCT TGCAGCTGTCAGGGGAGGGGAGGCGGGGGCTGATGTCAGGA GGGATAtAAATAGTGCCGACGGCTGGGGGCCCTGTCTCCCCT CGCCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCCTCC TCCGTGCGCCCCGCCAGCCTCGCCCCGCGCCGTC (SEQ ID NO: 15)
Desmin (1kb)	taccacctgccccccacagctcctctctgtgccttgtttccagccatgcgttctctataataatccc gctctggtatttgggttgagcagctgtgtgcagggagatggttgggttgacatgcggctcctgaca aaacacaaaccctggtgtgtgtggcggtggtgtgtgagtaggggatgaatcaggaggagggg cAggggacccagggggcaggagccacacaaagtctgtgcgggggtgggagcgacatagcaat tggaaactgaaagcttatcagacccttctggaaatcagccactgtttataaactgaggccccaccct cgacagtaccggggaggaagaggcgctgactagtccagaggaaactgaggctcagggcCag ctcgcccatagacatacatggcaggcaggctttggccagatccctccgctgcccaggcgtctccct gcccctcccttctgcctagagacccccaccctcaagcctggctggttttgcctgagacccaaacctt tcgacttcaagagaatatttaggaacaaggtggttagggccttctgggaacaggccttgaccctta agaaatgaccacaaagtctctccttgaccacaaaaggggaccctcaactaaagggaagcctcttctg ctgtctcccctgacccactccccccacccaggacgaggagataaccagggtgaaagaggcc cgcttgggggctgcagacatgcttgcctgcccggcgaaggattggcaggcttgcctgcacag gacccccgctggctgactcaggggcgagggcctTttcgggggagctggcctccccgccccac ggccacgggccccttcttggcaggacagcgggatcttcagctgtcaggggaggggagggcg ggggctgatgtcaggaggagatacaaatagtgcgacggctggggggccctGTCTCCCCCTCG CCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCCTCCTC CGTGCGCCCCGCCAGCCTCGCCCCG (SEQ ID NO: 16)
Short human desmin reporter	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCTTGTTTCC CAGCCATGCGTTCTCCTCTATAAATACCCGCTCTGGTATTTGG GGTTGGCAGCTGTTGCTGCCAGGGAGATGGTTGGGTTGACAT GCGGCTCCTGACAAAACACAAACCCCTGGTGTGTGTGGGCGT GGGTGGTGTGAGTAGGGGGATGAATCAGGGAGGGGGCGGGG GACCCAGGGAGGGGCGCAGGCCTTTTGCGGGGGAGCTGGCC TCCCCGCCCCACGGCCACGGGCCGCCCTTTCTTGGCAGGAC AGCGGGATCTTGCAGCTGTCAGGGGAGGGGAGGCGGGGGCT GATGTCAGGAGGGATAtAAATAGTGCCGACGGCTGGGGGCC TGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCC GCCGCTCCTCCGTGCGCCCCGCCAGCCTCGCCCCGCGCCGTCA CGCGTc (SEQ ID NO: 22)
Kozak 1	Gccgccacc (SEQ ID NO: 11)
Kozak 2	gccACCatgg (SEQ ID NO: 18)
Kozak 3	gccgccACCatg (SEQ ID NO: 30)

Table 1B: Hybrid Regulatory Sequences:

Hybrid Regulatory Sequences	Sequences (SEQ ID NO)
MYPLF + CK6 + SV40 intron + Kozak 1	TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATG TCCTGTCAATTCCCACTCAGTGTCTGATCTCCTTCTCCTCACC TTTCCATCTCCCGTTTGACCCAGCTTCTGAGCTCTCCTCCC

	ATTCCCCTTTTGGAGTCCTCCTCCTCTCCCAGAACCCAGTAA TAAGTGGGCTCCTCCCTGGCCTGGACCCCGTGGTAACCCTA TAAGGCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGT GTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGGTGGCTCCTT CTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCC AGAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAG GAGGGGGGACAGGGCAGGGGGAGGCATCTTCCTCAGGACAT CGGGTCCTAGAGGccactacgggtctaggtgccatgtaaggaggcaaggcctggg acacccgagatgctgtgtataattaacccaacacctgctgcccccccccaacacctgctgct gagcctgagcggttaccacacccgggtgctgggtcttaggtctgtacaccatggaggagaagctc gctctaaaaataacacctgtccctgggtgatGGGCCcaaggctgtgggggactgagggcaggc tgtaacaggcttgggggccagggttatactgtcctgggactccaaagtattactgttccatgttccc ggcgaagggccagctgtccccgccagctagactcagcacttagtttaggaaccagttagcaagtc agcccttggggcagccatacaaggccatggggctgggcaagctgcacgcctgggtccgggtgg gcacggtgcccgggcaacgagctgaaagctcatctgctctcaggggccctccctggggacagcc cctcctggctagtcacacacctgtaggctcctctatataaccaggggcacagggtgcccCCGG GTCACCACCACGCGTcagGTAAGTTTAGTCTTTTTGTCTTTTAT TTCAGGTCCCAGATCTGGTGGTGGTGCAAATCAAAGAACTGC TCCTCAGTGGCTTCTAGGTGCTAGCgcccgcacc (SEQ ID NO: 19)
MYPL + CK1 + SV40 intron + Kozak 1	TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATG TCCTGTCAATTCCCACTCAGTGTCTGATCTCCTTCTCCTCACC TTTCCCATCTCCCGTTTGACCCAGCTTCCTGAGCTCTCCTCCC ATTCCCCTTTTGGAGTCCTCCTCCTCTCCCAGAACCCAGTAA TAAGTGGGCTCCTCCCTGGCCTGGACCCCGTGGTAACCCTA TAAGGCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGT GTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGGTGGCTCCTT CTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCC AGAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAG GAGGGGGGACAGGGCAGGGGGAGGCATCTTCCTCAGGACAT CGGGTCCTAGAGGCCACTACGGGTCTAGGCTGCCCATGTAAG GAGGCAAGGCCTGGGGACACCCGAGATGCCTGGTTATAATT AACCCAGACATGTGGCTGccccccccccAACACCTGCTGCCTGA GCCTCACccccacccgggtgctgggtcttaggtctgtacaccatggaggagaagctcgtct aaaaataacacctgtccctgggtgatGGGCCcaaggctgtgggggactgagggcaggctgtaa caggcttgggggccagggttatactgtcctgggactccaaagtattactgttccatgttcccggcg aagggccagctgtccccgccagctagactcagcacttagtttaggaaccagttagcaagtcagccc ttggggcagccatacaaggccatggggctgggcaagctgcacgcctgggtccgggtgggcaag gtgcccgggcaacgagctgaaagctcatctgctctcaggggccctccctggggacagcccctct ggctagtcacacacctgtaggctcctctatataaccaggggcacagggtgcccCCGGGTC ACCACCACGCGTcagGTAAGTTTAGTCTTTTTGTCTTTTATTTT AGGTCCCAGATCTGGTGGTGGTGCAAATCAAAGAACTGCTCC TCAGTGGCTTCTAGGTGCTAGCgcccgcacc (SEQ ID NO: 20)
MYPLF + CK7 + SV40 intron + Kozak 1	TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATG TCCTGTCAATTCCCACTCAGTGTCTGATCTCCTTCTCCTCACC TTTCCCATCTCCCGTTTGACCCAGCTTCCTGAGCTCTCCTCCC ATTCCCCTTTTGGAGTCCTCCTCCTCTCCCAGAACCCAGTAA TAAGTGGGCTCCTCCCTGGCCTGGACCCCGTGGTAACCCTA TAAGGCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTGGT GTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGGTGGCTCCTT

	CTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCAGTGCCC AGAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAG GAGGGGGGACAGGGCAGGGGGAGGCATCTTCCTCAGGACAT CGGGTCCTAGAGGccactacgggtctaggtgccccatgtaaggaggcaaggcctggg acacccgagatgctggttataattaaccagacatgtggctgccccccccccccacacctgctgc ctctaaaaataaccctgtccctggatccctgcatgcaagatcttgaacaaggctgtggggga ctgagggcaggctgtaacaggcttggggccagggttatacgtgctgggactccaaagtattac tgtccatgttccggcgaaggggccagctgtccccgccagctagactcagcacttagtttaggaacc agtgaagcaagtcagcccttggggcagccatacaaggccatggggctgggcaagctgcacgctg ggctccgggtgggcacgggtgccgggcaacgagctgaaagctcatctgctctcaggggccccctcc ctggggacagccccctctggctagtacacccctgtaggctcctctatataaccaggggcacagggg ctgcccctattctaccaccacccacagcaccagACGCGTcagGTAAGTTTAGTCT TTTTGTCTTTTATTTTCAGGTCCCAGATCTGGTGGTGGTGCAA TCAAAGAACTGCTCCTCAGTGGCTTCTAGGTGCTAGCgccgccac c (SEQ ID NO: 21)
hMYPLF enhancer - 111 bp (Short form 1) + CK6 promoter + SV40 intron + kozak 1	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGC AGCTGCTAAGTTTAGGGTGGCTCCTTCTccactacgggtctaggtgccccat gtaaggaggcaaggcctggggacacccgagatgctggttataattaacccaacacctgctgccc cccccccaacacctgctgctgagcctgagcggttaccacccccggctgctggtcttaggtct gtacaccatggaggagaagctcgtctaaaaataaccctgtccctgggtggaGGCCCaaggc tgtgggggactgagggcaggctgtaacaggcttggggccagggttatacgtgctgggactccc aaagtattactgttccatgttccggcgaaggggccagctgtccccgccagctagactcagcacttagt ttaggaaccagtgagcaagtcagcccttggggcagccatacaaggccatggggctgggcaagctg cacgctgggtccgggtgggcacgggtgccgggcaacgagctgaaagctcatctgctctcaggg gccccctctggggacagccccctctggctagtacacccctgtaggctcctctatataaccaggggc acaggggctgcccCCGGGTACCAACACGCGTcagGTAAGTTTAGTCT TTTTGTCTTTTATTTTCAGGTCCCAGATCTGGTGGTGGTGCAA ATCAAAGAACTGCTCCTCAGTGGCTTCTAGGTGCTAGCgccgcc acc (SEQ ID NO: 136)
hMYPLF enhancer - 111 bp (Short form 1) + CK1 promoter + SV40 intron + kozak 1	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGC AGCTGCTAAGTTTAGGGTGGCTCCTTCTCACTACGGGTCTAG GCTGCCCCATGTAAGGAGGCCAAGGCCTGGGGACACCCGAGAT GCCTGGTTATAATTAACCCAGACATGTGGCTGccccccccccAA CACCTGCTGCCTGAGCCTCACccccccccggctggtggtcttaggtctgtaca ccatggaggagaagctcgtctaaaaataaccctgtccctgggtggaGGCCCaaggctgtg gggactgagggcaggctgtaacaggcttggggccagggttatacgtgctgggactccaaagt attactgttccatgttccggcgaaggggccagctgtccccgccagctagactcagcacttagtttagg aaccagtgagcaagtcagcccttggggcagccatacaaggccatggggctgggcaagctgcacg cctgggtccgggtgggcacgggtgccgggcaacgagctgaaagctcatctgctctcaggggccc ctccctggggacagccccctctggctagtacacccctgtaggctcctctatataaccaggggcacag gggctgcccCCGGGTACCAACACGCGTcagGTAAGTTTAGTCTTT TTGTCTTTTATTTTCAGGTCCCAGATCTGGTGGTGGTGCAAATC AAAGAACTGCTCCTCAGTGGCTTCTAGGTGCTAGCgccgccacc (SEQ ID NO: 137)
hMYPLF enhancer - 111 bp (Short form 1) + CK7	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGC AGCTGCTAAGTTTAGGGTGGCTCCTTCTccactacgggtctaggtgccccat

promoter + SV40 intron + kozak 1	gtaaggaggcaaggcctggggacacccgagatgcctggtataattaaccagacatgtggctgcc cccccccccaacacctgtgcctctaaaaataacctgtccctgggtgatccctgcatgcgaagat cttcgaacaaggctgtgggggactgagggcaggctgtaacaggctgggggccagggttatactg gcctgggactcccaagattactgttccatgttcccggaagggccagctgtccccgccagctag actcagcacttagtttaggaaccagtgcgaagtcagccctggggcagccatacaaggccatggg gctgggcaagctgcacgcctgggtccgggtgggcacgggtggcggaacgagctgaaagctc atctgctcagggggccctccctggggacagccctcctggtagtcacacctgtaggctcctctat ataaccaggggcacaggggctgcctcattctaccaccacctccacagcaccagACGCGTca gGTAAGTTTAGTCTTTTTGTCTTTTATTTTCAGGTCCCAGATCTG GTGGTGGTGCAAATCAAAGAACTGCTCCTCAGTGGCTTCTAG GTGCTAGCgccgccacc (SEQ ID NO: 138)
708 bp Desmin enhancer (full- length) + Desmin proximal (full- length) + SIE + Kozak 1	ACCCATGCCTCCTCAGGTACCCCTGCCCCACAGCTCCTCT CCTGTGCCTTGTTCCTCAGCCATGCGTTCTCCTCTATAAATAC CCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAG ATGGTTGGGTGACATGCGGCTCCTGACAAAACACAAACCCC TGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATC AGGGAGGGGGCGGGGGACCCAGGGGGCAGGAGCCACACAA AGTCTGTGCGGGGGTGGGAGCGCACATAGCAATTGAAACT GAAAGCTTATCAGACCCCTTTCTGGAAATCAGCCCACTGTTTA TAAACTTGAGGCCCCACCCTCGAGATAACCAGGGCTGAAAG AGGCCCCGCTGGGGGCTGCAGACATGCTTGCTGCCTGCCCTG GCGAAGGATTGGCAGGCTTGCCCGTCACAGGACCCCCGCTGG CTGACTCAGGGGCGCAGGCCTCTTGCGGGGAGCTGGCCTCC CCGCCCCACGGCCACGGGCCGCCCTTTCTGGCAGGACAGC GGGATCTTGACAGCTGTGAGGGAGGGGAGGCGGGGGCTGAT GTCAGGAGGGATATAAATAGTGCCGACGGCTGGGGGCCCTGT CTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCCCGC CGCCTCCTCCGTGCGCCCGCCAGCCTCGCCCGCGCCGTCACG CGTcaggttaagtggtgtaaaacaagtgtaattgagtgatagagtactgtctgccgcagattcta aaaataaccacctggacagcagcaccgaaggtcttggcggtccctgttctcttccctggcttcag ctcaccgtgggtccattcataggtatttagtatgctgtatatattttcaggtgggatccACGCGTgc cgccacc (SEQ ID NO: 23)
448 bp Desmin enhancer (truncated) + Desmin proximal (truncated) + SIE + Kozak 1	CCCACAGCTCCTCTCCTGTGCCTTGTTTCCCAGCCATGCGTTC TCCTCTATAAATACCCGCTCTGGTATTTGGGGTTGGCAGCTGT TGCTGCCAGGGAGATGGTTGGGTGACATGCGGCTCCTGACA AAACACAAACCCCTGGTGTGTGTGGGCGTGGGTGGTGTGAGT AGGGGGATGAATCAGGGAGGGGGCGGGGGACCCAGGGAGG GGCGCAGGCCTCTTGCGGGGGAGCTGGCCTCCCCGCCCCAC GGCCACGGGCCGCCCTTTCTGGCAGGACAGCGGGATCTTGC AGCTGTGAGGGGAGGGGAGGCGGGGGCTGATGTCAGGAGGG ATAAATAAGTGCCGACGGCTGGGGGCCCTGTCTCCCCCTCGCC GCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCCTCCTCCG TGCGCCCCGACGCTCGCCCGCGCCGTCACGCGTcaggttaagtggc tgtaacaagaagtgtaattgagtgatagagtactgtctgccgcagattctaaaataaccacctgga cagcagcaccgaaggtcttggcggtccctgttctcttccctggcttcagctcaccgtgggtccattc ataggtatttagtatgctgtatatattttcaggtgggatccACGCGTgccgccacc (SEQ ID NO: 24)
111 bp MYLPF + 448 bp Desmin	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGC

(truncated) + SIE + Kozak 1	AGCTGCTAAGTTTAGGGTGGCTCCTTCTCTAGTCCCACAGCTC CTCTCCTGTGCCTTGTTCCTCCAGCCATGCGTTCTCCTCTATAA ATACCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAG GGAGATGGTTGGGTTGACATGCGGCTCCTGACAAAACACAA ACCCCTGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGAT GAATCAGGGAGGGGGCGGGGGACCCAGGGAGGGGCGCAGG CCTCTTGCAGGGGAGCTGGCCTCCCCGCCCCACGGCCACGG GCCGCCCTTTCCTGGCAGGACAGCGGGATCTTGCAGCTGTCA GGGGAGGGGAGGCGGGGGCTGATGTCAGGAGGGGATAAAT AGTGCCGACGGCTGGGGGGCCCTGTCTCCCTCGCCGCATCCA CTCTCCGGCCGGCCGCCTGCCCGCCGCCTCCTCCGTGCGCCC GCCAGCCTCGCCCGCGCCGTCACGCGTcaggttaagttggctgtaacaaagt tgaatttgagttgatagagtactgtctgccgccagattctaaaaataaccacctggacagcagcacc aaggtctttggcgggtccctgttctcttccctggcttcagctcaccgtgggtccattcataggtatttagt atgctgtatatattttcaggtgggatccACGCGTgccgccacc (SEQ ID NO: 25)
51 bp MYLPF + 448 Desmin (truncated) + SIE + Kozak 1	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGCTG CTGTCTGAGCTAGTCCCACAGCTCCTCTCCTGTGCCTTGTTC CCAGCCATGCGTTCTCCTCTATAAATACCCGCTCTGGTATTTG GGGTTGGCAGCTGTTGCTGCCAGGGAGATGGTTGGGTTGACA TGCGGCTCCTGACAAAACACAAACCCCTGGTGTGTGTGGGCG TGGGTGGTGTGAGTAGGGGGATGAATCAGGGAGGGGGCGGG GGACCCAGGGAGGGGCGCAGGCCTCTTGCAGGGGAGCTGGC CTCCCCGCCCCACGGCCACGGGCCGCCCTTTCCTGGCAGGA CAGCGGGATCTTGCAGCTGTCAGGGGAGGGGAGGCGGGGGC TGATGTCAGGAGGGGATAAATAGTGCCGACGGCTGGGGGCC CTGTCTCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCC CGCCGCCTCCTCCGTGCGCCCCGCCAGCCTCGCCCGCGCCGTC ACGCGTcaggttaagttggctgtaacaaagtgaatttgagttgatagagtactgtctgccgccag attctaaaaataaccacctggacagcagcaccgaaggtctttggcgggtccctgttcttccctgg cttcagctcaccgtgggtccattcataggtatttagtctgtatatattttcaggtgggatccACGC GTgccgccacc (SEQ ID NO: 26)
60 bp MYLPF + 448 bp Desmin (truncated) + SIE + Kozak 1	GCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCTAA GTTTAGGGTGGCTCCTTCTCTAGTCCCACAGCTCCTCTCCTGT GCCTTGTTCCTCCAGCCATGCGTTCTCCTCTATAAATACCCGCT CTGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGATGGT TGGGTTGACATGCGGCTCCTGACAAAACACAAACCCCTGGTGT TGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATCAGGG AGGGGGCGGGGGACCCAGGGAGGGGCGCAGGCCTCTTGCAG GGGAGCTGGCCTCCCCGCCCCACGGCCACGGGCCGCCCTTT CCTGGCAGGACAGCGGGATCTTGCAGCTGTCAGGGGAGGGG AGGCGGGGGCTGATGTCAGGAGGGGATAAATAGTGCCGACG GCTGGGGGGCCCTGTCTCCCTCGCCGCATCCACTCTCCGGCC GGCCGCCTGCCCGCCGCCTCCTCCGTGCGCCCCGCCAGCCTCG CCCGCGCCGTCACGCGTcaggttaagttggctgtaacaaagtgaatttgagttgatag agtactgtctgccgccagattctaaaaataaccacctggacagcagcaccgaaggtctttggcgggt ccctgttcttccctggcttcagctcaccgtgggtccattcataggtatttagtctgtatatattttc aggtgggatccACGCGTgccgccacc (SEQ ID NO: 27)
CK8e + SIE + Kozak 2	TGGATCTCTGCAGCCCTCCTTAGGTGCCCATGTAAGGAGGCA AGGCCTGGGGACACCCGAGATGCCTGGTTATAATTAACCCAG

	ACATGTGGCTGCCCCCCCCCCCCAACACCTGCTGCCTCTAAA AATAACCCCTGCATGCCATGTTCCCGGCGAAGGGCCAGCTGTC CCCCGCCAGCTAGACTCAGCACTTAGTTTAGGAACCAGTGAG CAAGTCAGCCCTTGGGGCAGCCCATAACAAGGCCATGGGGCT GGGCAAGCTGCACGCCTGGGTCCGGGGTGGGCACGGTGGCC GGGCAACGAGCTGAAAGCTCATCTGCTCTCAGGGGCCCCCTCC CTGGGGACAGCCCCTCCTGGCTAGTCACACCCTGTAGGCTCC TCTATATAACCCAGGGGCACAGGGGCTGCCCTCATTCTACCA CCACCTCCACAGCACAGACAGACACTCAGGAGCCAGCCAGC ACGCGTcaggttaagttggctgtaacaaagttgaatttgagttgatagagtactgtctgccgccag attctaaaaataaccacctggacagcagcacccaaggtcttggcgggtccctgttctcttccctgg cttcagctcaccgtggtccattcataggtatttagtatgctgtatatattttcaggtggCgccgccAC Catgg (SEQ ID NO: 28)
1 kb Desmin + SIE + Kozak 2	tacccttgcacacagctcctctgtgcttgggttccagccatgcgttctctataaaatacc gctctggtatttgggttggcagctgttgcctgagggagatgggtgggtgacatgcggtcctgaca aaacacaaaccctggtgtgtgtggcggtgggtgtgagtaggggatgaatcagggaggggg cAggggacccagggggcaggagccacacaaagtctgtgcgggggtgggagcgacatagcaat tggaaactgaaagcttatcagacccttctggaatcagccactgttataaactgaggccccacct cgacagtaccggggagggaagaggcctgcactagtcagagggaactgaggtcagggcCag ctgccccatagacatacatggcaggcaggcttggccaggatccctccgctgcccaggcgtctccct gccctcccttctgcctagagacccccacctcaagcctggctggtcttgcctgagacccaaacctt tcgacttcaagagaatatttaggaacaaggtggttagggccttctgggaacaggccttgaccttta agaaatgacccaaagtctctccttgacaaaaaggggaccctcaactaaagggaagcctcttctg ctgtctcccctgacccactccccccacccaggacgaggagataaccagggtgaaaggaggcc cgcttgggggctgcagacatgcttgcctgcccggcgaaggattggcaggcttgcctgacag gaccccgctggctgactcagggggcgcaggcctTttcgggggagctggcctccccgccccac ggccacgggccccttcttgcaggacagcgggatcttgcagctgtcaggggaggggagggcg ggggctgatgtcaggaggatataatagtgccgacggctgggggcctGTCTCCCCCTCG CCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCCTCCTC CGTGCGCCCGCCAGCCTCGCCCGGCAGgtaagttggctgtaacaaagttga atttgagttgatagagtactgtctgccgcagattctaaaaataaccacctggacagcagcacccaag gtcttggcgggtccctgttctcttccctggcttcagctcaccgtggtccattcataggtatttagtatg ctgtatatattttcaggtCTAGCgccgccACCatgg (SEQ ID NO: 29)

[0090] Myosin light chain phosphorylatable fast skeletal muscle (MYLPF) is a regulatory component/enhancer that is derived from the human muscle-specific gene, *MYLPF*, encoding the myosin light chain. *MYLPF* is typically highly expressed in fast twitch muscle while having lower expression in low or slow twitch muscle. It is one of numerous potent muscle-specific transcriptional *cis*-regulatory modules (CRMs) containing clusters of transcription factor binding sites. It is one of seven different skeletal muscle-specific *cis*-regulatory modules designated as *Sk-CRMs* associated with genes that are highly expressed in skeletal muscle. As such, *MYLPF* is also known as *Sk-CRM4* (Sarcar et al., Nat. Commun. 10, 492(2019)). The MYLPF gene corresponds to the corresponds to the following position in the human genome (hg19), chr16:30,383,321-30,383,755. In some embodiments, a nucleic acid

sequence encoding human MYLPF comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

```

ACTCCTTGC TTCTTTCAG CCGGAGCCGC TGCCTTGCC CCCGGAGACT GAAGACATGg      60
tgagtggagaa tcttccaacc cctgcccaga tcccttagac ctcagggcag cctcaccctt      120
tgaaagaaag tagctggttc ccagtccaac aaaaagttaa aaggatgctc atctctttct      180
cgccatcact gatggaaaac agactgtgtg atggcctaaa ggatccaatc aatgataatc      240
cgccagacaca ttcaactcac agggggccag tccctctcaa agcctgatcc attcattcct      300
gtactccatc ttctaaggc cctctggcac cagccaggca ataactccca tcccatttta      360
caaatgagaa agtagaggct caggccaggt gctgtggttc acgctgtaa tcccagcact      420
ttggggaggcc aaggtggttg gatcacttga agccaggagt tcgagaccag cctggccaac      480
aggcgaaacc ccgtctctac taaaaataca aaaattaaact gggcggtgtg gtgcattgct      540
gtaatcccag ctactacgga ggctggggca taagaatcgc ttgaaccccg gaggcagagg      600
ttgcggttaag ccgagatcgc agcaccgcac tccagcctgg gtgacagagc aagactctgt      660
ctcaaaaaaa aaaaaaaaaa gaaaaagaaa gtagaggctc aaacagggca aatgagtggc      720
tgggggtacag gtggagggac tgggtccatgg gccctcttgt tcaacccctca cctccagGC      780
gggaagcttcc acccattgtg ggggatggag agttggagac ccaggagcca cagacgggaa      840
tttggggaat gcaatcaaa gcaagccctc cagcaaaagag gtcccaagga aggggagact      900
caggccaaaag ggctgcttta ggaacaggga ctccaagaat tctgactctt aatccattgc      960
ccccagagag gtggagggac tgggtccatgg gccctcttgt tcaacccctca cctccagGC      1020
ACCCAAGAGG GCCAAGAGAA GGACAGTAGA GGGCGGAAGC TCCAGCGTCT TCTCCATGTT      1080
CGACCAGACT CAGATCCAGG AGTTCAAAGA Ggtgggtgag gggacgggga aactggaagg      1140
ctgacccaaa gcagccctgc tggccctctc cctggaatgt gcacctgctt gaaggagggg      1200
aaagggtttag aattccttcc tcccctctct gcttgggtgt gaagaggaca gttggctgag      1260
agccagcatc tgatcctcct ggggtaggga tgetgaggct gggccatggg ggacctaac      1320
ctccccacca cgccctccc cagGCCCTCA CTGTGATCGA CCAGAACCGT GATGGTATTA      1380
TAGACAAGGA GGACCTTCGG GACACCTTCG CAGCCATGGg tgagcccccct acccccaggt      1440
gggaatattc aagctgcag agtctaggaa attcagtggc cttgggttcc agccctgtca      1500
gctccaccctt gggcctcagt ctaccagct gaaaaatgaa tgggatcagc tgaattagt      1560
ggaaagcgag gcccgggccc aggagcctgc tccagcccat gcttccccc accccctcca      1620
gGCCGCCTCA ATGTGAAGAA TGAGGAGTTG GATGCCATGA TGAAGGAAGC CAGCGGTCCC      1680
ATCAACTTCA CCGTCTTCT GACCATGTTG GGGGAGAAGC TCAAGGgtga gttgagtgc      1740
ctgggtccaa gggcccggt gccttttccc agacccttca gggaccctc agcttcatgt      1800
gccctctgac cccacagGT GCCGACCTG AGGATGTGAT CACCGGAGCC TTCAAGGTCT      1860
TGGACCTTGA GGGAAAGGGC ACCATCAAGA AGAAGTTgta agcatcggt cccccacct      1920
tccgaccctt cccccactcc accagctgct cccacactga cagcctcctt aaatttctac      1980
caaggctggg cctgtggct caggcctata atcccagtgc tttgggaggc agaggcagga      2040
ggatttactt gaggacagga gtttgagatc agcctggata acacagcatg accctgttcc      2100
taaaaaactta actaaaaaat tcgtgggctg gccgggcaca gtggctcatg cctgtaatcc      2160
cagcaacttt ggagggcaa gcgggcggt catgtgaggt caggagtcca agaacagcct      2220
ggccaacacg gtgaaacccc gtctctacta aaaatacaaa aattagctgg gcgtggtggc      2280
ggggcgctgt agtcccagct actagggaga ctgggacatg agaatcgctt gaaccctgga      2340
gtcggagggtt gcagtgaacc aagatcacgc cactgccctc caacctgggc gacaatagca      2400
aaactccgtc tccaaaaaaa agaaaaaaa aaattcgcgg gagtgggtgc aaacacctat      2460
agtcccagct actccggagg ctgaggcagg aagatccctt gagcccaaga gttccaggtt      2520
gcagtgaacc gtgatcaaac cactgcattc cagcctgggc gacagggtga gacctgtct      2580
caaaacaaac aaaaacccaa aaattcctac gaagggttag gcactgccag aggtaaagga      2640
gtagttaaga ttgtgtaaat gggcggttaag attgcatttt cacagcaaat gtggcgagg      2700
agggcaaaag cgagtgtctg ggaggccaat gaaggaaagta gagtgaacctg ggttttcagg      2760
ggaggctggg gctggcatcg gggatgaggt gcgagggagt ggaaggaac caggaaacca      2820
atcgcaactc ccttgtgtc acctctctcc agCCTGGAGG AGCTGCTGAC CAGCAGTGT      2880
GACCGCTTCT CCCAGGAGGA Ggtgagtggg gaccggggcg gggccggagc agcaaaggg      2940
ctggggcccg ggagactaag ggcctgtgcg agacgcccct acgtctttcc ccagATCAAG      3000
AACATGTGGG CGGCCCTCCC CCCCAGCTG GCGCGCAACG TCGACTACAA AAACATCTGC      3060
TACGTCATCA CGCACGGCGA CGCCAAGGAC CAGGAGTAGG GGCACCCGCG GGCCTCCGT      3120
GCCCGACGCT TCTGTTCCGC CCGACCTCCA CCCCggctcc CAATAAAATT TAACTGATCT      3180
TTGTTTCTTA (SEQ ID NO: 31).

```

[0091] In some embodiments, an MYLPF enhancer comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 1. In one

embodiment, a MYLPF enhancer comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 2. In one embodiment, a MYLPF enhancer comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 3. In one embodiment, a MYLPF enhancer comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 4.

[0092] Hybrid regulatory sequences of the present disclosure can comprise at least one CK regulatory sequence. In some embodiments, the at least one CK regulatory sequence comprises at least one CK6 enhancer/promoter sequence, at least one CK1 enhancer/promoter sequence, at least one CK7 enhancer/promoter sequence, at least one CK8e promoter sequence, or any combination thereof.

[0093] Hybrid regulatory sequences of the present disclosure can comprise at least one CK6 enhancer/promoter sequence. In some embodiments, a CK6 enhancer/promoter sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 8.

[0094] Hybrid regulatory sequences of the present disclosure can comprise at least one CK1 enhancer/promoter sequence. In some embodiments, a CK1 enhancer/promoter sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 9.

[0095] Hybrid regulatory sequences of the present disclosure can comprise at least one CK7 enhancer/promoter sequence. In some embodiments, a CK7 enhancer/promoter sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 10.

[0096] Hybrid regulatory sequences of the present disclosure can comprise at least one CK8e promoter sequence. In some embodiments, a CK8e promoter sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%,

85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 17.

[0097] In some embodiments The CK6 regulatory component is a mutated enhancer region plus a CK6 promoter region. The CK7 regulatory component is a CK7 enhancer and a CK7 promoter comprising an additional 3' sequence for cardiac expression. The regulatory cassettes of murine muscle creatine kinase (CK) and alpha-myosin heavy-chain genes promote varying levels of expression in skeletal and cardiac muscle (Salva et al., Mol. Ther. 15, 2 (2007)). In one embodiment, the hybrid regulatory sequence comprises a MYLPF enhancer sequence and a CK6 enhancer-promoter sequence.

[0098] In one embodiment, the hybrid regulatory sequence comprise a MYLPF enhancer, a CK enhancer/promoter sequence, and an intronic sequence. In another embodiment, the hybrid regulatory sequence comprises a MYLPF enhancer sequence, a CK6 enhancer-promoter sequence, and an SV40 intron sequence.

[0099] In one embodiment, a hybrid regulatory sequence of the disclosure comprises a MYLPF enhancer sequence and a creatine kinase (CK) enhancer/promoter sequence.

[0100] Hybrid regulatory sequences comprising a truncated desmin promoter combined with a SIE (small intronic enhancer) are disclosed herein. The truncated desmin promoter is derived from the human sequence by deleting poorly conserved regions. The SIE was derived from primarily the first intron of the mouse muscle creatine kinase gene Piekarczyk et al. Mol Ther Methods Clin Dev. 15 (2019). In some embodiments, the hybrid regulatory sequence comprises a MYLPF enhancer sequence, a truncated desmin promoter sequence, and a SIE sequence.

[0101] Enhancers, core promoters, promoter/enhancers and intronic sequences can be modularly assembled to form hybrid regulatory sequence and can be utilized in a rAAV genome or other vector-based approach to achieve various levels and balances of expression of a transgene from cardiac and skeletal muscle. Specific combination of this collection of regulatory sequences enables desired tissue selective expression levels for a given indication, transgene, AAV capsid, and route of delivery thereby providing a more potent and efficacious therapeutic effect.

[0102] In some embodiments, the muscle selective hybrid regulatory sequence is cardiac muscle selective. In one embodiment, the muscle selective hybrid regulatory sequence is skeletal muscle selective.

[0103] Preferential selectivity of skeletal muscle expression versus cardiac muscle expression or vice versa utilizing the combination of hybrid regulatory sequences disclosed herein is represented in the below expression levels table:

				Protein Expression	Protein Expression
Enhancer	Promoter	Intron	Exemplary Construct	Cardiac	Skeletal
	DES		A02239	**	*
MYLPF	modified DES	SIE	A04959	*****	***
MYLPF	CK6	SV40	A05288	*	***
MYLPF	CK7	SV40		****	***
MYLPF	CK1	SV40		**/**	***

[0104] In some embodiments, a hybrid regulatory sequence can comprise at least one hMYLPF enhancer sequence, at least one CK6 enhancer/promoter sequence, at least one SV40 intron sequence, and at least one Kozak sequence. In some embodiments, a hybrid regulatory sequence can comprise, from 5' to 3', at least one hMYLPF enhancer sequence, at least one CK6 enhancer/promoter sequence, at least one SV40 intron sequence, and at least one Kozak sequence.

[0105] In some embodiments, a hybrid regulatory sequence of the disclosure comprises a MYLPF enhancer sequence. In some embodiments, a skeletal muscle selective hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK6 (muscle creatine kinase 6) enhancer-promoter sequence set forth in SEQ ID NO: 8, an SV40 intron set forth in SEQ ID NO: 7, and a kozak sequence set forth in SEQ ID NO: 11. In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 136.

[0106] In some embodiments, a skeletal muscle selective hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK1 (muscle creatine kinase 1) enhancer-promoter sequence set forth in SEQ ID NO: 9, an SV40 intron set forth in SEQ ID NO: 7, and a kozak sequence set forth in SEQ ID NO: 11. In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 137.

[0107] In some embodiments, a skeletal muscle selective hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK7 (muscle creatine kinase 7) enhancer-promoter sequence set forth in SEQ ID NO: 10, an SV40 intron set forth in SEQ ID NO: 7, and a kozak sequence set forth in SEQ ID NO: 11. In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 138.

[0108] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 19.

[0109] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 20.

[0110] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 21.

[0111] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 23.

[0112] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 24.

[0113] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 25.

[0114] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 26.

[0115] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 27.

[0116] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 28.

[0117] In some embodiments, a hybrid regulatory sequence comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 29.

[0118] *PUF(CUG) Compositions*

[0119] RNA-targeting gene therapy compositions and methods for treating myotonic dystrophy type 1 (DM1) are provided here. Such compositions comprise hybrid regulatory promoter-enhancer-intron combinations, hybrid regulatory sequences of the disclosure, which drive selective and preferential expression of transgenes, including CUG-repeat targeting PUF polypeptides in skeletal muscle cells over cardiac muscle cells.

[0120] The disclosure provides a recombinant AAV (rAAV) vector comprising a hybrid regulatory sequence of the disclosure operably linked to a transgene. The disclosure also provides an rAAV vector comprising a hybrid regulatory sequence of the disclosure operably linked to a CUG-repeat RNA targeting PUF polypeptide.

[0121] The disclosure provides an rAAV vector comprising a first inverted terminal repeat (ITR) sequence; a hybrid regulatory sequence of the disclosure; a nucleic acid sequence encoding a transgene; and a second ITR sequence.

[0122] The disclosure provides an rAAV vector comprising a first inverted terminal repeat (ITR) sequence; a hybrid regulatory sequence of the disclosure; a nucleic acid sequence encoding a PUF polypeptide; a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence; an SV40 polyA sequence; and a second ITR sequence.

[0123] The disclosure provides an rAAV vector comprising a first inverted terminal repeat (ITR) sequence; an MYLPF enhancer sequence; a CK6 promoter sequence; an SV40 intron sequence; a nucleic acid sequence encoding a PUF polypeptide; a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence; an SV40 polyA sequence; and a second ITR sequence.

[0124] In some embodiments, the first ITR sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 32.

[0125] In some embodiments, the second ITR sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 34.

[0126] In some embodiments, the MYLPF enhancer sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 1.

[0127] In some embodiments, the CK6 promoter sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 8.

[0128] In some embodiments, the SV40 intron sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 35.

[0129] In some embodiments, the PUF polypeptide sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 45.

[0130] In some embodiments, the WPRE sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 38.

[0131] In some embodiments, the SV40 polyA sequence comprises a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the nucleic acid sequence set forth in SEQ ID NO: 39.

[0132] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 2. The vector, A05288 comprises from 5' to 3' the elements set forth in Table 2.

Table 2: ssAAV MYLPF_CK6_SV40intron_PUFCUG_WPRE_SV40pA A05288 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequence
5' ITR	cctgcaggcagctgcgcgctcgtcgtcactgaggccgccccgggcgtcggcgga cctttggtcggcgccctcagtgagcgcgagcgcgcagagagggagtggccaa ctccatcactaggggttct (SEQ ID NO: 32)
MYLPF enhancer (CMR4)	TTCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAG CCCCATGTCTGTCAATTCCCACTCAGTGTCTGATC TCCTTCTCCTCACCTTTCCCATCTCCCGTTTGACCCA

	GCTTCCTGAGCTCTCCTCCCATTCCTCTTTTGGAGT CCTCCTCCTCTCCCAGAACCCAGTAATAAGTGGGCT CCTCCCTGGCCTGGACCCCGTGGTAACCCTATAAG GCGAGGCAGCTGCTGTCTGAGGCAGGGAGGGGCTG GTGTGGGAGGCTAAGGGCAGCTGCTAAGTTTAGGG TGGCTCCTTCTCTCTTCTTAGAGACAACAGGTGGCT GGGGCCTCAGTGCCCAGAAAAGAAAATGTCTTAGA GGTATCGGCATGGGCCTGGAGGAGGGGGGACAGGG CAGGGGGAGGCATCTTCCTCAGGACATCGGGTCCT AGAGG (SEQ ID NO: 1)
CK6 (creatine kinase)	ccactacgggtctaggctgcccatgtaaggaggcaaggcctggggacacccgagat gcctgggtataattaacccaacacctgctgcccccccccaacacctgctgcctg agcctgagcgggttaccacacccgggtgcctgggtcttaggctgtacaccatggag gagaagctcgtctaaaaataacacctgtccctgggtggatGGGCCCaaggctgt gggggactgagggcaggctgtaacaggcttggggccagggtctatacgtgcctg ggactcccaagttactgttccatgttcccggcgaagggccagctgtccccgcca gctagactcagcacttagtttaggaaccagtgagcaagtcagcccttggggcagccc atacaaggccatggggctgggcaagctgcacgcctgggtccgggtgggacaggt gccccgggcaacgagctgaaagctcatctgctctcagggggccccctggggacag ccccctcgtggctagtcacacccctgtaggctcctctatataacccaggggcacagggg ctgcccCCGGGGTCACCACC (SEQ ID NO: 8)
SV40 intron	cagGTAAGTTTAGTCTTTTTGTCTTTTATTTTCAGGTCC CAGATCTGGTGGTGGTGCAAATCAAAGAACTGCTC CTCAGTGGCTTCTAGGT (SEQ ID NO: 35)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	gaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaccactgaaaa actacgc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-CUG	GGACGAAGCCGACTCTTGGAAGACTTCAGAAACAA TCGGTATCCGAACCTTCAGCTGAGAGAAATTGCTGG TCACATCATGGAATTTTCTCAAGATCAACATGGAAG CCGGTTTATTGAACTTAAACTCGAACGAGCCACCCC GGCCGAAAGGCAATTGGTGTTCATGAAATTCTTCA GGCCGCATACCAACTCATGGTTGATGTTTTTGGGAA CTATGTTATTCAAAAGTTTTTTGAGTTCGGGTCCT GGAGCAAAAGTTGGCATTGGCAGAGCGAATCCGGG GCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTA GTTATGTCATTTCGAAAGCACTCGAGTTCATACCAT CAGATCAACAGAATGAGATGGTGCGGGAGCTGGAT GGGCATGTTTTGAAATGCGTGAAAGACCAAAACGG TAGCTACGTAGTTGAGAAATGCATCGAATGCGTCC AACCACAGTCTCTCCAATTTATTATAGATGCATTTA AGGGTCAGGTTTTTCGCGCTTTCTACGCACCCGTATG GGAACCGAGTGATTCAGAGAATCTTGAGCACTGC CTGCCGGATCAGACACTCCCTATCTTGAGGAATTG CACCAGCATACCGAACAATTGGTGCAAGATCAATA CGGTTTCATATGTTATTCGGCACGTTCTTGAGCATGG AAGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGA TTAGAGGTAACGTATTGGTGCTCTCACAAACAAAT

	TTGCATCTAATGTGGTGGAGAAATGTGTTACTCATG CTTCTAGAACGGAAAGGGCAGTTCTCATAGACGAA GTTTGCACAATGAATGATGGTCCTCATAGCGCACTT TATACCATGATGAAGGACCAGTATGCAAACATATGT CGTCCAGAAAATGATCGATGTGGCGGAGCCCCGGTC AACGGAAAATCGTGATGCACAAAATCCGACCTCAC ATTGCTACACTCAGAAAATACACGTATGGAAAACA TATTCTGGCTAAGCTGGAGAAATATTACATGAAGA ATGGAGTGGATCTGGGGTaa (SEQ ID NO: 45)
WPRE	aatcaacctctggattacaaaatttgtgaaagattgactggtattcttaactatgttgctcc ttttacgctatgtggatagctgctttaatgcctttgtatcatgctattgctcccgatggc tttcattttctcctctgtataaatcctggttgctgtctctttatgaggagttgtggcccggtt gtcaggcaacctggcggtgtgactgtgtttgctgacgaacccccactgggtggg gcattgccaccactgtcagctcctttccgggactttcgctttccccctccctattgccac ggcggaactcatcgccgctgccttgcctgctgacaggggctcgctgttgg gcactgacaattccgtgtgtgttcggggaatcatcgcttcttctggctgctgcct gtgttgccacctggattctgcgaggacgtccttctgctacgtccctcgccctcaat ccagcggaccttctcccgcgctgctgcccgtctgcccctcttccgctcttc gccttcgccctcagacgagtcggatctcccttgggcccgcctccccgc (SEQ ID NO: 38)
SV40-PA	aactgtttattgcagcttataatggttacaataaagcaatagcatcacaatttcacaa ataaagcattttttcactgcattctagttgtggtttgtccaaactcatcaatgtatctta (S EQ ID NO: 39)
3' ITR	aggaacccctagtgtgaggtggccactccctctctgcgcgctcgtcgtcactga ggccggcgaccaaaggctcggcgacggcggttggccggcgccctcagt gagcgagcgagcgcgagctgcctgcagg (SEQ ID NO: 34)

[0133] In some embodiments, rAAV vector A05288 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical:

cctgcaggcagctgcgcgctcgtcgtcactgaggccgccccggcgctggggcagccttggctgccccggcctcagtgagcgagc
gagcgcgagagaggagtgccaactccatcactaggggttctGCGGCCGGTCGCGTCTAGTACTAGTT
TCTGAGTCCTCTAAGGTCCCTCACTCCCAACTCAGCCCCATGTCCTGTCAATTCCC
ACTCAGTGTCTGATCTCCTTCTCCTACCTTTCCCATCTCCCGTTTGACCCAGCTT
CCTGAGCTCTCCTCCCATTTCCCTTTTTGGAGTCCTCCTCCTCTCCCAGAACCCAG
TAATAAGTGGGCTCCTCCCTGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGG
CAGCTGCTGTCTGAGGCAGGGAGGGGCTGGTGTGGGAGGCTAAGGGCAGCTGCT
AAGTTTAGGGTGGCTCCTTCTCTCTTCTTAGAGACAACAGGTGGCTGGGGCCTCA
GTGCCAGAAAAGAAAATGTCTTAGAGGTATCGGCATGGGCCTGGAGGAGGGGG
GACAGGGCAGGGGGAGGCATCTTCTCAGGACATCGGGTCCTAGAGGccactacgggt
ctaggctgccccatgtaaggaggcaaggcctggggacacccgagatgcctggttataaftaaccacaacacctgctgcccccccccc
caacacctgctgcctgagcctgagcggttaccacccccggctggtggtcttaggctctgtacacatggaggagaagctcgtctta

aaaataaccctgtccctggtgatGGCCCCaaggctgtgggggactgagggcaggctgtaacaggctgggggccagggtt
atacgtgcctgggactcccaaagtattactgttccatgttccggcgaaaggccagctgtccccgccagctagactcagcacttagtt
aggaaccagtgtgagcaagtcagccctggggcagccatacaaggccatggggctgggcaagctgcacgcctgggtccgggtgg
gcacgggtccccgggcaacgagctgaaagctcatctgctctcaggggcccctccctggggacagcccctcctggctagtcacaccct
gtaggctcctctatataaccaggggcacaggggtgcccCCGGGTCACCACCACGCGTcagGTAAGTTT
AGTCTTTTTGTCTTTATTTTCAGGTCCCAGATCTGGTGGTGGTGCAAATCAAAGA
ACTGCTCCTCAGTGGCTTCTAGGTGCTAGCgccgccACCatggaccgtgtgctcaaagaagtgtga
aggaagcaaccctcctaaccactgaaaaactacgcggaGGATCGGGACGAAGCCGACTCTTGGAAG
ACTTCAGAAACAATCGGTATCCGAACCTTCAGCTGAGAGAAATTGCTGGTACAT
CATGGAATTTTCTCAAGATCAACATGGAAGCCGGTTTATTGAACTTAAACTCGAA
CGAGCCACCCCGGCCGAAAGGCAATTGGTGTTCATGAAATTCTTCAGGCCGCA
TACCAACTCATGGTTGATGTTTTTGGGAACTATGTTATTCAAAAGTTTTTTGAGTT
CGGGTCACTGGAGCAAAAGTTGGCATTGGCAGAGCGAATCCGGGGCCATGTTCT
GAGCCTCGCTCTCCAAATGTACGGTAGTTATGTCATTTCGCAAAGCACTCGAGTTC
ATACCATCAGATCAACAGAATGAGATGGTGCAGGAGCTGGATGGGCATGTTTTG
AAATGCGTGAAAGACCAAAACGGTAGCTACGTAGTTGAGAAATGCATCGAATGC
GTCCAACCACAGTCTCTCCAATTTATTATAGATGCATTTAAGGGTCAGGTTTTCG
CGCTTTCTACGCACCCGTATGGGAACCGAGTGATTTCAGAGAATCTTGAGCACTG
CCTGCCGGATCAGACACTCCCTATCTTGGAGGAATTGCACCAGCATACCGAACA
ATTGGTGCAAGATCAATACGGTTCATATGTTATTCGGCACGTTCTTGAGCATGGA
AGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGATTAGAGGTAACGTATTGGTG
CTCTCACAACACAAATTTGCATCTAATGTGGTGGAGAAATGTGTTACTCATGCTT
CTAGAACGGAAAGGGCAGTTCTCATAGACGAAGTTTGCACAATGAATGATGGTC
CTCATAGCGCACTTTATACCATGATGAAGGACCAGTATGCAAACATATGTCGTCCA
GAAAATGATCGATGTGGCGGAGCCCGGTCAACGGAAAATCGTGATGCACAAAAT
CCGACCTCACATTGCTACACTCAGAAAATACACGTATGGAAAACATATTCTGGCT
AAGCTGGAGAAATATTACATGAAGAATGGAGTGGATCTGGGGtaatcGCGGCCGCT
CGAGGATTATAAGGATGACGACGATAAATTCGTCGAGCACCACCACCACCA
CTAATAAGGTTTATCCGATCCACCGGATCTAGATAAGATATCCGATCCACCGGAT
CTAGATAACTGATCATAATCAGCCATAACCACATTTGTAGAGGTTTTACTTGCTTT
AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTT
GTTGTTaatcaacctctggattacaaaattgtgaaagattgactgtattcttaactatgttgccttttacgctatgtggatagctg
ctttaatgcctttgtatcatgctattgcttcccgatggctttcattttctctccttgataaatcctgggtgctgtctctttatgaggagttgtgg

cccgtgtcaggcaacgtggcgtggtgtgactgtgttctgacgaacccccactggtggggcattgccaccacctgtcagctcct
 ttccgggaactttcgtttccccctccctattgccacggcggaactcatcgccgctgccttccccgctgctggacaggggctcggtgtt
 gggcactgacaattccgtggtgtgtcggggaaatcatcgctccttcttggctgctcgctgtgtgccacctggattctgcgcgggac
 gtcttctgctacgtcccttggccctcaatccagcggaccttccctcccgccgctgctgccgctctgcggcctcttccgcttctcg
 ccttcgcccctcagacgagtcggatctcccttgggcccgcctccccgcaactgtttattgcagcttataatggtfacaataaagcaatag
 catcacaatttcacaaataaagcattttttactgcattctagtgtggtttgtccaaactcatcaatgtatcttaACGCGGTAAC
 CACGTGCGGACCCAACGGCCGCaggaaccctagtgtgaggtggccactccctctctgcgcgctcgctcgctc
 actgaggccgggaccaaaggctgccccacgcccgggctttgcccggcgccctcagtgagcgagcgagcgcgagctgcct
 gcagg (SEQ ID NO: 47).

[0134] In some embodiments, rAAV vector A05288 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 139.

[0135] In some embodiments, an rAAV vector of the present disclosure can comprise, consist essentially of, or consist of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 150.

[0136] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 3. The vector, A02239 comprises from 5' to 3' the elements set forth in Table 3.

Table 3: ssAAV hDesmin PUF CUG WPRE SV40pA A02239 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequences
5' ITR	Cctgcaggcagctgcgcgctcgctcactgaggccgcccggcgctgggacaccttgggt cgccccggcctcagtgagcgagcgagcgcgagagggagtgccaactccatcactagggg ttct (SEQ ID NO: 32)
Human Desmin promoter	tacccccctgccccccacagctcctctcctgtgccttgttccagccatgcgttctcctataaatac ccgctctggtatttggggtggcagctgttctgctgccaggagatggttgggtgacatgcggctcct gacaaaacacaaacccctggtgtgtgtggcggtgggtggtgtgagtaggggatgaatcaggga gggggcAggggacccagggggcaggagccacacaaagtctgtcggggggtgggagcgcac atagcaattggaaactgaaagcttatcagacccttctggaaatcagccactgtttataaactgag gccccaccctcgacagtaccggggaggaagaggcctgcactagtccagagggaactgagg ctcagggcCagctcgccatagacatacatggcaggcaggcttggccaggatccctccgctg ccaggcgtctccctgcctcccttctgctagagacccccaccctcaagcctggctggtctttgc ctgagacccaaacctcttcgactcaagagaatatttaggaacaaggtggttagggccttcttg gaacaggccttgacccttaagaaatgacccaaagtctctccttgacaaaaaggggaccctcaa actaaagggaagcctctctctgctgtctccctgacccactccccccaccacaggacgagga gataaccagggtgaaagaggcccgctgggggctgcagacatgctgtgctgcctgcctggcg aaggattggcaggcttgcccgacagggacccccgctggctgactcaggggcgagggccttttg cgggggagctggcctccccgccccacggccacgggcccgttcttggcaggacagcggg atcttgacgtgtcaggggaggggagggcggggctgatgtcaggagggatacaaatagtcgg acggctgggggcccGTCTCCCTCGCCGCATCCACTCTCCGCGCG

	GCCGCCTGCCCGCCGCCTCCTCCGTGCGCCCGCCAGCCTCG CCCG (SEQ ID NO: 16)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	Gaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaaccactgaaaaactacgc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-CUG	GGACGAAGCCGACTCTTGGAAGACTTCAGAAACAATCGGT ATCCGAACCTTCAGCTGAGAGAAATTGCTGGTCACATCATG GAATTTTCTCAAGATCAACATGGAAGCCGGTTTATTGAACT TAAACTCGAACGAGCCACCCCGGCCGAAAGGCAATTGGTG TTCAATGAAATTCTTCAGGCCGCATACCAACTCATGGTTGA TGTTTTTGGGAAGTATGTTATTCAAAGTTTTTTGAGTTTCGG GTCACTGGAGCAAAAGTTGGCATTGGCAGAGCGAATCCGG GGCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTAGTTA TGTCATTCGCAAAGCACTCGAGTTCATACCATCAGATCAAC AGAATGAGATGGTGCGGGAGCTGGATGGGCATGTTTTGAA ATGCGTGAAAGACCAAAACGGTAGCTACGTAGTTGAGAAA TGCATCGAATGCGTCCAACACAGTCTCTCCAATTTATTAT AGATGCATTTAAGGGTCAGGTTTTTCGCGCTTCTACGCACC CGTATGGGAACCGAGTGATTCAGAGAATCTTGAGCACTG CCTGCCGGATCAGACACTCCCTATCTTGGAGGAATTGCACC AGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATA TGTTATTCGGCACGTTCTTGAGCATGGAAGGCCAGAGGAC AAGTCAAAGATCGTCGCTGAGATTAGAGGTAACGTATTGG TGCTCTCACAAACACAAATTTGCATCTAATGTGGTGAGAAA TGTGTTACTCATGCTTCTAGAACGGAAAGGGCAGTTCTCAT AGACGAAGTTTGCACAATGAATGATGGTCCTCATAGCGCA CTTTATACCATGATGAAGGACCAGTATGCAAATATGTCGT CCAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGAAA ATCGTGATGCACAAAATCCGACCTCACATTGCTACACTCAG AAAATACACGTATGGAAAACATATTCTGGCTAAGCTGGAG AAATATTACATGAAGAATGGAGTGGATCTGGGGTaa (SEQ ID NO: 45)
WPRE	Aatcaacctctggattacaaaattgtgaagattgactggtattcttaactatgtgctcctttacgct atgtggatacgtgctttaatgctttgtatcatgctattgcttcccgtatggctttcattttctcctctg tataaatcctggtgctgtctctttatgaggagttgtggcccggtgtcaggcaacgtggcgtggtgtg cactgtgttgctgacgcaacccccactggtggggcattgccaccacctgtcagctcctttccggg actttcgctttccccctcctattgccacggcggaactcatcgccgctgcttggccgctgctgga caggggctcgctgttgggcactgacaattccgtggtgtgtcggggaaatcatgctccttccttg gctgctgcctgtgttgccacctggattctgcgaggacgtccttctgctacgtcccttcggccctc aatccagcggaccttcctcccgcggcctgctgcccgtctgcccgtcttccgcgttcttcgcttc gccctcagacgagtcggatctccctttgggcccctccccgc (SEQ ID NO: 38)
SV40-PA	Aacttgttattgcagcttataatggttacaaataagcaatagcatcacaatttcacaaataaagc attttttactgcattctagtgtgtgtgttgcacaaactcatcaatgtatctta (SEQ ID NO: 39)
3'ITR	Aggaaccctagtgtgagttggccactccctctctgcgcgctcgtcgtcactgaggccgg gcgaccaaaggtcgcccgacggcggttggccggcgccctcagtgagcgagcgagcgc gcagctgcctgcagg (SEQ ID NO: 34)

[0137] In some embodiments, rAAV vector A02239 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

```
cctgcaggcagctgcgcgctcgtcgtcactgaggccgccccggcgctggcgacctttggtcgccggcctcagtgagcgagc
gagcgcgcagagaggagtgccaaactccatcactaggggttctGCGGCCGGTCGCGTCTAGTACTAGTta
ccccctgccccccacagctcctcctgtgccttgtttccagccatgcgttctcctataaatacccgctctggtatttggggttggcag
ctgttctgcccaggagatggttgggttgacatgcggctcctgacaaaaacacaaacccctggtgtgtgtggcggtgggtgtgtgagt
agggggatgaatcaggagggggcAgggggacccagggggcaggagccacacaaagtctgtcggggggtgggagcgcacata
gcaattggaaactgaaagcttatcagaccctttctggaaatcagcccactgtttataaactgaggccccaccctcgacagtaccgggg
aggaagaggcctgcactagtccagagggaactgaggtcagggcCagctcgcccatagacatacatggcaggcaggctttgg
ccaggatccctccgctgccaggcgtctccctgccctccctcctcctgcctagagacccccaccctcaagcctggctggtctttgcctgag
acccaaacctcttcgacttcaagagaatatttaggaacaagggtgttagggccttctggaacaggccttgaccctttaagaaatga
cccaaagtctccttgacaaaaaggggaccccaactaaagggaagcctctcttctgctgtctcccctgacccactccccccac
cccaggacgaggagataaccagggtgaaagaggcccgctgggggctgcagacatgcttgcctgcctggcgaaggattgg
caggcttccccgtcacaggacccccgtggctgactcaggggcgcaggccttttgcgggggagctggcctccccgccccacggc
cacgggcccgccttcttgaggacagcgggatcttcagctgtcaggggaggggagggcgggggctgatgtcaggaggggatac
aaatagtccgacggctgggggcctGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCT
GCCCCGCCCTCCTCCGTGCGCCCCGCCAGCCTCGCCCCGGCTAGCgccgccACCatggac
cgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaaccactgaaaaactacgcggaGGATCGGGACGAAG
CCGACTCTTGGAAGACTTCAGAAACAATCGGTATCCGAACCTTCAGCTGAGAGA
AATTGCTGGTCACATCATGGAATTTTCTCAAGATCAACATGGAAGCCGGTTTATT
GAACTTAAACTCGAACGAGCCACCCCGGCCGAAAGGCAATTGGTGTTCATGAA
ATTCTTCAGGCCGCATACCAACTCATGGTTGATGTTTTTGGGAAGTATGTTATTCA
AAAGTTTTTTGAGTTCGGGTCCTGGAGCAAAAGTTGGCATTGGCAGAGCGAAT
CCGGGGCCATGTTCTGAGCCTCGCTCTCAAATGTACGGTAGTTATGTCATTTCG
AAAGCACTCGAGTTCATACCATCAGATCAACAGAATGAGATGGTGGGGAGCTG
GATGGGCATGTTTTGAAATGCGTGAAAGACCAAAACGGTAGCTACGTAGTTGAG
AAATGCATCGAATGCGTCCAACCACAGTCTCTCCAATTTATTATAGATGCATTTA
AGGGTCAGGTTTTTCGCGCTTTCTACGCACCCGTATGGGAACCGAGTGATTCAGAG
AATCTTGAGCACTGCCTGCCGGATCAGACACTCCCTATCTTGAGGAATTGCAC
CAGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATATGTTATTCGGCACG
TTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGATTAGAG
GTAACGTATTGGTGCTCTCACAACACAAATTTGCATCTAATGTGGTGGAGAAATG
```

TGTTACTCATGCTTCTAGAACGGAAGGGCAGTTCTCATAGACGAAGTTTGCACA
 ATGAATGATGGTCCTCATAGCGCACTTTATACCATGATGAAGGACCAGTATGCAA
 ACTATGTCGTCCAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGGAATTCG
 TGATGCACAAAATCCGACCTCACATTGCTACACTCAGAAAATACACGTATGGAA
 AACATATTCTGGCTAAGCTGGAGAAATATTACATGAAGAATGGAGTGGATCTGG
 GGtaacGCGGCCGCTCGAGGATTATAAGGATGACGACGATAAATTCGTGAGCAC
 CACCACCACCACCTAATAAGGTTTATCCGATCCACCGGATCTAGATAAGATAT
 CCGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAG
 GTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAA
 TGAATGCAATTGTTGTTGTTaatcaacctctggattacaaaatttgtaaagattgactggtattcttaactatgttgctc
 cttttacgtatgtggatacgtctttaatgcctttgtatcatgtattgcttccgtatggctttcatttctcctcctgtataaatcctggttg
 ctgtctctttatgaggagttgtggccgtgtcaggcaacgtggcggtgtgactgtgttgctgacgcaacccccactggttggggc
 attgccaccacctgtcagctcctttccgggactttcgtttccccctccctattgccacggcggaactcatcgccgctgccttgcccgt
 gctggacaggggctcggctgttgggactgacaattccgtggtgtgtcggggaaatcatcgctccttcttggtgctgcctgctgttg
 ccacctggattctgcgcgggacgtccttctgctacgtcccttcggccctcaatccagcggaacttcttcccggcgctgctgccggct
 ctgcgccctcttccgctcttcgcttcgcccctcagacgagtcggatctccctttggggcgccctcccgcgaactgtttattgagcttata
 atggtfacaataaagcaatagcatcacaatttcacaataaagcatttttctactgcattctagtgtggtttgtccaaactcatcaatgta
 tcttaACGCGGTAACACGTGCGGACCCAACGGCCGCaggaaccctagtgtgagttggccactcc
 ctctctgcgcgctcgtcgtcactgaggccggcgaccaaaggtcgccgacgcccgggctttgcccggcgccctcagtgcg
 gagcgagcgcgagctgcctgcagg (SEQ ID NO: 48).

[0138] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 4. The vector, A04949 comprises from 5' to 3' the elements set forth in Table 4.

Table 4: scAAV short hDesmin_ SIE intron PUF CUG_ WPRE3_ SV40pA A04949 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequences
5' ITR	Ctgcgcgctcgtcgtcactgaggccgcccgggcaagccccggcgctcgggagacatttggtcgcggcctcagtgagcgagcgagcgagagagggagtggggtt (SEQ ID NO: 33)
Short Human Desmin promoter	TGGCCTGGACCCCGTGGTAACCCTATAAGGCGAGGCA GCTGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCC TTGTTTCCCAGCCATGCGTTCTCCTCTATAAATACCCGC TCTGGTATTGGGGTTGGCAGCTGTTGCTGCCAGGGAG ATGGTTGGGTTGACATGCGGCTCCTGACAAAACAAAA CCCCTGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGG GATGAATCAGGGAGGGGGCGGGGACCCAGGGAGGGG CGCAGGCCTTTTTCGGGGGAGCTGGCCTCCCCGCCCCC ACGGCCACGGGCCGCCCTTTCCTGGCAGGACAGCGGGA

	TCTTGCAGCTGTCAGGGGAGGGGAGGCGGGGGCTGATG TCAGGAGGGGATAAATAAGTGCCGACGGCTGGGGGCCC TGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCT GCCCCGCCGCTCCTCCGTGCGCCCCGCCAGCCTCGCCCCG GCCGTACGCGTc (SEQ ID NO: 22)
SIE intron	Aggtaagttggctgtaaacaagtgaatttgagttgatagagtactgtctgccgccagattct aaaaataaccacctggacagcagcaccgaaggctttggcgggtccctgtctcttccctg gcttcagctcaccgtggccattcataggtatttagtatgctgtatatattttcaggt (SEQ ID NO: 6)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	Gaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaaccactgaaaaaacta cgc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-CUG	GGACGAAGCCGACTCTTGGAAGACTTCAGAAACAATCG GTATCCGAACCTTCAGCTGAGAGAAATTGCTGGTCACA TCATGGAATTTTCTCAAGATCAACATGGAAGCCGGTTTA TTGAACTTAACTCGAACGAGCCACCCCGGCCGAAAGG CAATTGGTGTTCAATGAAATTCTTCAGGCCGCATACCAA CTCATGGTTGATGTTTTTGGGAACTATGTTATTCAAAAG TTTTTTGAGTTCGGGTCCTGGAGCAAAAGTTGGCATTG GCAGAGCGAATCCGGGGCCATGTTCTGAGCCTCGCTCT CCAAATGTACGGTAGTTATGTCATTTCGCAAAGCACTCG AGTTCATACCATCAGATCAACAGAATGAGATGGTGCGG GAGCTGGATGGGCATGTTTTGAAATGCGTGAAAGACCA AAACGGTAGCTACGTAGTTGAGAAATGCATCGAATGCG TCCAACCACAGTCTCTCCAATTTATTATAGATGCATTTA AGGGTCAGGTTTTTCGCGCTTTCTACGCACCCGTATGGGA ACCGAGTGATTACAGAGAATCTTGGAGCACTGCCTGCCG GATCAGACACTCCCTATCTTGGAGGAATTGCACCAGCA TACCGAACAATTGGTGCAAGATCAATACGGTTCATATG TTATTCGGCACGTTCTTGAGCATGGAAGGCCAGAGGAC AAGTCAAAGATCGTCGCTGAGATTAGAGGTAACGTATT GGTGCTCTCACAACACAAATTTGCATCTAATGTGGTGG AGAAATGTGTTACTCATGCTTCTAGAACGGAAAGGGCA GTTCTCATAGACGAAGTTTGCACAATGAATGATGGTCCT CATAGCGCACTTTATAACCATGATGAAGGACCAGTATGC AACTATGTGCTCCAGAAAATGATCGATGTGGCGGAGC CCGGTCAACGGAAAATCGTGATGCACAAAATCCGACCT CACATTGCTACACTCAGAAAATACACGTATGGAAAACA TATTCTGGCTAAGCTGGAGAAAATATTACATGAAGAATG GAGTGGATCTGGGGtaa (SEQ ID NO: 45)
WPRE3 (short)	GATAATCAACCTCTGGATTACAAAATTTGTGAAAGATT GACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGT GGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCT TCCCGTATGGCTTTTCAATTTCTCCTCCTTGTATAAATCCT GGTTAGTTCTTGCCACGGCGGAACATCGCCGCCTGCC TTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACT GACAATCCGTGG (SEQ ID NO: 132)

SV40-PA	Aacttgtttattgcagcttataatggtfacaataaagcaatagcatcacaatttcacaaataa agcattttttcactgcattctagtgtggtttgtccaaactcatcaatgtatctta (SEQ ID NO: 39)
3' ITR	Aggaacccctagtgtgagttggccactccctctctgcgcgctcgcctcactgaggc cgggcgaccaaaggctcgccgacgcccgggctttgccggggcgccctcagtgagcgag cgagcgcgagctgcctgcagg (SEQ ID NO: 34)

[0139] In some embodiments, rAAV vector A04949 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

ctgcgcgctcgcctcactgagggcccgccgggcaaagcccgggcgtcgggcgacctttggtcgccggcctcagtgagcgagc
gagcgcgagagaggagtggttCTTCGAAACACCGGTTGGCCTGGACCCCCGTGGTAACC
CTATAAGGCGAGGCAGCTGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCT
TGTTTCCCAGCCATGCGTTCTCCTCTATAAATACCCGCTCTGGTATTTGGGGTTGG
CAGCTGTTGCTGCCAGGGAGATGGTTGGGTTGACATGCGGCTCCTGACAAAACA
CAAACCCCTGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATCAGGG
AGGGGGCGGGGGACCCAGGGAGGGGGCGCAGGCCTTTTGCGGGGGAGCTGGCCT
CCCCGCCCCACGGCCACGGGCCGCCCTTTCCTGGCAGGACAGCGGGATCTTGC
AGCTGTCAGGGGAGGGGAGGCGGGGGCTGATGTCAGGAGGGGATAaAAATAGTGC
CGACGGCTGGGGGCCCTGTCTCCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCT
GCCCCGCCCTCCTCCGTGCGCCCCGCCAGCCTCGCCCCGCGCCGTCACGCGTcagta
agttggctgtaacaaagttgaattgagttgatagagtactgtctgccccagattctaaaaataaccacctggacagcagcacccaa
ggtctttggcggtccctgttctcttccctggcttcagctaccgtggtccattcataggtatttagtatgctgtatatattttcaggtGC
TAGCgccgccACCatggaccgtgtgctcaaaagaagtgtgaaggaagcaacctcctaaccactgaaaaactacgcgga
GGATCGGGACGAAGCCGACTCTTGAAGACTTCAGAAACAATCGGTATCCGAAC
CTTCAGCTGAGAGAAATTGCTGGTCACATCATGGAATTTTCTCAAGATCAACATG
GAAGCCGGTTTATTGAACTTAACTCGAACGAGCCACCCCGGCCGAAAGGCAAT
TGGTGTTCATGAAATCTTCAGGCCGCATACCAACTCATGGTTGATGTTTTTGG
GAACTATGTTATTCAAAAGTTTTTTGAGTTCGGGTCACTGGAGCAAAAGTTGGCA
TTGGCAGAGCGAATCCGGGGCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTA
GTTATGTCATTCGCAAAGCACTCGAGTTCATACCATCAGATCAACAGAATGAGAT
GGTGCGGGAGCTGGATGGGCATGTTTTGAAATGCGTGAAAGACCAAAACGGTAG
CTACGTAGTTGAGAAATGCATCGAATGCGTCCAACCACAGTCTCTCCAATTTATT
ATAGATGCATTTAAGGGTCAGGTTTTTCGCGCTTTCTACGCACCCGTATGGGAACC
GAGTGATTGAGAGAATCTTGGAGCACTGCCTGCCGGATCAGACACTCCCTATCTT
GGAGGAATTGCACCAGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATA

TGTTATTCGGCACGTTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAGATCGT
 CGCTGAGATTAGAGGTAACGTATTGGTGCTCTCACAACACAAATTTGCATCTAAT
 GTGGTGGAGAAATGTGTTACTCATGCTTCTAGAACGGAAAGGGCAGTTCTCATA
 GACGAAGTTTGCACAATGAATGATGGTCCTCATAGCGCACTTTATACCATGATGA
 AGGACCAGTATGCAAACCTATGTCGTCCAGAAAATGATCGATGTGGCGGAGCCCCG
 GTCAACGGAAAATCGTGATGCACAAAATCCGACCTCACATTGCTACACTCAGAA
 AATACACGTATGGAAAACATATTCTGGCTAAGCTGGAGAAATATTACATGAAGA
 ATGGAGTGGATCTGGGG^{taa}GATAATCAACCTCTGGATTACAAAATTTGTGAAAGA
 TTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTT
 AATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTA
 TAAATCCTGGTTAGTTCTTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCCG
 TGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGG^{Gaacttg}ttattgcagcttat
 aatgggtacaataaagcaatagcatcacaaattcacaaataaagcattttttactgcattctagttgtggttgccaactcatcaatgt
 atcttaCACGTGCGGACCCAACGGCCGC^{Gaggaacc}ctagtgatggagtggccactccctctctgcgcgctcgc
 tcgctcactgaggccggcgaccaaaggtcgcccgcgcccgggctttgcccggcgccctcagtgagcgagcgagcgcgag
 ctgcctgcagg (SEQ ID NO: 49).

[0140] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 5. The vector, A04959 comprises from 5' to 3' the elements set forth in Table 5.

Table 5: ssAAV insulator9 _short_hDesmin_ SIE intron_ PUFCUG_WPRE_SV40pA A04959 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequences
5' ITR	Cctgcaggcagctgcgcgctcgtcgtcactgaggccgcccgggctggggcgaccttt ggtcgcccggcctcagtgagcgagcgagcgcgagagaggagtggccaactccatcact aggggttct (SEQ ID NO: 32)
Insulator	TATACTCACCCGGACACCGACGTGAACCTGCAGCCTCCC CCGGGGAGCACGCATGGGTCCCAGTTGCCTTACAGTCTG ATCACGGATGCTGTTGCCTAGGCAACGTGAGGGTGGCCC TGGCTCCTTGATCTTCCTGGTGGCTCTGGGGGACCATGA CGGAGATGCGGATAATTCAGGTCAGGGCAGCAGCCGGC CGGCAGGGGGCGACAAGCACACTGCTCATGAACGGTTG GCTCTGCCCACGTGCACACGGCGTTTCCACATTAACATT CATTTAATAATAATAAAAGAATAAGCAACGAGAGAGAA AGCCAGAGATCTACATGACTAGACTTGCTGCATTTCTAA ATTACCAGGTACGACGTGAGCCGAGCATCTGCCCCCTTG CCCTGACCAGGTCCCTAGT (SEQ ID NO: 41)
Short Human Desmin promoter	TGGCCTGGACCCCCGTGGTAACCTATAAGGCGAGGCAG CTGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCT TGTTTCCCAGCCATGCGTTCTCCTCTATAAATACCCGCTC TGGTATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGATG

	GTTGGGTTGACATGCGGCTCCTGACAAAACACAAACCCC TGGTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATG AATCAGGGAGGGGGCGGGGGACCCAGGGAGGGGGCGCA GGCCTTTTGCAGGGGAGCTGGCCTCCCCGCCCCACGGC CACGGGCCGCCCTTTCCTGGCAGGACAGCGGGATCTTGC AGCTGTCAGGGGAGGGGAGGCGGGGGCTGATGTCAGGA GGGATAAATAAGTGCCGACGGCTGGGGGCCCTGTCTCC CCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCC GCCTCCTCCGTGCGCCCGCCAGCCTCGCCCGCGCCGTCA CGCGTc (SEQ ID NO: 22)
SIE intron	Aggtaagttggctgtaacaaagttgaatttgagttgatagagtactgtctgccgccagattcta aaaataaccaccctggacagcagcaccgaaggtcttggcgggtccctgttctcttcctgg cttcagctcaccgtggtccattcataggtatttagtatgctgtatatattttcaggt (SEQ ID NO: 6)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	Gaccgtgtgctcaaaagaagtgctgaaggaagcaaccctcctaaccactgaaaaactac gc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-CUG	GGACGAAGCCGACTCTTGGAAGACTTCAGAAACAATCG GTATCCGAACCTTCAGCTGAGAGAAATTGCTGGTCACAT CATGGAATTTTCTCAAGATCAACATGGAAGCCGTTTAT TGAACCTAAACTCGAACGAGCCACCCCGGCCGAAAGGC AATTGGTGTTCAATGAAATTCTTCAGGCCGCATACCAAC TCATGGTTGATGTTTTTGGGAAGTATGTTATTCAAAGTT TTTTGAGTTCGGGTCACTGGAGCAAAAGTTGGCATTGGC AGAGCGAATCCGGGGCCATGTTCTGAGCCTCGCTCTCCA AATGTACGGTAGTTATGTCATTCGCAAAGCACTCGAGTT CATACCATCAGATCAACAGAATGAGATGGTGCAGGAGC TGGATGGGCATGTTTTGAAATGCGTGAAAGACCAAAAC GGTAGCTACGTAGTTGAGAAATGCATCGAATGCGTCCAA CCACAGTCTCTCCAATTTATTATAGATGCATTTAAGGGTC AGGTTTTTCGCGCTTTCTACGCACCCGTATGGGAACCGAG TGATTCAGAGAATCTTGGAGCACTGCCTGCCGGATCAGA CACTCCCTATCTTGGAGGAATTGCACCAGCATACCGAAC AATTGGTGCAAGATCAATACGGTTCATATGTTATTTCGGC ACGTTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAG ATCGTCGCTGAGATTAGAGGTAACGTATTGGTGCTCTCA CAACACAAATTTGCATCTAATGTGGTGGAGAAATGTGTT ACTCATGCTTCTAGAACGGAAGGGCAGTTCTCATAGAC GAAGTTTGCACAATGAATGATGGTCCTCATAGCGCACTT TATACCATGATGAAGGACCAGTATGCAAACATATGTCGTC CAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGGAA AATCGTGATGCACAAAATCCGACCTCACATTGCTACACT CAGAAAATACACGTATGGAAAACATATTCTGGCTAAGCT GGAGAAATATTACATGAAGAATGGAGTGGATCTGGGGta a (SEQ ID NO: 45)
WPRE	Aatcaacctctggattacaaatttgtaagattgactggattcttaactatgttgctcctttac gctatgtggatacgtgctttaatgcctttgtatcatgctattgcttcccgatggcttcttctcc tcctgtataaatcctggttctctttatgaggagttgtggcccggttcagggaacgtggc

	gtggtgtgcactgtgtttgctgacgaacccccactggtggggcattgccaccacctgtcagc tcctttccgggactttcgtttccccctccctattgccacggcggaactcatcgccgcctgcctt gcccgtgctggacaggggctcggctgttgggcaactgacaattccgtggtgtgtcggggaa atcatcgctcctttcctggctgctgcctgtgttgcacctggattctgcggggacgtccttctg ctacgtcccttcggccctcaatccagcgaccttcctcccggcctgctgcccggctctgcg gcctcttcggcgtcttcgccttcgccctcagacgagtcggatctcccttggggccgcctccccg c (SEQ ID NO: 38)
SV40-PA	Aacttggttattgcagcttataatggttacaataaagcaatagcatcacaatttcacaaataa gcattttttcactgcattctagtgtgtgttgcacaaactcatcaatgtatctta (SEQ ID NO: 39)
3'ITR	Aggaaccctagtgtgagttggccactccctctctgcgcgctcgtcgtcactgaggcc ggcgaccaaaggtcgcccgacggcggttggccggcctcagtgagcgagcg agcgcgagctgcctgcagg (SEQ ID NO: 34)

[0141] In some embodiments, rAAV vector A04959 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

cctgcaggcagctgcgcgctcgtcgtcactgaggccgccccggcgctggcgaccttggtcgccccgcctcagtgagcgagc
gagcgcgagagagggagtggccaactccatcactaggggtcctGCGGCCGGTCGCGTCTAGTActagTAT
ACTCACCCGGACACCGACGTGAACCTGCAGCCTCCCCCGGGGAGCACGCATGGG
TCCCAGTTGCCTTACAGTCTGATCACGGATGCTGTTGCCTAGGCAACGTGAGGGT
GGCCCTGGCTCCTTGATCTTCCTGGTGGCTCTGGGGGACCATGACGGAGATGCGG
ATAATTCAGGTCAGGGCAGCAGCCGGCCGGCAGGGGGCGACAAGCACACTGCTC
ATGAACGGTTGGCTCTGCCCACGTGCACACGGCGTTTCCACATTAACATTCATTT
AATAATAATAAAAGAATAAGCAACGAGAGAGAAAGCCAGAGATCTACATGACT
AGACTTGCTGCATTTCTAAATTACCAGGTACGACGTGAGCCGAGCATCTGCCCCC
TTGCCCTGACCAGGTCCCTAGTTGGCCTGGACCCCCGTGGTAACCCCTATAAGGCG
AGGCAGCTGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCTTGTTCCTCAG
CCATGCGTTCTCCTCTATAAATACCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGC
TGCCAGGGAGATGGTTGGGTTGACATGCGGCTCCTGACAAAACACAAACCCCTG
GTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATCAGGGAGGGGGCGG
GGGACCCAGGGAGGGGCGCAGGCCTTTTGCGGGGGAGCTGGCCTCCCCGCCCCC
ACGGCCACGGGCGGCCCTTTCCTGGCAGGACAGCGGGATCTTGACAGCTGTCAGG
GGAGGGGAGGCGGGGGCTGATGTCAGGAGGGATtAAATAGTGCCGACGGCTGG
GGGCCCTGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCC
TCCTCCGTGCGCCCCGCCAGCCTCGCCCCGCGCCGTCACGCGTcaggttaagttggctgtaaacaaa
gttgaatttgagttgatagagtactgtctgccgccagattctaaaaataaccacctggacagcagcaccgaaggtcttggcgggtccc
tgttctcttccctggcttcagctcaccgtggtccattcataggtatttagtatgctgtatatattttcaggtGCTAGCgccgccAC

CatggaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaccactgaaaaactacgcggaGGATCGGGAC
GAAGCCGACTCTTGGAAGACTTCAGAAACAATCGGTATCCGAACCTTCAGCTGA
GAGAAATTGCTGGTCACATCATGGAATTTTCTCAAGATCAACATGGAAGCCGGTT
TATTGAACTTAAACTCGAACGAGCCACCCCGGCCGAAAGGCAATTGGTGTTCAA
TGAAATTCCTTCAGGCCGCATACCAACTCATGGTTGATGTTTTTGGGAACTATGTT
ATTCAAAAGTTTTTTGAGTTCGGGTCACTGGAGCAAAAGTTGGCATTGGCAGAGC
GAATCCGGGGCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTAGTTATGTCAT
TCGCAAAGCACTCGAGTTCATACCATCAGATCAACAGAATGAGATGGTGCGGGA
GCTGGATGGGCATGTTTTGAAATGCGTGAAAGACCAAAACGGTAGCTACGTAGT
TGAGAAATGCATCGAATGCGTCCAACCACAGTCTCTCCAATTTATTATAGATGCA
TTTAAGGGTCAGTTTTTCGCGCTTTCTACGCACCCGTATGGGAACCGAGTGATTC
AGAGAATCTTGGAGCACTGCCTGCCGGATCAGACACTCCCTATCTTGGAGGAATT
GCACCAGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATATGTTATTCGG
CACGTTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGATT
AGAGGTAACGTATTGGTGCTCTCACAACACAAATTTGCATCTAATGTGGTGGAGA
AATGTGTTACTCATGCTTCTAGAACGGAAAGGGCAGTTCTCATAGACGAAGTTTG
CACAATGAATGATGGTCCTCATAGCGCACTTTATACCATGATGAAGGACCAGTAT
GCAAACATATGTCGTCCAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGGAAA
ATCGTGATGCACAAAATCCGACCTCACATTGCTACACTCAGAAAATACACGTATG
GAAAACATATTCTGGCTAAGCTGGAGAAATATTACATGAAGAATGGAGTGGATC
TGGGGtaatcGCGGCCGCTCGAGGATTATAAGGATGACGACGATAAATTCGTGAG
CACCACCACCACCACCTAATAAGGTTTATCCGATCCACCGGATCTAGATAAGA
TATCCGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTA
GAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATA
AAATGAATGCAATTGTTGTTGTTaatcaacctctggattacaaaatttgtaaagattgactggattcttaactatgt
tgctccttttacgctatgtggatacgtgcttaatgcctttgtatcatgctattgctcccgatggctttcattttctcctcctgtataaatcct
gggtgctgtctctttatgaggagttgtggcccggtgtcaggcaacgtggcgtggtgtgactgtgtttgctgacgcaacccccactgggtg
gggcattgccaccacgtcagctcctttccgggactttcgtttccccctccctattgccacggcggaactcatcgccgctgccttgc
ccgctgctggacaggggctcggctgttggcactgacaattccgtggtgtgtcgggaaatcatcgtcctttccttgctgctgcctg
tgttgccacctggattctgcgaggacgtccttctgtacgtcccttcggccctcaatccagcggaacttcttcccgcggtgctgccc
ggctctgcggccttctccgctcttcgcttcgccctcagacgagtcggatctccctttgggcccctccccgcaactgtttattgcagc
ttataatggttacaaataaagcaatagcatcacaatttcacaaataaagcattttttcactgcattctagtgtggtttgtccaaactcatca
atgtatcttaACGCGGTAACCACGTGCGGACCCAACGGCCGCaggaaccctagtgatggagttggcca

ctccctctctgcgcgctcgtcgtcactgaggccggcgaccaaaggtcgcccgacgcccgggctttgcccggggcgccctcagtg
agcgagcgagcgcgcagctgcctgcagg (SEQ ID NO: 50).

[0142] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 6. The vector, A04546 comprises from 5' to 3' the elements set forth in Table 6.

Table 6: ssAAV insulator9_short_hDesmin_SIE intron_PUFCUGmyc tag_WPRE_SV40pA_A04546 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequences
5' ITR	Cctgcaggcagctgcgcgctcgtcgtcactgaggccggcgaccaaaggtcgcccgacgcccgggctttgcccggggcgccctcagtg cgcccggcctcagtgagcgcgcagcgcgcagagagggagtggccaactccatcactagggg ttcct (SEQ ID NO: 32)
Insulator	TATACTACCCGGACACCGACGTGAACCTGCAGCCTCCCC GGGGAGCACGCATGGGTCCCAGTTGCCTTACAGTCTGATC ACGGATGCTGTTGCCTAGGCAACGTGAGGGTGGCCCTGGC TCCTTGATCTTCCTGGTGGCTCTGGGGGACCATGACGGAGA TGCGGATAATTCAGGTCAGGGCAGCAGCCGGCCGGCAGGG GGCGACAAGCACACTGCTCATGAACGGTTGGCTCTGCCA CGTGCACACGGCGTTTCCACATTAACATTCAATTAATAATA ATAAAAGAATAAGCAACGAGAGAGAAAGCCAGAGATCTA CATGACTAGACTTGCTGCATTTCTAAATTACCAGGTACGAC GTGAGCCGAGCATCTGCCCCCTTGCCCTGACCAGGTCCCTA GT (SEQ ID NO: 41)
Short Human Desmin promoter	TGGCCTGGACCCCCGTGGTAACCCTATAAGGCGAGGCAGC TGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCTTGT TTCCCAGCCATGCGTTCTCCTCTATAAAATACCCGCTCTGGT ATTTGGGGTTGGCAGCTGTTGCTGCCAGGGAGATGGTTGG GTTGACATGCGGCTCCTGACAAAACACAAACCCCTGGTGT GTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATCAGGG AGGGGGCGGGGGACCCAGGGAGGGGGCGCAGGCCTTTTGCG GGGGAGCTGGCCTCCCCGCCCCACGGCCACGGGCGGCC TTTCCTGGCAGGACAGCGGGATCTTGCAGCTGTCAGGGGA GGGGAGGCGGGGGCTGATGTCAGGAGGGATaAAATAGTG CCGACGGCTGGGGGGCCCTGTCTCCCCTCGCCGCATCCACTC TCCGGCCGGCCGCTGCCCCGCCCTCCTCCGTGCGCCCCG CAGCCTCGCCCCGCGCCGTCACGCGTc (SEQ ID NO: 22)
SIE intron	Aggtaagttggctgtaacaaagttgaatttgagttgatagagtactgtctgccccagattctaaa aataaccaccctggacagcagcaccgaaggtcttggcgggtccctgttctcttccctggcttcag ctcaccgtggtccattcataggtatttagtatgctgtatatattttcaggt (SEQ ID NO: 40)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	Gaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaccactgaaaaactacgc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-CUG	GGACGAAGCCGACTCTTGGAAGACTTCAGAAACAATCGGT ATCCGAACCTTCAGCTGAGAGAAATTGCTGGTCACATCATG GAATTTTCTCAAGATCAACATGGAAGCCGGTTTATTGAACT TAAACTCGAACGAGCCACCCCGGCCGAAAGGCAATTGGTG

	TTCAATGAAATTCTTCAGGCCGCATACCAACTCATGGTTGATGTTTTTGGGAACTATGTTATTCAAAGTTTTTTGAGTTCGGTCACTGGAGCAAAAGTTGGCATTGGCAGAGCGAATCCGGGGCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTAGTTATGTCATTTCGCAAAGCACTCGAGTTCATACCATCAGATCAACAGAATGAGATGGTGCGGGAGCTGGATGGGCATGTTTTGAAATGCGTGAAAGACCAAAACGGTAGCTACGTAGTTGAGAAAATGTCATCGAATGCGTCCAACCACAGTCTCTCCAATTTATTATAGATGCATTTAAGGGTTCAGGTTTTTCGCGCTTTCTACGCACCCTATGGGAACCGAGTGATTTCAGAGAATCTTGGAGCACTGCCTGCCGGATCAGACACTCCCTATCTTGGAGGAATTGCACCAGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATATGTTATTCGGCACGTTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGATTAGAGGTAACGTATTGGTGCTCTCACAAACAAAATTTGCATCTAATGTGGTGGAGAAATGTGTTACTCATGCTTCTAGAACGAAAGGGCAGTTCTCATAGACGAAGTTTGCACAATGAATGATGGTCCTCATAGCGCACTTTATACCATGATGAAGGACCAGTATGCAAACATATGTCGTCCAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGAAAATCGTGATGCACAAAATCCGACCTCACATTGCTACACTCAGAAAATACACGTATGGAAAACATATTCTGGCTAAGCTGGAGAAATATTACATGAAGAATGGAGTGGATCTGGGG (SEQ ID NO: 46)
Linker	GGCGGAAGT (SEQ ID NO: 42)
Myc tag	GAGCAAAAAGTATTAGTGAAGAAGATCTC (SEQ ID NO: 43)
Stop codon	TAA
WPRES	Aatcaacctctggattacaaaatttgtaagattgactggattcttaactatgtgctcctttacgctatgtggatacgtgctttaatgcctttgtatcatgctattgctcccgatggctttcattttctcctctgtataaatcctggtgctgtctctttatgaggagttgtggcccggtgtcaggcaacgtggcgtggtgtgcactgtgttgctgacgaacccccactgggtggggcattgccaccactgtcagctcctttccgggactttcgtttccccctcctattgccacggcggaactcatcgccgctgccttggccgctgtggaagggggctggctgtgggactgacaattccgtggtgtgtcggggaaatcatgctcctttccttgctgctgcctgtgttgccacctggattctgcgcgggacgtccttctgtacgtccctcgccctcaatccagcggaccttccccggcctgtgcgggctgtgcggccttcccgcttctgccttcgccctcagacgagtcggatctcccttggggcgcctccccgc (SEQ ID NO: 38)
SV40-PA	Aactgtttattgcagcttataatggttacaaataaagcaatagcatcacaatttcacaaataaagcatttttactgcattctagttgtgtgtgttgcacaaactcatcaatgtatctta (SEQ ID NO: 39)
3'ITR	Aggaaccctagtgtgagtggtggcactccctctctgcgcgctcgtcgtcactgaggccgggacgacaaaggctgcccagcggcggtttgcccggcgccctcagtgagcgagcgcgcagctgcctgcagg (SEQ ID NO: 34)

[0143] In some embodiments, rAAV vector A04546 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

cctgcaggcagctgcgcgctcgtcgtcactgaggccgcccggcgctggcgacctttggtcgcccggcctcagtgagcgagc
gagcgcgcagagaggagtgccaaactcatcactaggggttctGCGGCCGGTCGCGTCTAGTActagTAT

ACTCACCCGGACACCGACGTGAACCTGCAGCCTCCCCCGGGGAGCACGCATGGG
TCCCAGTTGCCTTACAGTCTGATCACGGATGCTGTTGCCTAGGCAACGTGAGGGT
GGCCCTGGCTCCTTGATCTTCCTGGTGGCTCTGGGGGACCATGACGGAGATGCGG
ATAATTCAGGTCAGGGCAGCAGCCGGCCGGCAGGGGGCGACAAGCACACTGCTC
ATGAACGGTTGGCTCTGCCCACGTGCACACGGCGTTTCCACATTAACATTCATTT
AATAATAATAAAAGAATAAGCAACGAGAGAGAAAAGCCAGAGATCTACATGACT
AGACTTGCTGCATTTCTAAATTACCAGGTACGACGTGAGCCGAGCATCTGCCCCC
TTGCCCTGACCAGGTCCCTAGTTGGCCTGGACCCCCGTGGTAACCCTATAAGGCG
AGGCAGCTGCTGTCTGAGCTAGtCCCACAGCTCCTCTCCTGTGCCTTGTTTCCCAG
CCATGCGTTCTCCTCTATAAATACCCGCTCTGGTATTTGGGGTTGGCAGCTGTTGC
TGCCAGGGAGATGGTTGGGTTGACATGCGGCTCCTGACAAAACACAAACCCCTG
GTGTGTGTGGGCGTGGGTGGTGTGAGTAGGGGGATGAATCAGGGAGGGGGCGG
GGGACCCAGGGAGGGGCGCAGGCCTTTTGCGGGGGAGCTGGCCTCCCCGCCCCC
ACGGCCACGGGCGGCCCTTTCCTGGCAGGACAGCGGGATCTTGACAGCTGTCAGG
GGAGGGGAGGCGGGGGCTGATGTCAGGAGGGATtAAATAGTGCCGACGGCTGG
GGGCCCTGTCTCCCCTCGCCGCATCCACTCTCCGGCCGGCCGCCTGCCCGCCGCC
TCCTCCGTGCGCCCCGCCAGCCTCGCCCCGCGCCGTCACGCGTcaggttaagttggctgtaaacaaa
gtgaatttgagttgataagtagtactgtctgccgccagattctaaaaataaccaccctggacagcagcacccaaggtcttggcgggtccc
tgttctcttccctggcttcagctcaccgtgggtccattcataggtatttagtatgtgtatatattttcaggtGCTAGCgccgccAC
CatggaccgtgtgtcaaaagaagtgtgaaggaagcaaccctctaaccactgaaaaactacgcggaGGATCGGGAC
GAAGCCGACTCTTGGAAGACTTCAGAAACAATCGGTATCCGAACCTTCAGCTGA
GAGAAATTGCTGGTCACATCATGGAATTTTCTCAAGATCAACATGGAAGCCGGTT
TATTGAACTTAAACTCGAACGAGCCACCCCGGCCGAAAGGCAATTGGTGTTCAA
TGAAATCTTCAGGCCGCATACCAACTCATGGTTGATGTTTTTGGGAACTATGTT
ATTCAAAAGTTTTTTGAGTTCGGGTCACTGGAGCAAAAGTTGGCATTGGCAGAGC
GAATCCGGGGCCATGTTCTGAGCCTCGCTCTCCAAATGTACGGTAGTTATGTCAT
TCGCAAAGCACTCGAGTTCATACCATCAGATCAACAGAATGAGATGGTGCGGGA
GCTGGATGGGCATGTTTTGAAATGCGTGAAAGACCAAAACGGTAGCTACGTAGT
TGAGAAATGCATCGAATGCGTCCAACCACAGTCTCTCCAATTTATTATAGATGCA
TTTAAGGGTCAGGTTTTTCGCGCTTTCTACGCACCCGTATGGGAACCGAGTGATTC
AGAGAATCTTGAGCACTGCCTGCCGGATCAGACACTCCCTATCTTGAGGAATT
GCACCAGCATACCGAACAATTGGTGCAAGATCAATACGGTTCATATGTTATTCGG
CACGTTCTTGAGCATGGAAGGCCAGAGGACAAGTCAAAGATCGTCGCTGAGATT

AGAGGTAACGTATTGGTGCTCTCACAAACACAAATTTGCATCTAATGTGGTGGAGA
AATGTGTTACTCATGCTTCTAGAACGGAAGGGCAGTTCTCATAGACGAAGTTTG
CACAATGAATGATGGTCCTCATAGCGCACTTTATACCATGATGAAGGACCAGTAT
GCAAACATATGTCGTCCAGAAAATGATCGATGTGGCGGAGCCCGGTCAACGGAAA
ATCGTGATGCACAAAATCCGACCTCACATTGCTACACTCAGAAAATACACGTATG
GAAAACATATTCTGGCTAAGCTGGAGAAATATTACATGAAGAATGGAGTGGATC
TGGGGGGCGGAAGTGAGCAAAAACCTGATTAGTGAAGAAGATCTCtaacGCGGCCG
CTCGAGGATTATAAGGATGACGACGATAAATTCGTCGAGCACCACCACCACCAC
CACTAATAAGGTTTATCCGATCCACCGGATCTAGATAAGATATCCGATCCACCGG
ATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCT
TTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTG
TTGTTGTTaatcaacctctggattacaaaatttgtaaagattgactggattcttaactatgttgctccttttacgctatgtggatacg
ctgctttaatgcctttgtatcatgctattgctcccgatggctttcattttctctccttgataaatcctgggtgctgtctctttatgaggagttgt
ggcccggttgtaggcaacgtggcggtgtgactgtgtttgctgacgcaacccccactgggtggggcattgccaccacctgtcagct
cctttccgggactttcgtttccccctccctattggcacggcggaactcatcgccgctgccttgcccgtgctggacaggggctcggc
tgttgggcactgacaattccgtggtgtgtcggggaaatcatcgtcctttccttggtgctgcctgtgttgccacctggattctgcgcgg
gacgtccttctgctacgtcccttcggccctcaatccagcggaaccttccttcccgcggcctgctgccggctctgcggccttctccgctct
tcgcttcgcctcagacgagtcggatctcccttgggccgctccccgcaactgtttattgcagcttataatggtfacaataaagcaat
agcatcacaatttcacaataaagcatttttactgcattctagtgtgtgttgcctcaactcatcaatgtatcttaACGCGGTAA
CCACGTGCGGACCCAACGGCCGCaggaacccctagtgtgaggtggccactccctctctgcgcgctcgctgc
tcactgagggccggcgaccaaagtcgcccgcacccgggcttggccggcgccctcagtgcgcgagcgagcgcgagctgc
ctgcagg (SEQ ID NO: 51).

[0144] In one embodiment, an rAAV vector comprising a CUG repeat targeting PUF operably linked to a hybrid regulatory sequence is set forth in Table 7. The vector, A04298 comprises from 5' to 3' the elements set forth in Table 7.

Table 7: ssAAV_hDesmin_PUF-lambda_myc tag_WPRE_SV40pA A04298 elements from 5' to 3'

Plasmid Element	Nucleic Acid Sequences
5' ITR	Cctgcaggcagctgcgcgctcgtcgtcactgagccgccccggcgctcggcgacc tttggtcgcggccctcagtgcgcgagcgagcgcgagagaggagtgccaactcca tcactaggggttcct (SEQ ID NO: 32)
Human Desmin promoter	tacccctgccccacagctcctcctgtgccttgttccagccatgcgttctctata aatacccgctctggtatttggggtggcagctgttgcctgagggagatggttgggtgac atgcggctcctgacaaaacacaaacctgggtgtgtgtggcgctgggtgtgtgagtag ggggatgaatcaggagggggcAggggacccagggggcaggagccacacaaagtc

	tgtgcgggggtgggagcgcacatagcaattggaactgaaagcttatcagaccctttctg gaaatcagcccactgtttataaacttgaggccccaccctcgacagtaccggggaggaag agggcctgcactagtcagagggaaactgaggctcagggcCagctcgcccatagacat acatggcaggcaggctttggccaggatccctccgctgcccaggcgtctccctgccctccc ttctgcctagagacccccaccctcaagcctggctggtctttgcctgagacccaaacctctt cgactcaagagaatatttaggaacaagggtggttagggcctttctgggaacaggccttg accctttaagaatgacccaaagtctctccttgacaaaaaggggacctcaactaaag ggaagcctctcttctgctgtctcccctgacccactccccccaccccaggacgaggaga taaccagggtgaaagaggcccgctgggggctgcagacatgcttgcctgccctggg cgaaggattggcaggcttgcctgtcacaggacccccgctggctgactcaggggagcag gcctTttgcggggagctggcctccccgccccacggccacggggcgccctttctgg caggacagcgggatcttgcagctgtcagggaggaggaggcggggctgatgtcagga gggatacaaatagtcgcgacggctgggggacctGTCTCCCTCGCCGCA TCCACTCTCCGCGCCGCGCCTGCCGCGCCTCCTCC GTGCGCCCCGCCAGCCTCGCCCCG (SEQ ID NO: 16)
Kozak	gccgccACCatg (SEQ ID NO: 30)
Rb NLS	Gaccgtgtgctcaaaagaagtgtgaaggaagcaaccctcctaaccactgaaaaaact acgc (SEQ ID NO: 36)
Linker	ggaGGATCG (SEQ ID NO: 37)
PUF-lambda	GGCAGATCTAGACTCCTCGAGGACTTTAGAAACAACA GATACCCTAATCTGCAGCTGAGGGAAATCGCCGGCCA CATTATGGAGTTCAGCCAAGACCAGCACGGCAATAGA TTCATCCAGCTCAAACCTGGAGAGGGCCACACCCGCCG AGAGGCAGCTGGTGTTCACGAGATTCTGCAAGCCGC CTACCAACTGATGGTGGATGTCTTCGGATGCTACGTCA TCCAGAAGTTCTTCGAATTTGGCTCTCTGGAGCAGAA GCTGGCCCTCGCTGAGAGAATCAGAGGACACGTGCTG AGCCTCGCTCTGCAGATGTATGGATCTAGAGTGATCA GAAAGGCTCTGGAGTTCATCCCCAGCGACCAGCAGAA CGAGATGGTGAGGGAACTCGATGGCCATGTCCTCAAG TCGTCAAGGATCAGAACGGCAGCTACGTGGTGGAGA AGTGTATCGAGTGTGTGCAACCCCAATCTCTGCAGTTC ATCATCGACGCCTTCAAGGGCCAAGTGTTCTGCTCTGA GCACCCATCCCTACGGCAGCAGAGTGATCGAAAGAAT CCTCGAGCATTGTCTGCCCCGACCAGACACTGCCTATTC TGGAGGAGCTGCACCAGCATAACCGAGCAACTGGTGCA AGACCAGTACGGCAACTACGTCATTACGACAGTGCTC GAACATGGAAGACCCGAAGACAAGAGCAAAATCGTG GCCGAGATCAGAGGCAATGTCCTCGTGCTGAGCCAGC ACAAGTTTGCTTGCAACGTGGTGCAGAAAGTGCCTGAC ACACGCTTCTAGAACCGAGAGGGCCGTGCTGATTGAC GAGGTGTGACCATGAACGACGGCCCCCATTCGCCC TCTACACCATGATGAAGGACCAATACGCTAGCTACGT CGTGAGGAAGATGATTGACGTGGCTGAGCCCGGCCAG AGGAAGATCGTGATGCATAAGATTAGACCCCATATCG CCACACTGAGGAAGTACACATACGGCAAGCACATTCT GGCCAAGCTCGAAAAGTACTATATGAAAAATGGCGTC GATCTGGGC (SEQ ID NO: 44)
Linker	GGCGGAAGT (SEQ ID NO: 42)

Myc tag	GAGCAAAACTGATTAGTGAAGAAGATCTC (SEQ ID NO: 43)
Stop codon	TAA
WPRE	Aatcaacctctggattacaaaattgtgaaagattgactggattcttaactatgttgctccttt tacgctatgtggatacgtgctttaatgcctttgtatcatgctattgcttcccgtatggctttcat tttctcctcctgtataaatcctggttgctgtctctttatgaggagttgtggcccggttgctaggc aacgtggcgtgggtgtgactgtgtttgctgacgcaacccccactgggtggggcattgccac cacctgtcagctcctttccggacttgcgttccccctccctattgccacggcggaactcat cgccgctgccttgcccgtgctggacaggggctcggtgttgggcactgacaattccgt ggtgttgcgggaaatcatcgtccttctccttggtgctcgcctgtgttgcacctggattct gcggggacgtccttctgctacgtccctcggccctcaatccagcgacgttcttcccgc ggcctgctgcggctcgcggccttccgcgtcttcgcttcgcctcagacgagtcgg atcctccttggggcgctccccgc (SEQ ID NO: 38)
SV40-PA	Aactgtttattgcagcttataatggttacaataaagcaatagcatcacaaattcacaaat aaagcatttttctactgcattctagttgtgttgcctcaactcatcaatgtatctta (SEQ ID NO: 39)
3' ITR	Aggaacccctagtgtgaggtggccactccctctctgcgcgtcgtcgtcactgagg ccgggacgacaaagggtcggcgacggcggttggccggcggtcagtgagcg agcgagcgcgcagctgcctgcagg (SEQ ID NO: 34)

[0145] In some embodiments, rAAV vector A04298 comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to one or more of the following nucleotide sequences:

[0146] cctgcaggcagctgcgcgtcgtcgtcactgaggccgcccggcgctggcgaccttggctgcccggcctcagt
gagcgagcgagcgcgcagagaggagtgccaactccactaggggttctGCGGCCGGTCGCGTCTAGTA
CTAGTtacccttgcacccacagctcctctctgtgccttgttccagccatgcgttctctctataaataccgctctggtattgg
ggttggcagctgttgcctcaggagatggttgggtgacatgcggctcctgacaaaacacaaacccctggtgtgtgtggcggtgggt
ggtgtgagtaggggatgaatcaggagggggcAggggacccagggggcaggagccacacaaagtctgtgcgggggtggga
gcgcacatagcaattggaaactgaaagcttatcagacccttctggaaatcagccactgtttataaactgaggccccaccctgcaca
gtaccggggaggaagaggcctgcactagtccagagggaactgaggtcagggcCagctcgccatagacatacatggcaggc
aggcttggccaggatccctccgctgcaggcgtctcctgcctccctcctgcctagagacccccaccctcaagcctggctggtct
ttgcctgagacccaaacctctcactcaagagaataatttaggaacaaggtggttagggccttctgggaacaggccttgaccctta
agaaatgacccaaagtctctccttgacaaaaaggggaccctcaactaaagggaagcctctcttctgctgtctccctgacccactc
ccccacccacaggacgaggagataaccagggtgaaagaggcccgcctgggggctgcagacatgcttgcctgcctggcg
aaggattggcaggcttggcgtcacaggacccccgtggctgactcagggcgagggcctTtgcgggggagctggcctccccgc
ccccacggccacggcgcccttctggcaggacagcgggatcttgcagctgtcaggggaggggagggcggggctgatgtcag
gagggatacaaatagtgccgacggctgggggccctGTCTCCCCTCGCCGCATCCACTCTCCGGCCGG
CCGCCTGCCCGCCGCTCCTCCGTGCGCCCGCCAGCCTCGCCCGGCTAGCgccgccA

CCatggaccgtgtgctcaaaagaagtgtgaaggaagcaacctcctaaaccactgaaaaactacgcggaGGATCGGGC
AGATCTAGACTCCTCGAGGACTTTAGAAACAACAGATACCCTAATCTGCAGCTG
AGGGAAATCGCCGGCCACATTATGGAGTTCAGCCAAGACCAGCACGGCAATAGA
TTCATCCAGCTCAAACCTGGAGAGGGCCACACCCGCCGAGAGGCAGCTGGTGTTC
AACGAGATTCTGCAAGCCGCCTACCAACTGATGGTGGATGTCTTCGGATGCTACG
TCATCCAGAAGTTCTTCGAATTTGGCTCTCTGGAGCAGAAGCTGGCCCTCGCTGA
GAGAATCAGAGGACACGTGCTGAGCCTCGCTCTGCAGATGTATGGATCTAGAGT
GATCAGAAAGGCTCTGGAGTTCATCCCCAGCGACCAGCAGAACGAGATGGTGAG
GGAATCGATGGCCATGTCCTCAAGTGCCTCAAGGATCAGAACGGCAGCTACGT
GGTGGAGAAGTGTATCGAGTGTGTGCAACCCCAATCTCTGCAGTTCATCATCGAC
GCCTTCAAGGGCCAAGTGTTCGCTCTGAGCACCCATCCCTACGGCAGCAGAGTG
ATCGAAAGAATCCTCGAGCATTGTCTGCCCCGACCAGACACTGCCTATTCTGGAGG
AGCTGCACCAGCATAACCGAGCAACTGGTGCAAGACCAGTACGGCAACTACGTCA
TTCAGCACGTGCTCGAACATGGAAGACCCGAAGACAAGAGCAAAATCGTGGCCG
AGATCAGAGGCAATGTCCTCGTGCTGAGCCAGCACAAAGTTTGCTTGCAACGTGGT
GCAGAAGTGCGTGACACACGCTTCTAGAACCGAGAGGGCCGTGCTGATTGACGA
GGTGTGCACCATGAACGACGGCCCCCATTCGCCCTCTACACCATGATGAAGGA
CCAATACGCTAGCTACGTGCTGAGGAAGATGATTGACGTGGCTGAGCCCCGCCA
GAGGAAGATCGTGATGCATAAGATTAGACCCCATATCGCCACACTGAGGAAGTA
CACATACGGCAAGCACATTCTGGCCAAGCTCGAAAAGTACTATATGAAAAATGG
CGTCGATCTGGGCGGCGGAAGTGAGCAAAAACCTGATTAGTGAAGAAGATCTCAT
CCGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAG
GTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAA
TGAATGCAATTGTTGTTGTTaatcaacctctggattacaaaattgtgaagattgactggattcttaactatgttgctc
cttttacgctatgtggatacgtgctttaatgcctttgtatcatgctattgcttccgctatggcttctttctcctcctgtataaatcctggtg
ctgtctctttatgaggagtgtggccggtgtcaggcaacgtggcgtggtgtgcactgtgttgctgacgcaacccccactggttggggc
attgccaccacctgtcagctccttccgggacttccgttcccccctcctattgccacggcggaactcatcgccgctgccttggccgct
gctggacaggggctcggctgttgggcactgacaattccgtggtgtgtcggggaaatcatcgtccttcccttggctgctcgctgtgtg
ccacctggattctgcgcggaacgtccttctgctacgtcccttcggccctcaatccagcggaccttcttcccgcgccctgctgccggct
ctgcggcctcttccgctcttcgcttccctcagacgagtcggatccttcttggccgctcctcccgcaactgtttattgcagcttata
atggttacaataaagcaatagcatcacaatttcacaaataaagcatttttctactgcattctagttgtgtgttgcacaaactcatcaatga
tcttaACGCGGTAACACGTGCGGACCCAACGGCCGCaggaaccctagtgatggagttggccactcc

ctctctgcgcgctcgcctcactgaggccggcgaccaaaggctgcccgcgcccgggctttgcccggcgccctcagtgagc
gagcgcgcgcgagctgcctgcagg (SEQ ID NO: 52)

[0147] AAV vectors

[0148] An "AAV vector" as used herein refers to a vector comprising, consisting essentially of, or consisting of one or more nucleic acid molecules and one or more AAV inverted terminal repeat sequences (ITRs). In some embodiments, the nucleic acid molecule encodes for a CUG-repeat targeting protein and/or composition of the disclosure. Such AAV vectors can be replicated and packaged into infectious viral particles when present in a host cell that provides the functionality of rep and cap gene products; for example, by transfection of the host cell. In some embodiments, AAV vectors contain a promoter, at least one nucleic acid that may encode at least one protein or RNA, and/or an enhancer and/or a terminator within the flanking ITRs that is packaged into the infectious AAV particle. The encapsidated nucleic acid portion may be referred to as the AAV vector genome. Plasmids containing AAV vectors may also contain elements for manufacturing purposes, e.g., antibiotic resistance genes, origin of replication sequences etc., but these are not encapsidated and thus do not form part of the AAV particle.

[0149] In some embodiments, an AAV vector can comprise at least one nucleic acid molecule encoding a CUG-repeat targeting composition of the disclosure. In some embodiments, an AAV vector can comprise at least one regulatory sequence. In some embodiments, an AAV vector can comprise at least one AAV inverted terminal (ITR) sequence. In some embodiments, an AAV vector can comprise a first ITR sequence and a second ITR sequence. In some embodiments, an AAV vector can comprise at least one promoter sequence. In some embodiments, an AAV vector can comprise at least one enhancer sequence. In some embodiments, an AAV vector can comprise at least one polyA sequence. In some embodiments, an AAV vector can comprise at least one linker sequence. In some embodiments, an AAV vector of the disclosure can comprise at least one nuclear localization signals. In some embodiments, an AAV vector of the disclosure can comprise a CUG-repeat targeting PUF or PUMBY protein, peptide, or fragment thereof. In some embodiments, an AAV vector of the disclosure can comprise a Cas protein, peptide, or fragment thereof. In some embodiments, an AAV vector of the disclosure can comprise an endonuclease protein, peptide, or fragment thereof. In some embodiments, an AAV vector of the disclosure can comprise a guide RNA, in some cases a CUG-repeat targeting guide RNA. In some embodiments, AAV vectors of the disclosure can comprise a fusion protein

comprising one or more elements of the disclosure, including, but not limited to, a CUG-repeat targeting protein (such as a Cas, PUF, or PUMBY) and an endonuclease. Optionally, fusion proteins of the AAV vector can further comprise a linker amino acid sequence between the one or more elements of the disclosure.

[0150] In some embodiments, a AAV vector can comprise a first AAV ITR sequence, a promoter sequence, a CUG-repeat targeting composition nucleic acid molecule, a regulatory sequence and a second AAV ITR sequence. In some embodiments, an AAV vector can comprise, in the 5' to 3' direction, a first AAV ITR sequence, a promoter sequence, a transgene nucleic acid molecule, and a second AAV ITR sequence.

[0151] AAV Capsids

[0152] Viral vectors comprising CUG-repeat targeting PUF compositions of the disclosure can comprise a capsid sequence that is isolated or derived from an AAV of serotype AAVPHP.B, AAVrh.74, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 (AAVrh10), AAV11 or AAV12. In some embodiments, the AAV9 capsid is a AAV9 variant comprising one or more mutations and/or modifications. In some embodiment, the AAV9 variant capsid is a myoAAV capsid. In one embodiment, the AAV vector comprises a modified capsid. In one embodiment the AAV vector is an AAV2-Tyr mutant vector. In one embodiment the AAV vector comprises a capsid with a non-tyrosine amino acid at a position that corresponds to a surface-exposed tyrosine residue in position Tyr252, Tyr272, Tyr275, Tyr281, Tyr508, Tyr612, Tyr704, Tyr720, Tyr730 or Tyr673 of wild-type AAV2. See also WO 2008/124724 incorporated herein in its entirety. In some embodiments, the AAV vector comprises an engineered capsid. AAV vectors comprising engineered capsids include without limitation, AAV2.7m8, AAV9.7m8, AAV2 2tYF, and AAV8 Y733F). In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant (rAAV). In some embodiments, the viral vector is self-complementary (scAAV). In some embodiments, the rAAV is a single-stranded AAV (ssAAV).

[0153] AAV9 and other muscle-specific capsids such as MyoAAV4A are disclosed herein for use in combination with the hybrid regulatory sequences and CUG-repeat targeting PUF compositions of the disclosure. The disclosure provides modified AAV capsid proteins. Modified AAV capsid proteins of the disclosure can comprise a peptide insertion targeting integrin. In some embodiments, the integrin-targeting peptide comprises an RGD-motif. RGD sequences are known in the art, and include, for example, the motif RGDXXXX which

may be inserted in an AAV viral vector for targeting via the integrin class of receptors. *see, e.g., Michelfelder et al., PLoS One. 2009; 4(4): e5122* which is incorporated herein by reference in its entirety for example of RGD sequences that may be used in modified AAV capsids described herein. RGD-motif peptide insertions into VR8 of AAV9 has been shown to increase mouse muscle transduction (See Weinmann et al. Nature Communications, 11:5432 which is incorporated herein by reference in its entirety). In some embodiments, the RGD peptide comprises a subsequence Y or F amino acid to produce an RGDY or RGDF motif. RGDY or RGDF motifs have been demonstrated to produce enhanced muscle transduction in non-human primates (NHP) (See Tabebordbar et al. Cell, 184, 19, 2021, 4919-4938 which is incorporated herein by reference in its entirety). In some embodiments, the RGD sequence comprises RGDGLS (SEQ ID NO: 140). In some embodiments, the RGD sequence comprises RGDGL (SEQ ID NO: 141). In some embodiments, the RGD sequence comprises RGDLP (SEQ ID NO: 142), SNSRGDYNL (SEQ ID NO: 143), ENRRGDFNNT (SEQ ID NO: 144), SRGDYNL (SEQ ID NO: 145), RGDYNL (SEQ ID NO: 146), RGDLP (SEQ ID NO: 147), or RGDYVGL (SEQ ID NO: 148).

[0154] In some embodiments, an AAV capsid of the disclosure comprises LBV201. LBV201 can comprise an RGD peptide insertion. In some embodiments, the RGD peptide insertion of LBV201 can comprise the amino acid sequence set forth in SEQ ID NO: 40 or SEQ ID NO: 141.

[0155] In some embodiments, a nucleic acid sequence encoding an LBV201 capsid comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCA
 TTCGCGAGTGGTGGGACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGC
 AAAAGCAGGACGACGGCCGGGTCTGGTGCTTCCTGGCTACAAGTACCTCGGAC
 CCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGATGCAGCGGCCC
 TCGAGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACC
 TGCGGTATAACCACGCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGT
 CTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAGGCCAAGAAGAGGGTTCTCG
 AACCTTTTGGTCTGGTTGAGGAAGGTGCTAAGACGGCTCCTGGAAAGAAACGTC
 CGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCGGGCATTGGCAAGACAG
 GCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAGT

CAGTCCCCGACCCACAACCTCTCGGAGAACCTCCAGCAACCCCCGCTGCTGTGG
GACCTACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAG
GCGCCGACGGAGTGGGTAATGCCTCAGGAAATTGGCATTGCGATTCCACATGGC
TGGGCGACAGAGTCATCACCACCAGCACCCGAACATGGGCCTTGCCACCTATA
ACAACCACCTCTACAAGCAAATCTCCAGTGCTTCAGGGGCCAGCAACGACAACC
ACTACTTCGGCTACAGCACCCCCTGGGGGTATTTTGATTTC AACAGATTCCACTG
CCATTTCTCACCACGTGACTGGCAGCGACTCATCAACAACAATTGGGGATTCCGG
CCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAAGTCAAGGAGGTCACGACG
AATGATGGCGTCACGACCATCGCTAATAACCTTACCAGCACGGTTCAAGTCTTCT
CGGACTCGGAGTACCAGTTGCCGTACGTCCTCGGCTCTGCGCACCAAGGGCTGCCT
CCCTCCGTTCCCGGCGGACGTGTTTCATGATTCCGCAGTACGGCTACCTAACGCTC
AACAATGGCAGCCAGGCAGTGGGACGGTCATCCTTTTACTGCCTGGAATATTTCC
CATCGCAGATGCTGAGAACGGGCAATAACTTTACCTTCAGCTACACCTTCGAGGA
CGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAA
TCCTCTCATCGACCAGTACCTGTATTACCTGAACAGAACTCAGAATCAGTCCGGA
AGTGCCCCAAAACAAGGACTTGCTGTTTAGCCGGGGGTCTCCAGCTGGCATGTCTG
TTCAGCCCCAAAACACTGGCTACCTGGACCCTGTTACCGGCAGCAGCGCTTTCTAA
AACAAAAACAGACAACAACAACAGCAACTTTACCTGGACTGGTGCTTCAAAATA
TAACCTTAATGGGCGTGAATCTATAATCAACCCTGGCACTGCTATGGCCTCACAC
AAAGACGACgAAGACAAGTTCTTTCCCATGAGCGGTGTCATGATTTTTGGAAAGG
AGAGCGCCGGAGCTTCAAACACTGCATTGGACAATGTCATGATCACAGACGAAG
AGGAAATCAAAGCCACTAACCCCGTGGCCACCGAAAGATTTGGGACTGTGGCAG
TCAATCTCCAGAGCAGCcgaggtgacctcggtctttctAGCACAGACCCTGCGACCGGAGATG
TGCATGTTATGGGAGCCTTACCTGGAATGGTGTGGCAAGACAGAGACGTATACC
TGCAGGGTCCTATTTGGGCCAAAATTCCTCACACGGATGGACACTTTCACCCGTC
TCCTCTCATGGGCGGCTTTGGACTTAAGCACCCGCCTCCTCAGATCCTCATCAA
AACACGCCTGTTCTGCGAATCCTCCGGCAGAGTTTTCGGCTACAAAGTTTGCTT
CATTCATCACCCAGTATTCCACAGGACAAGTGAGCGTGGAGATTGAATGGGAGC
TGCAGAAAGAAAACAGCAAACGCTGGAATCCCGAAGTGCAGTATACATCTAACT
ATGCAAAATCTGCCAACGTTGATTTCACTGTGGACAACAATGGACTTTTATACTGA
GCCTCGCCCCATTGGCACCCGTTACCTCACCCGTCCCCTGTAA (SEQ ID NO: 53).

[0156] In some embodiments, a LBV201 capsid comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDGRGLVLPGYKYL
 PFNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTS
 FGGNLGRAVFQAKKRVLEPFGLVEEGAKTAPGKKRPVEQSPQEPDSSSGIGKTGQQP
 AKKRLNFGQTGDSESVDPDQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGV
 GNASGNWHCDSTWLGDREVTTSTRTWALPTYNNHLYKQISSASGASNDNHYFGYST
 PWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTTNDGVTTIA
 NNLSTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRS
 SFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLNR
 QNQSGSAQNKDLLFSRGSPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNSNFTWT
 GASKYNLNGRESIINPGTAMASHKDDDEKFFPMSGVMIFGKESAGASNTALDNVMT
 DEEEIKATNPVATERFGTVAVNLQSSRGDLGLSSTD PATGDVHVMGALPGMVWQD
 RDVYLQGPWAKIPHTDGHFHPSPMLGGFGLKHPPQILIKNTPVPANPPAEFSATKF
 ASFITQYSTGQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTE
 PRPIGTRYLTRPL (SEQ ID NO: 54).

[0157] In some embodiments, a nucleic acid sequence encoding a MyoAAV4 capsid comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

atggctgccgatggttatcttcagattggctcgaggacaaccttagtgaaggaattcgcgagtggtgggcttgaacctggagcccc
 tcaaccaaggcaaatcaacaacatcaagacaacgctcgaggtctgtgcttcgggttacaataccttggacccggcaacggactc
 gacaagggggagccggtaacgcagcagacgcggcgccctcgagcacgacaaggcctacgaccagcagctcaaggccggag
 acaacccgtacctcaagtacaaccacgccgacgccgagttccaggagcggctcaaagaagatacgtctttgggggcaacctcggg
 cgagcagttcttcaggccaaaaagaggcttctgaacctcttggtctggttgaggaagcggctaagacggctcctggaaagaagagg
 cctgtagagcagtcctcaggaaccgactcctccggttattggcaaatcgggtgcacagcccgtaaaaagagactcaatttcg
 gtcagactggcgacacagagtcagtcacagacctcaaccaatcgagaaacctccgcagccccctcaggtgtgggatctcttaca
 tggcttcaggtggtggcgaccagtgagcagacaataacgaagtgccgatggagtggttagtctcgggaaattggcattgcgattc
 ccaatggctgggggacagagtcacaccaccagcaccgaacctgggcccctgccacctaacaatacacctctacaagcaaatctc
 caacagcacatctggaggatctcaaatgacaacgcctacttcggctacagacccccctgggggtattttgacttcaacagattccactg
 ccacttctaccacgtgactggcagcagactcatcaacaacaactggggattccggcctaagcgactcaactcaagctcttcaacattc
 aggtcaaaagaggttacggacaacaatggagtcaagaccatcgccaataaccttaccagcaggtccaggtcttcacggactcagact

atcagctcccgtagctgctgggtcggctcacgagggctgctcccgccgtccagcggacgtttcatgattcctcagtagcgggtat
ctgacgcttaatgatggaagccaggccgtgggtcgctcttttactgcctggaatatttccgctcgaaatgctaagaacgggtaac
aactccagttcagctacgagtttgagaacgtacctttccatagcagctacgctcacagccaaagcctggaccgactaatgaatccactc
atcgaccaatacttgtagtctctcaaagactattaacgggtctggacagaatcaacaaacgctaaaattcagtggtggccgaccagc
aactggctgtccagggaagaaactacatacctggaccagctaccgacaacacgtgtctcaaccactgtgactcaaaacaacac
agcgaatttgcttggcctggagcttcttctgggtctcaatggacgtaatagcttgatgaatcctggacctgctatggccagccacaaa
gaaggagaggaccgtttcttcttcttctggatcttaattttggcaaacaaggaaactggaagagacaacgtggatgaggacaaaagtc
atgataaccaacgaagaagaaataaaactactaaccggtagcaacggagtcctatggacaagtggccacaaaccaccagTCC
AATagtCGGGGTGACTACAATTCATCgcacaggcgcagaccggctgggttcaaaaccaaggaataactccg
ggtagtgggttggcaggacagagatgtgtacctgcaaggaccatttgggcaaaaattcctcacacggcaggcaactttcacccttctcc
gctgatgggaggggttggaaatgaagcaccgcctcctcagatcctcataaaaacacacctgtacctgcggatcctccaacggccttc
aacaaggacaagctgaactctttcatcaccagtagtattctactggccaagtgcagctggagatcgagtgaggagctgcagaaggaaaaac
agcaagcgtggaacccggagatccagtagcacttccaactattacaagtctaataatgtgaatttctgttaatactgaaggtgtatata
gtgaaccccgcccatggcaccagatactgactcgtaatctgtaa (SEQ ID NO: 55).

[0158] In some embodiments a RepCap plasmid encoding LBV201 comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the sequence set forth in SEQ ID NO: 149.

[0159] In some embodiments, a MyoAAV4 capsid comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

MAADGYLPDWLEDNLSEGIREWALKPGAPQPKANQQHQDNARGLVLPGYKYL
PGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDT
SFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQ
PAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSMTMASGGGAPVADNNEGADGV
GSSSGNWHCDSQWLGDREVITSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGY
STPWGYFDFNRFHCHFSRWDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKT
IANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVG
RSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQLDRLMNPLIDQYLYYLS
KTINGSGQNQQLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWP
GASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVM
TNEEEIKTTNPVATESYGQVATNHQSNRSGDYNSLAQAQTGWVQNQGILPGMVWQ
DRDVYLQGPWAKIPHTDGNFHPSPLMGGFGMKHPPQILIKNTPVPADPPTAFNKD

KLNSFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYS
EPRPIGTRYLTRNL (SEQ ID NO: 56).

[0160] In some embodiments, a nucleic acid sequence encoding an AAV9 capsid comprises, consists essentially of, or consists of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

atggctgccgatggttatcttccagattggctcgaggacaacctagtgagggaattcgcgagtggtgggcttgaacctggagcccc
tcaaccaaggcaaatcaacaacatcaagacaacgctcgaggtcttgcttcgggttacaaataacctggaccgggcaacggactc
gacaagggggagccggtcaacgcagcagacgcgggcgccctcgagcacgacaaggcctacgaccagcagctcaaggccggag
acaacctgtacctcaagtacaaccacgcccagccgaggttccaggagcggtcaaagaagatactctttgggggcaacctcggg
cgagcagcttccaggccaaaaagaggcttctgaacctcttggtctggttgaggagcggttaagacggctcctggaaagaagg
cctgtagagcagctcctcaggaacggactcctccggggtattggcaaatcgggtgcacagcccgtaaaaagagactcaatttcg
gtcagactggcgacacagagtcagtcacgacctcaaccaatcggagaacctcccgagccccctcaggtgtgggatctcttaca
tggcttcaggtggtggcgaccagtgaggagacaataacgaaggtgccgatggagtggttagtctcgggaaattggcattgcgattc
ccaatggctgggggacagagtcaccaccagcaccgaacctgggcccctgccacctacaacaatcacctctacaagcaaatctc
caacagcacatctggagatcttcaaatgacaacgcctacttcgggtacagacccccctgggggtattttgacttcaacagattccactg
ccacttctaccacgtgactggcagcgactcatcaacaactggggattccggcctaagcgactcaactcaagctcttcaacattc
aggtcaaagaggttacggacaacaatggagtcgaagaccatcgccaataaccttaccagcacggtccaggtcttcacggactcagact
atcagctcccgtacgtgctcgggtcgggtcacgagggctgcctcccgccgttccagcggacgtttcatgattcctcagtcagggtat
ctgacgcttaatatgatgaagccaggccgtgggtcgttcgtcctttactgcctggaatatttccgctcgcaaatgctaagaacgggtaac
aacttccagttcagctacgagtttgagaacgtaccttccatagcagctacgctcacagccaaagcctggaccgactaatgaatccactc
atcgaccaatactgtactatcttcaaaactattaacgggtctggacagaatcaacaacgctaaaattcagtggtggccggaccagc
aacatggctgtccagggaagaaactacatacctggaccagctaccgacaacaacgtgtctcaaccactgtgactcaaaacaacaac
agcgaatttcttggcctggagcttcttctgggctctcaatggacgtaatagtctgatgaatcctggacctgctatggccagccacaaa
gaaggagaggaccgtttcttcttctgttgatctttaattttggcaacaaggaactggaagagacaacgtggatgaggacaaaagtc
atgataaccaacgaagaagaaattaaaactactaaccggtagcaacggagtcctatggacaagtggccacaaaccaccagagtc
ccaagcacaggcgagaccggctgggttcaaaaccaaggaatacttccgggtatggtttggcaggacagagatgtgtacctgcaag
gaccttggggccaaaattcctcacacggacggcaacttcacccttctccgctgatgggaggggttggaatgaagcacccgcctcct
cagatcctcatcaaaaacacacctgtacctgcggatcctcaacggccttcaacaaggacaagctgaactcttcatcaccagttattct
actggccaagtacgctggagatcgagtgaggagctgcagaaggaaaacagcaagcgctggaacccggagatccagtagacttcca
actattacaagtctaataatgtgaatttctgttaatactgaagggtgtatatagtgaacccgccccattggcaccagatacctgactcgt
aatctgtaa (SEQ ID NO: 57).

[0161] In some embodiments, an AAV9 capsid comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to:

MAADGYLPDWLEDNLSEGIREWWALKPGAPQPKANQQHQDNARGLVLPGYKYLGP
PGNGLDKGEPVNAADAAALEHDKAYDQQLKAGDNPYLKYNHADADEFQERLKEDT
SFGGNLGRAVFQAKKRLLEPLGLVEEAAKTAPGKKRPVEQSPQEPDSSAGIGKSGAQ
PAKKRLNFGQTGDTESVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGV
GSSSGNWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISNSTSGSSNDNAYFGY
STPWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKT
IANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVG
RSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHQSGLDRLMNPIDQYLYLS
KTINGSGQNQQTLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWP
GASSWALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVM
TNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNQGILPGMVWQDRDVYLQ
GPIWAKIPHTDGNFHPSPLMGGFGMKHPPQILIKNTPVPADPPTAFNKDKLNSFITQ
YSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRIGTR
YLTRNL (SEQ ID NO: 58).

[0162] *Signal Sequences*

[0163] In some embodiments, a target RNA-binding protein of the disclosure comprises a signal sequence. In some embodiments, a target RNA-binding protein comprises one or more signal sequences. In some embodiments, the signal sequence is a nuclear localization sequence (NLS), a nuclear export signal (NES), or a combination thereof. In some embodiments, the signal sequence comprises one or more nuclear localization sequences (NLSs). In some embodiments, one or more NLS sequence comprises a sequence listed in Table 8. In some embodiments, the NLS signal sequence is a SV40 NLS signal sequence. In some embodiments, the SV40 NLS signal sequence is PKKKRKV (SEQ ID NO: 437).

[0164] Table 8: Nuclear Localization Sequences of the disclosure

Name	Amino acid Sequence	SEQ ID NO:
SV40-NLS	PKKKRKV	59
human H2B-NLS	GKKRKRSRK	60
yeast H2B-NLS	GKKRSKV	61
human p53-NLS	KRALPNNTSSSPQPKKKP	62
human-cmyc-NLS	PAAKRVKLD	63
human pRB-NLS	KRSAEGSNPPKPLKKLR	64
human Nucleoplasmin-NLS	KRPAATKKAGQAKKKKLDK	65

Human pRB-NLS (extended version)	DRVLRKRSAGSNPPKPLKKLR	66
----------------------------------	-----------------------	----

[0165] In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 59. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 60. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 61. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 62. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 63. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 64. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 65. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 66.

[0166] In some embodiments, the signal sequence comprises one or more NES sequences. In some embodiments, the one or more NES sequence comprises a sequence listed in Table 9.

[0167] Table 9: Nuclear Export Sequences of the disclosure

Name	Amino acid Sequence	SEQ ID NO:
HIV REV NES	LPPLERLTLD	67
Human PKI NES	LALKLAGLDI	68

[0168] In some embodiments, the NES comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 67. In some embodiments, the NLS comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in SEQ ID NO: 68.

[0169] In some embodiments, a target RNA-binding fusion protein of the disclosure comprises a tag sequence. In some embodiments, the tag sequence is a FLAG tag. In some embodiments, the FLAG tag sequence is DYKDDDDK (SEQ ID NO: 69).

[0170] *Linker Sequences*

[0171] In some embodiments, a target RNA-binding protein comprises a linker sequence. In some embodiments, the linker sequence may comprise or consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of amino acids in between. In some embodiments, the linker sequence comprises a linker sequence listed in Table 10.

[0172] Table 10. Linker Sequences of the disclosure

Linker Sequence (amino acid)	SEQ ID NO:
GGS	70
VDTANGS	71
VDTGNGS	72
SGSETPGTSESATPES	73
GGGGSGGGGS	74
GGGGSGGGGS	75
GGGGSGGGGS	76
EAAAKEAAK	77
EAAAKEAAAKEAAK	78
EAAAKEAAAKEAAAKEAAK	79
APAPAPAP	80
APAPAPAPAPAP	81
APAPAPAPAPAPAPAPAP	82
GGGGSEAAK	83
EAAAKGGGGS	84
GGGGSGGGGSEAAAKEAAK	85
EAAAKEAAKGGGGSGGGGS	86

RQTSPDPCPQLPLVPR	87
VDTGNWF	88
VDTANGSVDTGNNGS	89
ARNVEERLCL	90
AIELNPSNA	91
ICGSRNL	92
VLATDMSKH	93
FLRELPEP	94
LIPKDQYYC	95
AEAAAKEAAKA	96
AEAAAKEAAAKEAAKA	97
AEAAAKEAAAKEAAAKEAAA KA	98
YVEFEGEQGVDEGGVSGGGS	99
GSRNLDFQALEETTEYDGGY	100
ASSTSPVEISEWLDQKLTKSDR PEL	101
VNQCRREQSEDSTFYLG	102
AVSPLLLLTTNSSEGLSMGNY	103
LDEAYPGKKLLPDDPYEKACQ	104
SAAAATPAVRTVPQYKYAAGV RNPQQHLNAQPQVTMQQPAV HVQGQEPL	105
GGGGSEAAKGGGGS	106
GGGGSEAAKGGGGSEAAK GGGS	107
GGGGSSGSETGGTSESATGESG GGGS	108
SGSETPGTSESATPES	109

[0173] In some embodiments, a linker sequence comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to the amino acid sequence set forth in any one of SEQ ID NO: 70 – SEQ ID NO: 109.

[0174] *Promoter Sequences*

[0175] In embodiments, CUG targeting compositions of the disclosure comprise a hybrid regulatory sequence or combination of sequences disclosed herein. In some embodiments, CUG targeting compositions comprise a muscle-specific hybrid regulatory combination which is more selective for skeletal muscle than cardiac muscle in a manner such that there is a higher level of expression in skeletal muscle than in cardiac muscle. In some embodiments, the CUG targeting compositions comprise a muscle-specific hybrid regulatory combination which provides equally or balanced levels of expression for skeletal muscle and cardiac muscle.

[0176] PUF and PUMBY Architecture

[0177] The unique RNA recognition mode of PUF proteins (named for *Drosophila* Pumilio and *C. elegans* fem-3 binding factor) that are involved in mediating mRNA stability and translation are well known in the art. The PUF domain of human Pumilio1, also known in the art, binds tightly to cognate RNA sequences and its specificity can be modified. It contains eight PUF modules that recognize eight consecutive RNA bases with each module recognizing a single base. Since two amino acid side chains in each module recognize the Watson-Crick edge of the corresponding base and determine the specificity of that module, a PUF protein can be designed to specifically bind most 8 to 16-nt RNA. Wang *et al.*, *Nat Methods*. 2009; 6(11): 825-830. See also WO2012/068627 which is incorporated by reference herein in its entirety.

[0178] The modular nature of the PUF-RNA interaction has been used to rationally engineer the binding specificity of PUF domains (Cheong, C. G. & Hall, T. M. (2006) PNAS 103: 13635-13639; Wang, X. et al (2002) Cell 110: 501-512). However, only the successful design of PUF domains with modules that recognize adenine, guanine or uracil have been reported prior to the teachings of WO2012/06827 *supra*. While the wild-type PumHD does not bind cytosine (C), molecular engineering has shown that some of the Pum units can be mutated to bind C with good yield and specificity. See e.g., Dong, S. et al. Specific and modular binding code for cytosine recognition in Pumilio/FPF (PUF) RNA-binding domains, *The Journal of biological chemistry* 286, 26732-26742 (2011). Accordingly, PumHD is a modified version of the WT Pumilio protein that exhibits programmable binding to arbitrary 8-base sequences of RNA. Each of the eight units of PumHD can bind to all four RNA bases, and the RNA bases flanking the target sequence do not affect binding. See also the following for art-recognized RNA-binding rules of PUF design: Filipovska A, Razif MF, Nygård KK, & Rackham O. A universal code for RNA recognition by PUF proteins. *Nature chemical biology*, 7(7), 425-427 (2011); Filipovska A, & Rackham O. Modular recognition of nucleic acids by PUF, TALE and PPR proteins. *Molecular BioSystems*, 8(3), 699-708 (2012); Abil Z, Denard CA, & Zhao H. Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. *Journal of biological engineering*, 8(1), 7 (2014); Zhao Y, Mao M, Zhang W, Wang J, Li H, Yang Y, Wang Z, & Wu J. Expanding RNA binding specificity and affinity of engineered PUF domains. *Nucleic Acids Research*, 46(9), 4771-4782 (2018); Shinoda K, Tsuji S, Futaki S, & Imanishi M. Nested PUF Proteins: Extending Target RNA Elements for Gene Regulation. *ChemBioChem*, 19(2), 171-176

(2018); Koh YY, Wang Y, Qiu C, Opperman L, Gross L, Tanaka Hall TM, & Wickens M. Stacking Interactions in PUF-RNA Complexes. RNA, 17(4), 718-727 (2011).

[0179] As such, it is well known in the art that human PUM1 (1186 amino acids) contains an RNA-binding domain (RBD) in the C-terminus of the protein (also known as Pumilio homology domain PUM-HD amino acid 828-amino acid 1175) and that PUFs are based on the RBD of human PUM1. There are 8 structural modules of 36 amino acids (except module 7 which has 43 amino acids) for RNA binding and flanking N- and C- terminal regions important for protein structure and stability. Within each module, amino acids 12, 13, and 16 are important for RNA binding with 12 and 16 responsible for RNA base recognition. Amino acid 13 stacks with RNA bases and can be modified to tune specificity and affinity.

Alternatively, the PUF design may maintain amino acid 13 as human PUM1's native residue. In some embodiments of the PUF(CUG) or PUMBY(CUG) compositions disclosed herein, amino acid 13 (for stacking) will be engineered with an H and in other embodiments, will be engineered with a Y. In some embodiments, stacking residues may be modified to improve binding and specificity. Recognition occurs in reverse orientation as N- to C-terminal PUF recognizes 3' to 5' RNA. Accordingly, PUF engineering of 8 modules (8PUF), as known in the art, mimics a human protein. An exemplary 8-mer RNA recognition (8PUF) would be designed as follows: R1'-R1-R2-R3-R4-R5-R6-R7-R8-R8'. In one embodiment, an 8PUF is used as the RBD. In another embodiment, a variation of the 8PUF design is used to create a 14-mer RNA recognition (14PUF) RBD, 15-mer RNA recognition (15PUF) RBD, or a 16-mer RNA recognition (16PUF) RBD. In another embodiment, the PUF can be engineered to comprise a 4-mer, 5-mer, 6-mer, 7-mer, 8-mer, 9-mer, 10-mer, 11-mer, 12-mer, 13-mer, 14-mer, 15-mer, 16-mer, 24-mer, 30-mer, 36-mer, or any number of modules between. Shinoda et al., 2018; Criscuolo et al., 2020. Repeats 1-8 of wild type human PUM1 are provided herewith at SEQ ID NOS: 110-117, respectively. The nucleic acid sequence encoding the PUF domain from human PUM1 is SEQ ID NO: 118 and the amino acid sequence of the PUF domain from human PUM1 amino acids 828-1176 is SEQ ID NO: 119. See also US Patent 9,580,714 which is incorporated herein in its entirety.

[0180] PUF or PUMBY RNA-binding sequences disclosed in WO 2022/119979 are incorporated herein by reference in its entirety.

[0181] In some embodiments, a nucleic acid sequence encoding a CUG repeat targeting PUF sequence comprises, consists essentially of, or consists of a nucleic acid sequence at

least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to in SEQ ID NO: 44, SEQ ID NO: 45, or SEQ ID NO: 46.

[0182] In some embodiments, a CUG repeat targeting PUF sequence comprises, consists essentially of, or consists of an amino acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to in SEQ ID NO: 133, SEQ ID NO: 134, or SEQ ID NO: 135.

[0183] In some embodiments, a PUF or PUMBY RNA-binding sequence is capable of binding a toxic target CUG repeat sequence comprising UGCUGCUG (SEQ ID NO: 120).

[0184] In one embodiment, the target RNA sequence is selected from the group consisting of UGCUGCUGCUGCUG (SEQ ID NO: 121), UGCUGCUGCUGCUGC (SEQ ID NO: 122), and UGCUGCUGCUGCUGCU (SEQ ID NO: 123).

[0185] In one embodiment, the target RNA sequence is selected from the group consisting of CUGCUGCU (SEQ ID NO: 124), CUGCUGCUGCUGCU (SEQ ID NO: 125), CUGCUGCUGCUGCUG (SEQ ID NO: 126), and CUGCUGCUGCUGCUGC (SEQ ID NO: 127).

[0186] In one embodiment, the target RNA sequence is selected from the group consisting of GCUGCUGC (SEQ ID NO: 128), GCUGCUGCUGCUGC (SEQ ID NO: 129), GCUGCUGCUGCUGCU (SEQ ID NO: 130), and GCUGCUGCUGCUGCUG (SEQ ID NO: 131).

[0187] In some embodiments the RNA-binding protein or RNA-binding portion thereof is a PUMBY (Pumilio-based assembly) protein. RNA-binding protein PumHD, which has been widely used in native and modified form for targeting RNA, has been engineered into a protein architecture designed to yield a set of four canonical protein modules, each of which targets one RNA base. These modules (*i.e.*, Pumby, for Pumilio-based assembly) are concatenated in chains of varying composition and length, to bind desired target RNAs. In essence, PUMBY is a more simple and modular form of PumHD, in which a single protein unit of PumHD is concatenated into arrays of arbitrary size and binding sequence specificity. The specificity of such Pumby–RNA interactions is high, with undetectable binding of a Pumby chain to RNA sequences that bear three or more mismatches from the target sequence. *Katarzyna et al., PNAS, 2016; 113(19): E2579-E2588*. See also US 2016/0238593 which is incorporated by reference herein in its entirety.

[0188] In some embodiments of the compositions of the disclosure, the first RNA binding protein comprises a Pumilio and FBF (PUF) protein. In some embodiments, the first RNA binding protein comprises a Pumilio-based assembly (PUMBY) protein.

[0189] In some embodiments, the fusion protein disclosed herein comprises a linker between the at least two RNA-binding polypeptides. In some embodiments, the linker is a peptide linker. In one embodiment, the linker is VDTANGS (SEQ ID NO: 71). In some embodiments, the peptide linker comprises one or more repeats of the tri-peptide GGS. In other embodiments, the linker is a non-peptide linker. In some embodiments, the non-peptide linker comprises polyethylene glycol (PEG), polypropylene glycol (PPG), co-poly(ethylene/propylene) glycol, polyoxyethylene (POE), polyurethane, polyphosphazene, polysaccharides, dextran, polyvinyl alcohol, polyvinylpyrrolidones, polyvinyl ethyl ether, polyacryl amide, polyacrylate, polycyanoacrylates, lipid polymers, chitins, hyaluronic acid, heparin, or an alkyl linker.

[0190] In some embodiments, nucleic acid sequences encoding PUF proteins of the disclosure are codon optimized nucleic acid sequences. In some embodiments, the codon optimized sequence encoding a PUF protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased expression in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

[0191] In some embodiments, nucleic acid sequences encoding PUF proteins of the disclosure are codon optimized nucleic acid sequences. In some embodiments, the codon optimized sequence encoding a PUF protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased translation in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

[0192] In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein exhibits increased stability. In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein exhibits increased stability through increased resistance to hydrolysis. In some embodiments, the codon optimized sequence encoding a PUF protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased stability relative to a wild-type or non-codon optimized nucleic acid sequence. In some embodiments, the codon optimized sequence encoding a PUF protein exhibits at least 5%, at least 10%, at

least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased resistance to hydrolysis in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

[0193] In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein can comprise no donor splice sites. In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein can comprise no more than about one, or about two, or about three, or about four, or about five, or about six, or about seven, or about eight, or about nine, or about ten donor splice sites. In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein comprises at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten fewer donor splice sites as compared to a non-codon optimized nucleic acid sequence encoding the PUF protein.

[0194] Without wishing to be bound by theory, the removal of donor splice sites in the codon optimized nucleic acid sequence can unexpectedly and unpredictably increase expression of the PUF protein *in vivo*, as cryptic splicing is prevented. Moreover, cryptic splicing may vary between different subjects, meaning that the expression level of the PUF protein comprising donor splice sites may unpredictably vary between different subjects. Such unpredictability is unacceptable in the context of human therapy. Accordingly, the codon optimized nucleic acid sequences, which lacks donor splice sites, unexpectedly and surprisingly allows for increased expression of the PUF protein in human subjects and regularizes expression of the PUF protein across different human subjects.

[0195] In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein, can have a GC content that differs from the GC content of the non-codon optimized nucleic acid sequence encoding the PUF protein. In some embodiments, the GC content of a codon optimized nucleic acid sequence encoding a PUF protein is more evenly distributed across the entire nucleic acid sequence, as compared to the non-codon optimized nucleic acid sequence encoding the PUF protein.

[0196] Without wishing to be bound by theory, by more evenly distributing the GC content across the entire nucleic acid sequence, the codon optimized nucleic acid sequence exhibits a more uniform melting temperature (“*T_m*”) across the length of the transcript. The uniformity of melting temperature results unexpectedly in increased expression of the codon optimized nucleic acid in a human subject, as transcription and/or translation of the nucleic acid sequence occurs with less stalling of the polymerase and/or ribosome.

[0197] In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein can have fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence encoding the PUF protein. In some embodiments, a codon optimized nucleic acid sequence encoding a PUF protein can have at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least ten fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence the PUF protein.

[0198] Without wishing to be bound by theory, by having fewer repressive microRNA target binding sites, the codon optimized nucleic acid sequence encoding a PUF protein unexpectedly exhibits increased expression in a human subject.

Vectors

[0199] One type of vector is a "plasmid," which refers to a circular double stranded DNA loop into which additional DNA segments can be inserted, such as by standard molecular cloning techniques. Another type of vector is a viral vector, wherein virally -derived DNA or RNA sequences are present in the vector for packaging into a virus (e.g., retroviruses, replication defective retroviruses, adenoviruses, replication defective adenoviruses, and adeno-associated viruses). Viral vectors also include polynucleotides carried by a virus for transfection into a host cell. In some embodiments, the vector is a lentivirus (such as an integration-deficient lentiviral vector) or adeno-associated viral (AAV) vector. Vectors are capable of autonomous replication in a host cell into which they are introduced such as e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors and other vectors such as, e.g., non-episomal mammalian vectors, are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome.

[0200] In some embodiments, vectors such as e.g., expression vectors, are capable of directing the expression of genes to which they are operatively-linked. Common expression vectors are often in the form of plasmids. In some embodiments, recombinant expression vectors comprise a nucleic acid provided herein such as e.g., a guide RNA which can be expressed from a DNA sequence, and a nucleic acid encoding a Cas 13d protein, in a form suitable for expression of a protein in a host cell. Recombinant expression vectors include one or more regulatory elements, which may be selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the

nucleotide sequence of interest is linked to the regulatory element(s) in a manner that allows for expression of the nucleotide sequence such as e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell. Certain embodiments of a vector depend on factors such as the choice of the host cell to be transformed, and the level of expression desired. A vector can be introduced into host cells to thereby produce transcripts, proteins, or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein such as, e.g., CRISPR transcripts, proteins, enzymes, mutant forms thereof, fusion proteins thereof, etc.

[0201] In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a viral vector. In some embodiments, the viral vector comprises a sequence isolated or derived from a retrovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from a lentivirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adenovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant. In some embodiments, the viral vector is self-complementary.

[0202] The term "adeno-associated virus" or "AAV" as used herein refers to a member of the class of viruses associated with this name and belonging to the genus Dependoparvovirus, family Parvoviridae. Adeno-associated virus is a single-stranded DNA virus that grows in cells in which certain functions are provided by a co-infecting helper virus. General information and reviews of AAV can be found in, for example, Carter, 1989, Handbook of Parvoviruses, Vol. 1, pp. 169- 228, and Berns, 1990, Virology, pp. 1743-1764, Raven Press, (New York). It is fully expected that the same principles described in these reviews will be applicable to additional AAV serotypes characterized after the publication dates of the reviews because it is well known that the various serotypes are quite closely related, both structurally and functionally, even at the genetic level. (See, for example, Blacklowe, 1988, pp. 165-174 of Parvoviruses and Human Disease, J. R. Pattison, ed.; and Rose, Comprehensive Virology 3: 1-61 (1974)). For example, all AAV serotypes apparently exhibit very similar replication properties mediated by homologous rep genes; and all bear three related capsid proteins such as those expressed in AAV2. The degree of relatedness is further suggested by heteroduplex analysis which reveals extensive cross-hybridization between serotypes along the length of the genome; and the presence of analogous self-annealing

segments at the termini that correspond to "inverted terminal repeat sequences" (ITRs). The similar infectivity patterns also suggest that the replication functions in each serotype are under similar regulatory control. Multiple serotypes of this virus are known to be suitable for gene delivery; all known serotypes can infect cells from various tissue types.

[0203] AAV possesses unique features that make it attractive as a vector for delivering foreign DNA to cells, for example, in gene therapy. AAV infection of cells in culture is noncytopathic, and natural infection of humans and other animals is silent and asymptomatic. Moreover, AAV infects many mammalian cells allowing the possibility of targeting many different tissues in vivo. Moreover, AAV transduces slowly dividing and non-dividing cells, and can persist essentially for the lifetime of those cells as a transcriptionally active nuclear episome (extrachromosomal element). The AAV proviral genome is inserted as cloned DNA in plasmids, which makes construction of recombinant genomes feasible. Furthermore, because the signals directing AAV replication and genome encapsidation are contained within the ITRs of the AAV genome, some or all of the internal approximately 4.3 kb of the genome (encoding replication and structural capsid proteins, rep-cap) may be replaced with foreign DNA to generate AAV vectors. The rep and cap proteins may be provided in trans. Another significant feature of AAV is that it is an extremely stable and hearty virus. It easily withstands the conditions used to inactivate adenovirus (56° to 65°C for several hours), making cold preservation of AAV less critical. AAV may even be lyophilized. Finally, AAV-infected cells are not resistant to superinfection.

[0204] AAV (AAV or AAV vector) genomes of the invention comprise, consist essentially of, or consist of a nucleic acid molecule encoding a CUG-repeat targeting composition (such as a PUF, PUMBY, or RNA-guided protein) and one or more AAV ITRs flanking the nucleic acid molecule. Production of pseudotyped AAV is disclosed in, for example, WO2001083692. Other types of AAV variants, for example rAAV with capsid mutations, are also contemplated. *See, e.g.,* Marsic et al., *Molecular Therapy*, 22(11): 1900-1909 (2014). The nucleotide sequences of the genomes of various AAV serotypes are known in the art.

[0205] In some embodiments of the compositions and methods of the disclosure, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector comprises an inverted terminal repeat sequence or a capsid sequence that is isolated or derived from an AAV of serotype AAVrh.74, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 (AAVrh10), AAV11 or AAV12. In some embodiments, the AAV9 is a AAV9 variant. In some embodiment, the

AAV9 variant is a myoAAV. In one embodiment, the AAV vector comprises a modified capsid. In one embodiment the AAV vector is an AAV2-Tyr mutant vector. In one embodiment the AAV vector comprises a capsid with a non-tyrosine amino acid at a position that corresponds to a surface-exposed tyrosine residue in position Tyr252, Tyr272, Tyr275, Tyr281, Tyr508, Tyr612, Tyr704, Tyr720, Tyr730 or Tyr673 of wild-type AAV2. See also WO 2008/124724 incorporated herein in its entirety. In some embodiments, the AAV vector comprises an engineered capsid. AAV vectors comprising engineered capsids include without limitation, AAV2.7m8, AAV9.7m8, AAV2 2tYF, and AAV8 Y733F). In some embodiments, a capsid protein can comprise an LBV201 capsid protein as disclosed herein, or variants thereof. In some embodiments, the capsid protein can comprise a myoAAV capsid protein, or variants thereof.

[0206] In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant (rAAV). In some embodiments, the viral vector is self-complementary (scAAV). In some embodiments, the rAAV is a single-stranded AAV (ssAAV).

[0207] In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a non-viral vector. In some embodiments, the vector comprises or consists of a nanoparticle, a micelle, a liposome or lipoplex, a polymersome, a polyplex or a dendrimer. In some embodiments, the vector is an expression vector or recombinant expression system. As used herein, the term “recombinant expression system” refers to a genetic construct for the expression of certain genetic material formed by recombination.

[0208] In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, an expression control element which is other than or combined with the hybrid regulatory sequences disclosed herein. An “expression control element” as used herein refers to any sequence that regulates the expression of a coding sequence, such as a gene. Exemplary expression control elements include but are not limited to promoters, enhancers, microRNAs, post-transcriptional regulatory elements, polyadenylation signal sequences, and introns. Expression control elements may be constitutive, inducible, repressible, or tissue-specific, for example. A “promoter” is a control sequence that is a region of a polynucleotide sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. In some embodiments, expression control by a promoter is tissue-

specific. In some embodiments, expression control by a promoter is constitutive or ubiquitous. Non-limiting exemplary promoters include a Pol III promoter such as, e.g., U6 and H1 promoters and/or a Pol II promoter e.g., SV40, CMV (optionally including the CMV enhancer), RSV (Rous Sarcoma Virus LTR promoter (optionally including RSV enhancer), CBA (hybrid CMV enhancer/ chicken β -actin), CAG (hybrid CMV enhancer fused to chicken β -actin), truncated CAG, Cbh (hybrid CBA), EF-1a (human elongation factor alpha-1) or EFS (short intron-less EF-1 alpha), PGK (phosphoglycerol kinase), CEF (chicken embryo fibroblasts), UBC (ubiquitin C), GUSB (lysosomal enzyme beta-glucuronidase), UCOE (ubiquitous chromatin opening element), hAAT (alpha-1 antitrypsin), TBG (thyroxine binding globulin), Desmin (full-length, modified or truncated), MCK (muscle creatine kinase), or creatine kinase (CK), C5-12 (synthetic muscle promoter), CK8e (creatin kinase 8), NSE (neuron-specific enolase), Synapsin, Synapsin-1 (SYN-1), opsin, PDGF (platelet-derived growth factor), PDGF-A, MecP2 (methyl CpG-binding protein 2), CaMKII (Calcium/ Calmodulin-dependent protein kinase II), mGluR2 (metabotropic glutamate receptor 2), NFL (neurofilament light), NFH (neurofilament heavy), n β 2, PPE (rat preproenkephalin), ENK (preproenkephalin), Preproenkephalin-neurofilament chimeric promoter, EAAT2 (glutamate transporter), GFAP (glial fibrillary acidic protein), MBP (myelin basic protein), human rhodopsin kinase promoter (hGRK1), β -actin promoter, dihydrofolate reductase promoter, MHCK7 (hybrid promoter of enhancer/ promoter regions of muscle creatine kinase and alpha myosin heavy-chain genes) and combinations thereof. An “enhancer” is a region of DNA that can be bound by activating proteins to increase the likelihood or frequency of transcription. Non-limiting exemplary enhancers and posttranscriptional regulatory elements include the CMV enhancer, MCK enhancer, R-U5' segment in LTR of HTLV-1, SV40 enhancer, the intron sequence between exons 2 and 3 of rabbit β -globin, and Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE). In some embodiments an intron is used to enhance promoter activity such as a UBB intron. In some embodiments, the UBB intron is used with an EFS promoter.

[0209] In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, vector elements such as an IRES or 2A peptide sites for configuration of “multicistronic” or “polycistronic” or “bicistronic” or tricistronic” constructs, i.e., having double or triple or multiple coding areas or exons, and as such will have the capability to express from mRNA two or more proteins from a single construct. Multicistronic

vectors simultaneously express two or more separate proteins from the same mRNA. The two strategies most widely used for constructing multicistronic configurations are through the use of an IRES or a 2A self-cleaving site. An “IRES” refers to an internal ribosome entry site or portion thereof of viral, prokaryotic, or eukaryotic origin which are used within polycistronic vector constructs. In some embodiments, an IRES is an RNA element that allows for translation initiation in a cap-independent manner. The term “self-cleaving peptides” or “sequences encoding self-cleaving peptides” or “2A self-cleaving site” refer to linking sequences which are used within vector constructs to incorporate sites to promote ribosomal skipping and thus to generate two polypeptides from a single promoter, such self-cleaving peptides include without limitation, T2A, and P2A peptides or other sequences encoding the self-cleaving peptides.

[0210] Exemplary vector configurations comprise a promoter or regulatory sequence (promoter/enhancer combination) driving the expression of the nucleic acid encoding the CUG-targeting PUF. In another embodiment, a vector configuration comprises a promoter driving expression of the RNA-binding protein. In another embodiment, the vector configuration comprises a linker and one or more tags.

[0211] In some embodiments, the vector is a viral vector. In some embodiments, the vector is an adenoviral vector, an adeno-associated viral (AAV) vector, or a lentiviral vector. In some embodiments, the vector is a retroviral vector, an adenoviral/retroviral chimera vector, a herpes simplex viral I or II vector, a parvoviral vector, a reticuloendotheliosis viral vector, a polioviral vector, a papillomaviral vector, a vaccinia viral vector, or any hybrid or chimeric vector incorporating favorable aspects of two or more viral vectors. In some embodiments, the vector further comprises one or more expression control elements operably linked to the polynucleotide. In some embodiments, the vector further comprises one or more selectable markers. In some embodiments, the AAV vector has low toxicity. In some embodiments, the AAV vector does not incorporate into the host genome, thereby having a low probability of causing insertional mutagenesis. In some embodiments, the AAV vector can encode a range of total polynucleotides from 4.5 kb to 4.75 kb. In some embodiments, exemplary AAV vectors that may be used in any of the herein described compositions, systems, methods, and kits can include an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV2-Tyr mutant vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7

vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an AAV.rh43 vector, a modified AAV.rh43 vector, an AAV.rh64R1 vector, and a modified AAV.rh64R1 vector, an AAV-Tyr mutant vector, and any combinations or equivalents thereof. In some embodiments, the AAV vector is an AAV9 variant such as a myoAAV vector. In some embodiments, the myoAAV vector is a myoAAV1 or a myoAAV4 vector. In some embodiments, the AAV is a single stranded AAV (ssAAV) or a self-complementary AAV (scAAV). In some embodiments, the lentiviral vector is an integrase-competent lentiviral vector (ICLV). In some embodiments, the lentiviral vector can refer to the transgene plasmid vector as well as the transgene plasmid vector in conjunction with related plasmids (e.g., a packaging plasmid, a rev expressing plasmid, an envelope plasmid) as well as a lentiviral-based particle capable of introducing exogenous nucleic acid into a cell through a viral or viral-like entry mechanism. Lentiviral vectors are well-known in the art (see, e.g., Trono D. (2002) *Lentiviral vectors*, New York: Springer-Verlag Berlin Heidelberg and Durand et al. (2011) *Viruses* 3(2):132-159 doi: 10.3390/v3020132). In some embodiments, exemplary lentiviral vectors that may be used in any of the herein described compositions, systems, methods, and kits can include a human immunodeficiency virus (HIV) 1 vector, a modified human immunodeficiency virus (HIV) 1 vector, a human immunodeficiency virus (HIV) 2 vector, a modified human immunodeficiency virus (HIV) 2 vector, a sooty mangabey simian immunodeficiency virus (SIV_{SM}) vector, a modified sooty mangabey simian immunodeficiency virus (SIV_{SM}) vector, a African green monkey simian immunodeficiency virus (SIV_{AGM}) vector, a modified African green monkey simian immunodeficiency virus (SIV_{AGM}) vector, an equine infectious anemia virus (EIAV) vector, a modified equine infectious anemia virus (EIAV) vector, a feline immunodeficiency virus (FIV) vector, a modified feline immunodeficiency virus (FIV) vector, a Visna/maedi virus (VNV/VMV) vector, a modified Visna/maedi virus (VNV/VMV) vector, a caprine arthritis-encephalitis virus (CAEV) vector, a modified caprine arthritis-encephalitis virus (CAEV) vector, a bovine immunodeficiency virus (BIV), or a modified bovine immunodeficiency virus (BIV).

[0212] Nucleic Acids

[0213] Provided herein are the nucleic acid sequences encoding RNA-binding CUG repeat-targeting systems disclosed herein for use in gene transfer and expression techniques described herein. It should be understood, although not always explicitly stated that the sequences provided herein can be used to provide the expression product as well as

substantially identical sequences that produce a protein that has the same biological properties. These “biologically equivalent” or “biologically active” or “equivalent” polypeptides are encoded by equivalent polynucleotides as described herein. They may possess at least 60%, or alternatively, at least 65%, or alternatively, at least 70%, or alternatively, at least 75%, or alternatively, at least 80%, or alternatively at least 85%, or alternatively at least 90%, or alternatively at least 95% or alternatively at least 98%, identical primary amino acid sequence to the reference polypeptide when compared using sequence identity methods run under default conditions. Specific polypeptide sequences are provided as examples of particular embodiments. Modifications to the sequences to amino acids with alternate amino acids that have similar charge. Additionally, an equivalent polynucleotide is one that hybridizes under stringent conditions to the reference polynucleotide or its complement or in reference to a polypeptide, a polypeptide encoded by a polynucleotide that hybridizes to the reference encoding polynucleotide under stringent conditions or its complementary strand. Alternatively, an equivalent polypeptide or protein is one that is expressed from an equivalent polynucleotide.

[0214] The nucleic acid sequences (e.g., polynucleotide sequences) disclosed herein may be codon-optimized which is a technique well known in the art. Codon optimization refers to the fact that different cells differ in their usage of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. By altering the codons in the sequence to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. It is also possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in a particular cell type. Codon usage tables are known in the art for mammalian cells, as well as for a variety of other organisms. Based on the genetic code, nucleic acid sequences can be generated. In some embodiments, such a sequence is optimized for expression in a host or target cell or a cell in which the disclosed methods are practiced (such as in a mammalian cell, e.g., a human cell). Codon preferences and codon usage tables for a particular species can be used to engineer isolated nucleic acid molecules (such as one encoding a protein having at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type protein) that takes advantage of the codon usage preferences of that particular species. In one example, a nucleic acid sequence is optimized for expression in human cells, such as one having at least 70%, at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 98%, or at

least 99% sequence identity to its corresponding wild-type or originating nucleic acid sequence. In some embodiments, an isolated nucleic acid molecule (which can be part of a vector) includes at least one protein coding sequence that is codon optimized for expression in a eukaryotic cell, or at least one protein coding sequence codon optimized for expression in a human cell. In one embodiment, such a codon optimized coding sequence has at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating sequence. In another embodiment, a eukaryotic cell codon optimized nucleic acid sequence encodes a protein having at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating protein. In another embodiment, a variety of clones containing functionally equivalent nucleic acids may be routinely generated, such as nucleic acids which differ in sequence but which encode the same protein sequence. Silent mutations in the coding sequence result from the degeneracy (i.e., redundancy) of the genetic code, whereby more than one codon can encode the same amino acid residue. Thus, for example, leucine can be encoded by CTT, CTC, CTA, CTG, TTA, or TTG; serine can be encoded by TCT, TCC, TCA, TCG, AGT, or AGC; asparagine can be encoded by AAT or AAC; aspartic acid can be encoded by GAT or GAC; cysteine can be encoded by TGT or TGC; alanine can be encoded by GCT, GCC, GCA, or GCG; glutamine can be encoded by CAA or CAG; tyrosine can be encoded by TAT or TAC; and isoleucine can be encoded by ATT, ATC, or ATA. Tables showing the standard genetic code can be found in various sources (see, for example, Stryer, 1988, Biochemistry, 3rd Edition, W.H. Freeman and Co., NY).

[0215] “Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogsteen binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

[0216] Examples of stringent hybridization conditions include: incubation temperatures of about 25°C to about 37°C; hybridization buffer concentrations of about 6x SSC to about 10x SSC; formamide concentrations of about 0% to about 25%; and wash solutions from about 4x

SSC to about 8x SSC. Examples of moderate hybridization conditions include: incubation temperatures of about 40°C to about 50°C; buffer concentrations of about 9x SSC to about 2x SSC; formamide concentrations of about 30% to about 50%; and wash solutions of about 5x SSC to about 2x SSC. Examples of high stringency conditions include: incubation temperatures of about 55°C to about 68°C; buffer concentrations of about 1x SSC to about 0.1x SSC; formamide concentrations of about 55% to about 75%; and wash solutions of about 1x SSC, 0.1x SSC, or deionized water. In general, hybridization incubation times are from 5 minutes to 24 hours, with 1, 2, or more washing steps, and wash incubation times are about 1, 2, or 15 minutes. SSC is 0.15 M NaCl and 15 mM citrate buffer. It is understood that equivalents of SSC using other buffer systems can be employed.

[0217] “Homology” or “identity” or “similarity” refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An “unrelated” or “non-homologous” sequence shares less than 40% identity, or alternatively less than 25% identity, with one of the sequences of the present invention.

[0218] Cells

[0219] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a prokaryotic cell.

[0220] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is a bovine, murine, feline, equine, porcine, canine, simian, or human cell. In some embodiments, the cell is a non-human mammalian cell such as a non-human primate cell.

[0221] In some embodiments, a cell of the disclosure is a somatic cell. In some embodiments, a cell of the disclosure is a germline cell. In some embodiments, a germline cell of the disclosure is not a human cell.

[0222] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a stem cell. In some embodiments, a cell of the disclosure is an embryonic stem cell. In some embodiments, an embryonic stem cell of the disclosure is not a human cell. In some embodiments, a cell of the disclosure is a multipotent stem cell or a pluripotent

stem cell. In some embodiments, a cell of the disclosure is an adult stem cell. In some embodiments, a cell of the disclosure is an induced pluripotent stem cell (iPSC). In some embodiments, a cell of the disclosure is a hematopoietic stem cell (HSC).

[0223] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a muscle cell. In some embodiments, a muscle cell of the disclosure is a myoblast or a myocyte. In some embodiments, a muscle cell of the disclosure is a cardiac muscle cell, skeletal muscle cell or smooth muscle cell. In some embodiments, a muscle cell of the disclosure is a striated cell. In one embodiment, a cell or cells of a patient treated with compositions disclosed herein include, without limitation, skeletal muscle (developing and mature muscle fibers and satellite cells), neuromuscular junction, cardiomyocytes, smooth muscle cells, peripheral nervous system (neurons), peripheral motor neurons, and/or sensory neurons.

[0224] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a fibroblast or an epithelial cell. In some embodiments, an epithelial cell of the disclosure forms a squamous cell epithelium, a cuboidal cell epithelium, a columnar cell epithelium, a stratified cell epithelium, a pseudostratified columnar cell epithelium or a transitional cell epithelium. In some embodiments, an epithelial cell of the disclosure forms a gland including, but not limited to, a pineal gland, a thymus gland, a pituitary gland, a thyroid gland, an adrenal gland, an apocrine gland, a holocrine gland, a merocrine gland, a serous gland, a mucous gland and a sebaceous gland. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of an organ including, but not limited to, a lung, a spleen, a stomach, a pancreas, a bladder, an intestine, a kidney, a gallbladder, a liver, a larynx or a pharynx. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of a blood vessel or a vein.

[0225] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a primary cell.

[0226] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a cultured cell.

[0227] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is in vivo, in vitro, ex vivo or in situ.

[0228] In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is autologous or allogeneic.

[0229] *Methods of Use*

[0230] The disclosure provides a method of treating a disease or disorder comprising administering to a subject a therapeutically effective amount of a composition of the disclosure. In one embodiment, the disclosure provides a method of treating DM1. The disclosure provides a method of treating myotonic dystrophy type 1 (DM1) in a subject in need thereof comprising administering an rAAV vector according to any embodiment of the disclosure, an AAV viral vector of any embodiment of the disclosure, or a pharmaceutical composition of any embodiment of the disclosure to the subject.

[0231] In some embodiments, the subject has a CUG microsatellite repeat expansion (MRE) in a *DMPK* RNA sequence. In some embodiments, the rAAV vector, AAV viral vector, or pharmaceutical composition comprises a PUF polypeptide that binds the *DMPK* CUG MRE RNA sequence .

[0232] The disclosure provides a method of treating a DM1 in a patient in need of such treatment comprising administering to the patient a therapeutically effective amount of a composition of the disclosure, wherein the composition modifies, reduces, destroys, knocks down or ablates a level of expression of a toxic CUG repeat RNA (compared to the level of expression of a toxic CUG repeat RNA treated with a non-targeting (NT) control or compared to no treatment). In one embodiment, the level of reduction of the target toxic CUG repeat RNA or toxic repeats encoded by the target RNA is compared to the level of reduction of the target RNA or toxic repeats encoded by the target RNA when treated with an RCas9 system. In another embodiment, the level of reduction is 1-fold or greater. In another embodiment, the level of reduction is 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold. In another embodiment, the level of reduction is 10-fold or greater. In another embodiment, the level of reduction is between 10-fold and 20-fold. In another embodiment, the level of reduction is 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, or 20-fold. In another embodiment, the gene therapy compositions disclosed herein when administered to a DM1 patient leads to 20%-100% destruction (or elimination) of the toxic CUG repeat RNA. In one embodiment, the % elimination of the toxic CUG repeat RNA is any of 20-99%, 25%-99%, 50%-99%, 80%-99%, 90%-99%, 95%-99%. In one embodiment, the % elimination is 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. In another embodiment, % elimination is complete elimination or 100% elimination of the toxic CUG repeat RNA.

[0233] In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the patient to be treated includes, without limitation, a disease or disorder

related to CTG microsatellite repeat expansion expression. In some embodiments, the disease or disorder is related to CTG microsatellite repeat expansion in the 3' untranslated region of the *DMPK* gene. In some embodiments of the compositions and methods of the disclosure, a disease or disorder of the disclosure is myotonic dystrophy type 1 (DM1).

[0234] In some embodiments of the methods of the disclosure, a subject of the disclosure has been diagnosed with DM1. In some embodiments, the subject of the disclosure presents at least one sign or symptom of DM1. At least one DM1 sign or DM1 symptom includes, without limitation, myotonia, muscle atrophy, centralized myonuclei, muscle strength recovery, GI distress, cardiac conduction defects, swallowing difficulty, respiratory capacity. In one embodiment, at least one sign or symptom of DM1 is ameliorated by treatment with the compositions disclosed herein. In some embodiments, the subject has a biomarker predictive of a risk of developing DM1. In some embodiments, the biomarker is a genetic mutation.

[0235] In some embodiments of the methods of the disclosure, a subject of the disclosure is female. In some embodiments of the methods of the disclosure, a subject of the disclosure is male. In some embodiments, a subject of the disclosure has two XX or XY chromosomes. In some embodiments, a subject of the disclosure has two XX or XY chromosomes and a third chromosome, either an X or a Y.

[0236] In some embodiments of the methods of the disclosure, a subject of the disclosure is a neonate, an infant, a child, an adult, a senior adult, or an elderly adult. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or 31 days old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 months old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or any number of years or partial years in between of age.

[0237] In some embodiments of the methods of the disclosure, a subject of the disclosure is a mammal. In some embodiments, a subject of the disclosure is a non-human mammal.

[0238] In some embodiments of the methods of the disclosure, a subject of the disclosure is a human.

[0239] In some embodiments of the methods of the disclosure, a therapeutically effective amount comprises a single dose of a composition of the disclosure. In some embodiments, a

therapeutically effective amount comprises a therapeutically effective amount comprises at least one dose of a composition of the disclosure. In some embodiments, a therapeutically effective amount comprises a therapeutically effective amount comprises one or more dose(s) of a composition of the disclosure.

[0240] In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount reduces a severity of a sign or symptom of the disease or disorder.

[0241] In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates the disease or disorder.

[0242] In some embodiments of the methods of the disclosure, a therapeutically effective amount prevents an onset of a disease or disorder. In some embodiments, a therapeutically effective amount delays the onset of a disease or disorder. In some embodiments, a therapeutically effective amount reduces the severity of a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount improves a prognosis for the subject.

[0243] In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject intramuscularly. In some embodiments, the composition of the disclosure is administered to the subject by an intravenous route. In some embodiments, the composition of the disclosure is administered to the subject by an injection or an infusion. In some embodiments, the composition is administered systemically. In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject locally.

[0244] *Pharmaceutical Compositions*

[0245] In some embodiments, the compositions disclosed herein are formulated as pharmaceutical compositions. Briefly, pharmaceutical compositions for use as disclosed herein may comprise a protein(s) or a polynucleotide encoding the protein(s), optionally comprised in an AAV, which is optionally also immune orthogonal, in combination with one or more pharmaceutically or physiologically acceptable carriers, diluents or excipients. Such compositions may comprise buffers such as neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, mannose, sucrose or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (e.g., aluminum hydroxide); and preservatives.

Compositions of the disclosure may be formulated for routes of administration, such as e.g., oral, enteral, topical, transdermal, intranasal, and/or inhalation; and for routes of administration via injection or infusion such as, e.g., intravenous, intramuscular, subpial, intrathecal, intrastriatal, subcutaneous, intradermal, intraperitoneal, intratumoral, intravenous, intraocular, and/or parenteral administration. In certain embodiments, the compositions of the present disclosure are formulated for intravenous administration.

[0246] The disclosure provides a pharmaceutical composition comprising a hybrid regulatory sequence of any embodiment of the disclosure, an rAAV vector of any embodiment of the disclosure, or a viral vector of any embodiment of the disclosure. The disclosure provides a pharmaceutical composition for treating DM1 comprising an rAAV vector comprising a nucleic acid set forth in SEQ ID NO: 47, SEQ ID NO: 139, or SEQ ID NO: 150.

[0247] Exemplary Embodiments

[0248] Embodiment 1: A nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence, and a CK6 (muscle creatine kinase 6) enhancer-promoter sequence.

[0249] Embodiment 2: The nucleic acid molecule of embodiment 1, wherein the hybrid regulatory sequence further comprises an intron sequence.

[0250] Embodiment 3: The nucleic acid molecule of embodiment 2, wherein the intron sequence is an SV40 (simian virus 40) intron sequence or an SIE (short intronic enhancer) sequence set forth in any one of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, or SEQ ID NO: 35.

[0251] Embodiment 4: The nucleic acid molecule of embodiment 3, wherein the intron sequence is an SV40 intron sequence.

[0252] Embodiment 5: The nucleic acid molecule of embodiment 4, wherein the hybrid regulatory sequence comprises a nucleic acid sequence set forth in SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 26, or SEQ ID NO: 27.

[0253] Embodiment 6: The nucleic acid molecule of embodiment 1, wherein the hybrid regulatory sequence is in operable linkage with a PUF(CUG) sequence.

[0254] Embodiment 7: The nucleic acid molecule of embodiment 1, wherein the PUF(CUG) sequence is set forth in SEQ ID NO: 44, SEQ ID NO: 45, or SEQ ID NO: 46.

- [0255] Embodiment 8: An AAV viral vector comprising the nucleic acid molecule of embodiment 1.
- [0256] Embodiment 9: An AAV viral vector of embodiment 8, comprising an AAV capsid protein.
- [0257] Embodiment 10: The AAV viral vector of embodiment 9, wherein the AAV capsid protein is an AAV1 capsid protein, an AAV2 capsid protein, an AAV4 capsid protein, an AAV5 capsid protein, an AAV6 capsid protein, an AAV7 capsid protein, an AAV8 capsid protein, an AAV9 capsid protein, an AAV10 capsid protein, an AAV11 capsid protein, an AAV12 capsid protein, an AAV13 capsid protein, an AAVPHP.B capsid protein, an AAVrh74 capsid protein or an AAVrh.10 capsid protein.
- [0258] Embodiment 11: The AAV viral vector of embodiment 10, wherein the AAV capsid protein is an AAV9 capsid protein.
- [0259] Embodiment 12: The AAV viral vector of embodiment 11, wherein the AAV9 capsid protein is a AAV9 capsid variant.
- [0260] Embodiment 13: The AAV viral vector of embodiment 12, wherein the AAV9 capsid variant is myoAAV.
- [0261] Embodiment 14: The AAV viral vector of embodiment 9, wherein the AAV capsid is LBV201 comprising an amino acid sequence set forth in SEQ ID NO: 54.
- [0262] Embodiment 15: A cell comprising the vector of any one of the preceding embodiments.
- [0263] Embodiment 16: A composition for treating DM1 comprising the promoter of embodiment 1 in operable linkage with a transgene.
- [0264] Embodiment 17: The composition of embodiment 15, wherein the transgene expresses PUF(CUG).
- [0265] Embodiment 18: A composition for treating DM1 comprising an rAAV vector comprising a nucleic acid set forth in SEQ ID NO: 47.
- [0266] Embodiment 19: A method of treating myotonic dystrophy type 1 (DM1) in a mammal comprising administering a composition or AAV vector according to any one of embodiments 1-17 to a toxic target CUG microsatellite repeat expansion (MRE) RNA sequence in tissues of the mammal whereby the level of expression of the toxic target RNA is reduced.
- [0267] Embodiment 20: The method of embodiment 18, wherein the composition or AAV vector is administered to the subject intravenously or intramuscularly.

[0268] Embodiment 21: The method of embodiment 19, wherein the composition or AAV vector is administered to the subject intravenously.

[0269] Embodiment 22: The method of embodiment 18, wherein the reduced level of expression of the toxic target RNA thereby ameliorates symptoms of DM1 in the mammal.

[0270] Embodiment 23: The method of embodiment 18, wherein the level of expression of the toxic target RNA is reduced compared to the reduction in the level of expression of untreated toxic target CUG RNA.

[0271] Embodiment 24: The method of embodiment 18, wherein the level of reduction is between 1-fold and 20-fold.

[0272] Embodiment 25: A nucleic acid molecule comprising a muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a promoter sequence, an enhancer sequence, and an intron sequence, wherein the enhancer sequence is a MYLPP (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence, and wherein the promoter sequence is muscle creatine kinase (CK) or desmin.

[0273] Embodiment 26: The nucleic acid molecule of embodiment 24, wherein the CK promoter is selected from the group consisting of CK1, CK6, and CK7.

[0274] Embodiment 27: The nucleic acid molecule of embodiment 24, wherein the Desmin promoter is selected from the group consisting of full-length Desmin and modified Desmin.

[0275] Embodiment 28: The nucleic acid molecule of embodiment 24, wherein the intron is a SV40 intron or an SIE intron set forth in any one of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, or SEQ ID NO: 35.

[0276] Embodiment 29: The nucleic acid molecule of embodiment 24, wherein the muscle selective hybrid regulatory sequence is cardiac muscle selective.

[0277] Embodiment 30: The nucleic acid molecule of embodiment 24, wherein the muscle selective hybrid regulatory sequence is skeletal muscle selective.

[0278] Embodiment 31: An AAV capsid comprising LBV201 comprising an amino acid sequence set forth in SEQ ID NO: 54.

[0279] Embodiment 32: A nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a MYLPP (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK6 (muscle creatine kinase 6) enhancer-promoter sequence set forth in SEQ ID NO: 8, an SV40 intron set forth in SEQ ID NO: 7, and a kozak sequence set forth in SEQ ID NO: 11.

[0280] Embodiment 33: The nucleic acid molecule of embodiment 32, wherein the nucleic acid molecule comprises the nucleic acid sequence set forth in SEQ ID NO: 136.

[0281] Embodiment 34: A nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK1 (muscle creatine kinase 1) enhancer-promoter sequence set forth in SEQ ID NO: 9, an SV40 intron set forth in SEQ ID NO: 7 and a kozak sequence set forth in SEQ ID NO: 11.

[0282] Embodiment 35: The nucleic acid molecule of embodiment 32, wherein the nucleic acid molecule comprises the nucleic acid sequence set forth in SEQ ID NO: 137.

[0283] Embodiment 36: A nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises a MYLPF (myosin light chain, phosphorylatable, fast skeletal muscle gene) enhancer sequence set forth in SEQ ID NO: 2, a CK7 (muscle creatine kinase 7) enhancer-promoter sequence set forth in SEQ ID NO: 10, an SV40 intron set forth in SEQ ID NO: 7 and a kozak sequence set forth in SEQ ID NO: 11.

[0284] Embodiment 37: The nucleic acid molecule of embodiment 32, wherein the nucleic acid molecule comprises the nucleic acid sequence set forth in SEQ ID NO: 138.

[0285] Embodiment 38: A nucleic acid sequence comprising an hMYPLF enhancer as set forth in one or more of SEQ ID NO: 1-4 in operable linkage with a PUF(CUG) sequence.

[0286] ***Examples***

[0287] Example 1: Modified desmin-based promoters improve expression and efficacy in vitro

[0288] Design of in vitro testing of modified desmin-based promoters in human patient-derived cardiomyocytes was as follows:

[0289] A02239 is the standard desmin promoter control and A04949, A04959, and A04546 contain modified desmin-based promoters (FIG. 1). All AAVs had a MOI of 1e6.

[0290] RNA was harvested from transduced cardiomyocytes at 7 days post transduction (FIG. 1). The RNA expression of the PUF-CUG was measured by ddPCR. The modified desmin-based promoters increased expression between 5 to 16-fold (FIG. 2).

[0291] Protein was harvested from transduced cardiomyocytes at 7 days post transduction (FIG. 1). The protein expression of the PUF-CUG was measured by western blot. The

desmin promoter in A02239 drove minimal PUF expression while the modified desmin-based promoters, A04949 and A04546, produced robust PUF protein levels (FIG. 3).

[0292] RNA CUG repeats sequester MBNL proteins which results in abnormal splicing. The PUF protein competes and displaces MBNL from the CUG repeats, and the free MBNL is then able to properly modulate splicing. The RNA splicing from the human patient-derived cardiomyocytes was analyzed across 5 different transcripts. The splicing correction was much more potent from the modified desmin-based promoters. This demonstrates the promoters drove greater protein expression which increase efficacy in the human disease context (FIG. 4).

[0293] Example 2: Increased PUF expression and efficacy in vivo from novel hybrid regulatory sequences

[0294] AAV9 A02239 and AAV9 A04959 were intravenously injected into HSALR mice and PUF RNA expression was measured at 4 weeks post injection by ddPCR. The modified desmin-based promoter increased expression 3.2x in the gastrocnemius and 7.7x in the heart (FIG. 5).

[0295] AAV9 A02239 and AAV9 A05288 were intravenously injected into HSALR mice and PUF RNA expression was measured at 4 weeks post injection by ddPCR. The modified MYLPF enhancer/CK6-based promoter increased expression 10x in the gastrocnemius and decreased expression 5x in the heart compared to the standard desmin promoter. This gives a more balanced expression between cardiac and skeletal muscle after AAV9 intravenous delivery in mice (FIG. 6).

[0296] AAV9 A05288 were intravenously injected into HSALR mice and PUF RNA expression was measured at 8 weeks post injection by ddPCR. This data was compared to the previous 4 weeks post injection expression data. At 8 weeks post injection there was more PUF expression in both gastrocnemius and the heart (FIG. 7).

[0297] The body weights of mice intravenously injected with AAV9 A05288 at a dose of 1.2×10^{14} vg/kg were tracked for 8 weeks. There was no change in the body weights compared to the vehicle control, which indicates that this serotype promoter and transgene combination is safe and well tolerated (FIG. 8).

[0298] Functional myotonic dystrophy can be measured in mice by EMG studies of myotonia. In the gastrocnemius of the HSALR mice there is a high frequency of myotonic runs after stimulation consistent with the disease phenotype. However, the treatment with

AAV9 A05288 significantly decreased the frequency of myotonic runs, which indicates a functional correction of the disease (FIG. 9).

[0299] RNA was harvested from the gastrocnemius and quadricep muscles of mice intravenously injected with AAV9 A05288 at both 4- and 8-weeks post injection. There was a significant reduction in CUG repeat containing RNA from the HSA transgene at both time points with a slightly greater reduction at 8 weeks. This demonstrates a reduction of the pathological RNA transcript in this mouse model (FIG. 10).

[0300] RNA was harvested from the gastrocnemius and quadricep muscles of mice intravenously injected with AAV9 A05288 at 8-weeks post injection. The splicing of Atp2a1 exon 22 was analyzed because it is a reliable marker in the HSALR model. AAV9 A05288 corrected splicing in both skeletal muscles (FIG. 11).

[0301] Example 3: Further increased PUF expression and efficacy in vivo from novel capsid

[0302] AAV9 A05288 and LBV201 A05288 were intravenously injected into HSALR mice at 1.2×10^{14} vg/kg and PUF RNA expression was measured by ddPCR. The novel LBV201 capsid resulted in higher PUF expression in both the cardiac and skeletal muscle tissues at 4 weeks post injection compared to AAV9 A05288 at 8 weeks. This demonstrates LBV201 is a more effective capsid for transducing muscle than AAV9 (FIG. 12).

[0303] AAV9 A05288 and LBV201 A05288 were intravenously injected into HSALR mice at 1.2×10^{14} vg/kg and PUF protein expression was measured by western blot. The novel LBV201 capsid resulted in higher PUF expression in the gastrocnemius and quadricep 4 weeks post injection compared to AAV9 A05288 at 8 weeks. This further demonstrates LBV201 is a more effective capsid for transducing muscle than AAV9 (FIG. 13).

[0304] RNA was harvested from the gastrocnemius and quadricep muscles of mice intravenously injected with AAV9 A05288 and LBV201 A05288. The novel LBV201 capsid resulted in greater relative reduction of the pathological CUG repeat-containing transcript (FIG. 14).

[0305] RNA was harvested from the gastrocnemius and quadricep muscles of mice intravenously injected with AAV9 A05288 and LBV201 A05288. The splicing of Atp2a1 exon 22 was analyzed because it is a reliable readout of splicing deficits in the HSALR model. LBV201 A05288 had a splicing correction at 4 weeks that was equivalent to AAV9 A05288 at 8 weeks post treatment (FIG. 15).

[0306] Protein was harvested from transduced human skeletal myotubes at 7 days post transduction. While PUF protein from AAV9 A02239 was only detected from the 5E5 vg/cell dose there was robust PUF protein from the AAV9 A05288 and LBV201 A05288 at the 5x lower dose of 1E5 vg/cell. This indicates that novel capsid LBV201 can transduce human myotubes (FIG. 16).

[0307] Example 4: Novel hybrid regulatory sequences functional in myoAAV4 capsid context

[0308] RNA was harvested from the gastrocnemius, quadriceps, tibialis anterior, and triceps muscles of mice intravenously injected with MyoAAV4A A05288 4 weeks post injection at a dose of 8E13 vg/kg. There was a significant reduction in CUG repeat containing RNA in all the skeletal muscles analyzed (FIG. 17).

[0309] RNA was harvested from the gastrocnemius, quadriceps, tibialis anterior, and triceps muscles of mice intravenously injected with MyoAAV4A A05288 4 weeks post injection at a dose of 8E13 vg/kg. The splicing of Atp2a1 exon 22 was analyzed because it is a reliable readout of splicing deficits in the HSALR model. There was significant splicing correction in all muscles except the triceps which had a high baseline (non-pathologic) level Atp2a1 exon 22 inclusion (FIG. 18).

[0310] PUF expression at the protein level from both AAV9 and MyoAAV4A mediated delivery of A05288 was compared. The PUF protein expression was much greater in the MyoAAV4A samples despite both the dose and duration of treatment being lower in the MyoAAV4A treatments. The MyoAAV4 showed expectedly higher expression at the higher dose (FIG. 19).

[0311] Example 5: Reversal of DM1-associated splicing events with PUF construct in various AAV capsids

[0312] RNA-seq was performed on HSA-LR skeletal muscle after intravenous treatment with AAV9, MyoAAV4A, or LBV201 packaged MYLPF-CK6-SV40intron-PUFCUG (A05288). Alternative splicing events were evaluated with Olego (open source, <https://zhanglab.c2b2.columbia.edu/index.php/Olego>) and a composite score of splicing index (exon inclusion fraction) was determined for 19 known DM1-associated missplicing events. Robust correction of alternative splicing of these 19 exons was observed with A05288 treatment in various capsid as the composite index neared wildtype (WT levels). (See FIG. 20 and FIG. 21A-21B.)

OTHER EMBODIMENTS

[0313] While particular embodiments of the disclosure have been illustrated and described, various other changes and modifications can be made without departing from the spirit and scope of the disclosure. The scope of the appended claims includes all such changes and modifications that are within the scope of this disclosure.

What is claimed is:

1. A recombinant adeno-associate viral (rAAV) vector comprising:
 - a) a first inverted terminal repeat (ITR) sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 32;
 - b) an MYLPF enhancer sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 1;
 - c) a CK6 promoter sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 8;
 - d) an SV40 intron sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 35;
 - e) a nucleic acid sequence encoding a PUF polypeptide, wherein the PUF polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 135;
 - f) a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 38;
 - g) an SV40 polyA sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 39; and
 - h) a second ITR sequence comprising a nucleic acid sequence at least 95% identical to SEQ ID NO: 34.
2. The rAAV vector of claim 1, wherein the rAAV vector comprises:
 - a) a first inverted terminal repeat (ITR) sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 32;
 - b) an MYLPF enhancer sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 1;
 - c) a CK6 promoter sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 8;
 - d) an SV40 intron sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 35;

- e) a nucleic acid sequence encoding a PUF polypeptide, wherein the PUF polypeptide comprises an amino acid sequence at least 99% identical to SEQ ID NO: 135;
 - f) a Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 38;
 - g) an SV40 polyA sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 39; and
 - h) a second ITR sequence comprising a nucleic acid sequence at least 99% identical to SEQ ID NO: 34.
3. The rAAV vector of claim 1 or claim 2, wherein the nucleic acid sequence encoding the PUF polypeptide sequence comprises a nucleic acid sequence at least 95% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 45, preferably wherein the nucleic acid sequence encoding the PUF polypeptide sequence comprises a nucleic acid sequence at least 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NO: 45.
4. The rAAV vector of any one of the preceding claims, wherein the rAAV vector comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to the nucleic acid sequence set forth in SEQ ID NO: 47, SEQ ID NO: 139, or SEQ ID NO: 150.
5. The rAAV vector of any one of the preceding claims, wherein the rAAV vector further comprises an AAV capsid protein,
- i) preferably wherein the AAV capsid protein is an AAV1 capsid protein, an AAV2 capsid protein, an AAV4 capsid protein, an AAV5 capsid protein, an AAV6 capsid protein, an AAV7 capsid protein, an AAV8 capsid protein, an AAV9 capsid protein, an AAV10 capsid protein, an AAV11 capsid protein, an AAV12 capsid protein, an AAV13 capsid protein, an AAVPHP.B capsid protein, an AAVrh74 capsid protein, a myoAAV capsid protein or an AAVrh.10 capsid protein, preferably wherein the AAV capsid protein is a myoAAV capsid protein; or
 - ii) preferably wherein the AAV capsid protein comprises an amino acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 54, preferably wherein the AAV

capsid protein is encoded by a nucleic acid sequence that is at least 95%, preferably at least 99%, identical to SEQ ID NO: 53.

6. A nucleic acid molecule comprising a skeletal muscle selective hybrid regulatory sequence, wherein the hybrid regulatory sequence comprises:
- a) a myosin light chain, phosphorylatable, fast skeletal muscle gene (MYLPPF) enhancer sequence or a desmin enhancer sequence, wherein the MYLPPF enhancer sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, and SEQ ID NO: 4, wherein the desmin enhancer sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical any one of SEQ ID NO: 12 and SEQ ID NO: 14, and
 - b) a creatine kinase (CK) promoter sequence or a desmin promoter sequence, wherein the CK promoter sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, or SEQ ID NO: 17, wherein the desmin promoter sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 16 or SEQ ID NO: 22.
7. The nucleic acid molecule of claim 6, wherein the hybrid regulatory sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 25, SEQ ID NO: 26, or SEQ ID NO: 27.
8. The nucleic acid molecule of claim 6 or claim 7, wherein the hybrid regulatory sequence further comprises an intron sequence, preferably wherein the intron sequence is an SV40 (simian virus 40) intron sequence or an SIE (short intronic enhancer), preferably wherein the intron sequence comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, or SEQ ID NO: 35.

9. The nucleic acid of any one of claims 6-8, wherein the skeletal muscle selective hybrid regulatory sequence comprises:

- a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 1;
- a CK6 promoter sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 8;
- an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7 and SEQ ID NO: 35; and
- a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11,

preferably wherein the nucleic acid molecule comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 136.

10. The nucleic acid of any one of claims 6-8, wherein the skeletal muscle selective hybrid regulatory sequence comprises:

- a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 2;
- a CK1 promoter sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 9;
- an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7; and
- a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11,

preferably wherein the nucleic acid molecule comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 137.

11. The nucleic acid of any one of claims 6-8, wherein the skeletal muscle selective hybrid regulatory sequence comprises:

- a MYLPF enhancer sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 2;
- a CK7 promoter sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 10;

an SV40 intron comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 7; and
a kozak sequence comprising a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 11,
preferably wherein the nucleic acid molecule comprises a nucleic acid sequence at least 95%, preferably at least 99%, identical to SEQ ID NO: 138.

12. The nucleic acid molecule of claims 6-11, wherein the hybrid regulatory sequence is in operable linkage with a nucleic acid sequence encoding a PUF polypeptide, preferably wherein the PUF polypeptide binds a CUG-repeat RNA sequence, preferably wherein the CUG-repeat RNA sequence is a *DMPK* CUG-repeat sequence.

13. The nucleic acid molecule of claim 12, wherein the PUF polypeptide comprises an amino acid sequence at least 95%, preferably at least 99%, identical to the amino acid sequence of any one of SEQ ID NO: 133, SEQ ID NO: 134, and SEQ ID NO: 135, preferably wherein the PUF polypeptide is encoded by a nucleic acid sequence that is at least 95%, preferably at least 99%, identical to any one of SEQ ID NO: 44, SEQ ID NO: 45, and SEQ ID NO: 46.

14. An rAAV vector comprising the nucleic acid molecule of any one of claims 6-13.

15. A pharmaceutical composition comprising the rAAV vector of any one of claims 1-5 and 14 or the nucleic acid molecule of any one of claims 6-13.

16. A method of treating myotonic dystrophy type 1 (DM1) in a subject, the method comprising administering to the subject the pharmaceutical composition of claim 15.

17. The pharmaceutical composition of claim 15 for the treatment of DM1 in a subject.

18. The method of claim 16 or the pharmaceutical composition for use of claim 17, wherein the subject has CUG microsatellite repeat expansion (MRE) in a *DMPK* RNA sequence.

19. The method of claim 16 or claim 18 or the pharmaceutical composition for use of claim 17 or claim 18, wherein the pharmaceutical composition is administered to the subject intravenously or intramuscularly, preferably wherein the pharmaceutical composition is administered intravenously.